frontiers | Research Topics

Monitoring depression and anxiety symptoms: Scales and measurements

Edited by
Michael Noll-Hussong and Jan Ilhan Kizilhan

Published in
Frontiers in Psychiatry

ISSN 1664-8714
ISBN 978-2-8325-2756-6
DOI 10.3389/978-2-8325-2756-6

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Monitoring depression and anxiety symptoms: Scales and measurements

Topic editors

Michael Noll-Hussong — Saarland University, Germany

Jan Ilhan Kizilhan — University of Duhok, Iraq

Citation

Noll-Hussong, M., Kizilhan, J. I., eds. (2023). *Monitoring depression and anxiety symptoms: Scales and measurements*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-2756-6

Table of contents

frontiers in Psychiatry

ORIGINAL RESEARCH
published: 11 February 2021
doi: 10.3389/fpsyt.2021.589422

Psychometrics of the Spanish Version of the Screen for Adult Anxiety Related Disorders (SCAARED)

Sarah Sánchez-Cueva[1], *Yurena Alonso-Esteban*[1], *Patricio Sánchez-Cueva*[2], *Boris Birmaher*[3] *and Francisco Alcantud-Marín*[1]*

[1] Department of Developmental and Educational Psychology, University of Valencia, Valencia, Spain, [2] College of Medicine, Pennsylvania State University, Hershey, PA, United States, [3] Department of Psychiatry, University of Pittsburgh Medical Center, Pittsburgh, PA, United States

Objectives: To translate and validate the Screen for Adult Anxiety Related Disorders (SCAARED) questionnaire into Spanish.

Method: The original SCAARED was translated into Spanish and administered to a non-clinical sample of 131 university students (92.4% women, mean age 22 years) in Valencia, Spain. To assess the concurrent validity of the SCAARED, the Depression, Anxiety and Stress Scale−21(DASS) and the Beck's Anxiety Inventory (BAI) were also administered. Test-retest reliability was evaluated 2 weeks after the first administration.

Results: The internal consistency of SCAARED was high ($\alpha = 0.91$) and the stability of the measurement was also high (test-retest 0.81). The results of the Exploratory Factor Analysis showed four factors comparable to the original SCAARED (generalized anxiety disorder, social phobia disorder, panic disorder, and separation anxiety disorder). The Area Under the Curve was excellent (0.88).

Conclusions: The Spanish version of the SCAARED showed good psychometric properties comparable to the original SCAARED suggesting that it may be a useful instrument to screen for anxiety disorders in Spanish-speaking adult populations. Future studies are needed to replicate these findings in larger community and clinical samples.

Keywords: anxiety disorders, anxiety measures, rating scales, tools translation, measure of anxiety in adults

OPEN ACCESS

Edited by:
Yuan-Pang Wang,
University of São Paulo, Brazil

Reviewed by:
Jose Antonio López-Pina,
University of Murcia, Spain
S. M. Yasir Arafat,
Enam Medical College, Bangladesh

***Correspondence:**
Francisco Alcantud-Marín
francisco.alcantud@uv.es

Specialty section:
This article was submitted to Mood and Anxiety Disorders, a section of the journal Frontiers in Psychiatry

Received: *30 July 2020*
Accepted: *19 January 2021*
Published: *11 February 2021*

Citation:
Sánchez-Cueva S, Alonso-Esteban Y, Sánchez-Cueva P, Birmaher B and Alcantud-Marín F (2021) Psychometrics of the Spanish Version of the Screen for Adult Anxiety Related Disorders (SCAARED). Front. Psychiatry 12:589422. doi: 10.3389/fpsyt.2021.589422

INTRODUCTION

Anxiety disorders including generalized anxiety disorder, social anxiety disorder, panic disorder, agoraphobia and specific phobia, are among the most common psychiatric disorders in youths and adults with 4–25% of people suffering from one or more anxiety disorders in their lives (1, 2).

There has been a growing interest in research on anxiety disorders in the last decade, partly due to greater recognition of their burden and the impact of untreated illnesses (3). Results from a recent review and meta-analysis indicate that the majority of anxiety disorders tend to have an early onset, generally in childhood or early adolescence (4, 5), and endure over time if not properly treated. Anxiety disorders experienced before or during early adulthood have been associated with poor psychosocial functioning (e.g., work), poor health, low life satisfaction, and less social relationships during adulthood (6, 7). In addition, there is substantial evidence to suggest that individuals with anxiety disorders are at risk to develop substance abuse (5, 8); chronical medical illness (8); depressive disorders (9, 10); suicide-related behaviors or other risky behaviors (11).

Unfortunately, anxiety disorders may be unrecognized, particularly when is comorbid with other disorders such major depression, making treatment ineffective. The high prevalence of anxiety disorders among youth and adults and the resulting consequences recommend early detection to identify anxiety symptoms in these age groups. One of the factors that influence the under recognition of anxiety disorders is the limitations of current screening instruments (12, 13) in typically developing adult's populations (14) and among people with neurodevelopmental disorders (15). The use of structured (or semi-structured) interviews to evaluate anxiety disorders is the procedure of choice for establishing the diagnosis of an anxiety disorder (16), but is time-consuming and requires extensive training from either Primary Health Care professionals, clinical psychologists, and researchers. Consequently, self-report measures are the most common method of anxiety assessment (12, 17). Still, access to good screening instruments (with good levels of reliability, validity, and diagnostic discrimination), including formats for various informants (e.g., parents, teachers, self-reporters) and are affordable and adapted to Spanish-speaking populations, is often limited.

Currently, the most popular instruments for assessing anxiety in adults, include the State-Trait Anxiety Inventory (STAI) (18) and the Beck Anxiety Inventory (BAI) (19), two empirically and widely validated instruments used in psychological research and clinical practice in Spain (20, 21). Also, the Depression, Anxiety, and Stress Scale (DASS-21) (22), has shown promise in the screening for anxiety symptoms. Several other validated and reliable anxiety measures for specific anxiety disorders exist including, among others, the Generalized Anxiety Disorder Questionnaire and Generalized Anxiety Disorder-7 Scale (GAD-Q-IV & GAD-7) (23); the Social Phobia Inventory (SPIN) (24), the Liebowitz Social Anxiety Scale (LSAS) (25); and the Panic Disorder Severity Scale (PDSS) (26).

One of the limitations of the above scales is that they mainly assess one or two specific anxiety disorders. Although informative, this may be problematic because anxiety disorders usually are comorbid within themselves (27). Recently, a screen for all DSM-5 (28) was developed, the Screen for Adult Anxiety Related Emotional Disorders (SCAARED) (29). The SCAARED is a 44-item self-report instrument that was adapted from the youth instrument, the Screen for Children Anxiety Related Emotional Disorders (SCARED) (30, 31), a rating scale developed to screen for DSM anxiety disorders in youth (32). Numerous studies and meta-analysis have examined the psychometric properties of the SCARED, indicating good psychometric properties for children and adolescents from various countries and on different language adaptations (32–34). The factorial structure of SCAARED shows a correspondence with the SCARED including four factors that correspond to the respective diagnostic categories of DSM-5, including agoraphobia, panic disorder, generalized anxiety, social anxiety, and separation anxiety disorder (29). The SCAARED has excellent internal consistency (α by Cronbach = 0.97).

In addition to its good psychometric properties, as eluded before, in contrast to the other available rating scales for anxiety disorders in adults which usually only include one anxiety disorder, the SCAARED includes all DSM-5 anxiety disorders. Moreover, the fact that the SCAARED was derived from the SCARED and share similar factors, allows to compare the scores of the two instruments between adults and youth and follow up studies from childhood into adulthood.

Many of the adult anxiety questionnaires noted above have been translated to Spanish [e.g.,: DASS-21 (35) STAI (36)], but not the SCAARED. Thus, the goals for this study were to: (1) To translate the (SCAARED) into Spanish and validate it in a non-clinical sample to verify its factor structure and its reliability (internal consistency and stability of the measurement); (2) To examine the concurrent validity of the Spanish translation of the SCAARED with the Depression, Anxiety and Stress Scale (DASS-21) (37) and the Beck Anxiety Inventory (BAI) (19) and (3) To analyze the construct validity by means of a factorial analysis to check the stability of the original model.

METHOD

Procedure and Participants

The sample comprised 131 college students (92.4% female, mean age 22 years old, all Caucasian), recruited from the University of Valencia, Spain, using non-probability and convenience sampling. Participants were informed about the PICCA project [Programa de Intervención Cognitivo-Conductual en Ansiedad (Cognitive Behavioral Intervention Program for Anxiety)], requesting their collaboration on a voluntary basis. Prior to data collection, the purpose, procedures, and expectations of the study were described to all participants All third-year students of the Faculty of Education (specialty of Therapeutic Pedagogy and Hearing and Language) and of the Faculty of Psychology and Speech Therapy (specialty of speech therapy), completed the Spanish version of the SCAARED at the same time they completed the BAI and the DASS-21. The administration of the battery was carried out during the rest time between two classes at the beginning of the second term of the 2019–20 academic year (first week of February). Participants completed the questionnaires independently, although in a collective/group session carried out in the presence of one of the investigators.

Following the completion of the questionnaires, we contacted the students who showed significant anxiety symptoms and agreed to be re-contacted by e-mail to participate in the clinical interview for confirmation of the diagnosis and, if appropriate, participate in PICCA. Participants who agreed returned for assessments at time 2 (15 days later) for administration retest reliability and diagnostic interview. The clinical interview was performed by a specialized psychologist (co-author of the study), administered a subset of relevant International Neuropsychiatric Interview chapters (38). This study was approved by the University of Valencia's Human Research Ethics Committee (H1549280336722). Consent was obtained from participants in accordance with the University of Valencia's Human Research Ethics Committee.

Measures

The SCAARED (29) was translated, following the guidelines for translation and adaptation of Psychological Assessment

instruments (39). After consulting the author and obtaining his consent, the English version of the questionnaire was initially translated by a bilingual psychologist, who proposed a first translation of the items into Spanish. In some cases, small adaptations were made since the literal translation could be misleading. A second translator performed the same task, to

TABLE 1 | Descriptive statistics of the 44 items of SCAARED.

Items	Response categories			Mean	Standard Deviation	Range	Skewness	Kurtosis	Corrected item-test correlation
	0	1	2						
1	47	59	26	0.84	0.732	2	0.260	−1.090	0.435
2	49	65	17	0.76	0.669	2	0.327	−0.780	0.474
3	65	48	18	0.64	0.713	2	0.651	−0.794	0.471
4	91	31	9	0.37	0.612	2	1.416	0.920	0.333
5	58	51	22	0.73	0.734	2	0.484	−1.009	0.320
6	103	24	4	0.24	0.498	2	1.935	3.003	0.331
7	20	81	30	1.08	0.615	2	−0.045	−0.337	0.462
8	12	47	77	1.50	0.661	2	−0.961	−0.211	0.452
9	63	46	22	0.69	0.745	2	0.581	−0.983	0.353
10	56	52	23	0.75	0.737	2	0.439	−1.048	0.561
11	73	40	19	0.58	0.723	2	0.832	−0.630	0.408
12	91	33	7	0.36	0.583	2	1.400	0.974	0.448
13	111	14	6	0.20	0.503	2	2.549	5.659	0.143
14	62	46	23	0.70	0.751	2	0.553	−1.031	0.237
15	114	14	3	0.15	0.420	2	2.847	7.845	0.472
16	74	41	16	0.56	0.703	2	0.871	−0.503	0.112
17	100	27	4	0.27	0.509	2	1.752	2.263	0.550
18	26	58	47	1.16	0.732	2	−0.260	−1.090	0.328
19	66	48	17	0.63	0.705	2	0.677	−0.738	0.415
20	71	41	19	0.60	0.730	2	0.778	−0.728	0.331
21	4	43	84	1.61	0.549	2	−1.019	0.024	0.472
22	57	51	23	0.74	0.740	2	0.457	−1.048	0.328
23	11	52	68	1.44	0.646	2	−0.712	−0.501	0.547
24	9	55	67	1.44	0.622	2	−0.655	−0.513	0.397
25	71	45	15	0.57	0.691	2	0.801	−0.546	0.466
26	99	24	8	0.31	0.580	2	1.765	2.073	0.223
27	72	41	18	0.59	0.722	2	0.809	−0.656	0.525
28	78	36	17	0.53	0.716	2	0.964	−0.415	0.413
29	24	60	47	1.18	0.718	2	−0.274	−1.018	0.454
30	34	68	29	0.96	0.695	2	0.051	−0.904	0.279
31	19	59	53	1.26	0.697	2	−0.403	−0.887	0.605
32	66	32	33	0.75	0.835	2	0.504	−1.384	0.379
33	17	56	58	1.31	0.692	2	−0.505	−0.817	0.274
34	43	51	37	0.95	0.783	2	0.081	−1.359	0.363
35	12	59	60	1.37	0.647	2	−0.527	−0.653	0.401
36	85	35	11	0.44	0.646	2	10.203	0.293	0.231
37	14	51	66	1.40	0.676	2	−0.681	−0.624	0.529
38	79	38	14	0.50	0.684	2	1.013	−0.210	0.486
39	29	64	38	1.07	0.715	2	−0.101	−1.019	0.539
40	89	27	15	0.44	0.692	2	1.299	0.304	0.382
41	39	50	42	1.02	0.789	2	−0.041	−1.387	0.512
42	70	44	17	0.60	0.710	2	0.769	−0.658	0.570
43	41	58	32	0.93	0.746	2	0.112	−1.184	0.325
44	37	61	33	0.97	0.733	2	0.048	−1.122	0.451

later reach a consensus on the modifications, and thus be able to propose a single translation. Finally, a back-translation into English was made, which was evaluated by the author of the scale to judge the adjustment of the terms used. The final version of the SCAARED in Spanish is the object of this study (see **Supplementary Material 1**).

To evaluate the validity (criteria, content, construct) of the SCAARED in Spanish, two existing anxiety self-reports were also administered. The DASS-21 is a short version derived from the full 42-item self-report scale DASS (37), which measures negative emotional states (Anxiety, Stress and Depression) with a selection of 7 elements from each construct. The DASS-21 has validated versions in Spanish, reporting adequate psychometric properties in the general adult population (40), in university students (22, 41), and in the clinical population (35). The Spanish version of the DASS-21 self-report instrument was used for this study (22).

The BAI (19) is a 21-item self-report instrument that is used for measuring the typical symptoms of anxiety disorders. It is designed to assess the severity of anxiety symptoms and is widely used in both clinical and research settings. Each item refers to symptoms experienced in the last week and is answered with a 4-point severity scale. The total score on the scale ranges from 0 to 63 points. There is a Spanish version (19) with excellent psychometric performance both in university students (42), in the general population general (43), and especially in the clinical population (44, 45).

Data Analysis

The descriptive statistical analyses were performed with the SPSS (V26) licensed by the University of Valencia (Spain), the dimensionality analysis will be carried out with the software Mplus 8.3 (46). First, the descriptive statistical and the psychometric analysis of the SCAARED items was performed by calculating of mean, standard deviation, range, kurtosis, asymmetry, and corrected item-test correlation of all items on the scale and the internal consistency indicators (Cronbach's Alpha) and correlations with the other measures. The stability of the measurement (test-retest validity) was calculated with in a subsample of participants ($n = 19$) 15 days later. Receiver Operating Characteristic (ROC curve analysis) were used to determine SCAARED relative diagnostic accuracy. To analyze the structure of SCAARED a Confirmatory Factor Analysis (CFA) was developed on the original model (29). Additionally, we propose to perform an Exploratory Factorial Analysis (EFA) on the same data in order to venture a possible dimensionality different from the proposal in SCAARED original model.

RESULTS

The descriptive results of the 44 items of the SCAARED are shown in **Table 1** in the Supplementary Material of this article (see **Supplement S1**). The value of the correlation between each item and the test shows low to medium values (0.11–0.60) and the indices of asymmetry and kurtosis that the distribution of response frequencies in the three item alternatives (Likert Type) show non-normal behavior. For the present validation study of the SCAARED in Spanish, we will consider the values of the original scale (29).

The Cronbach's alpha for the total SCAARED scale was adequate ($\alpha = 0.91$), and very acceptable internal consistency for the items in each of the four dimensions of the scale (PA/SO $\alpha = 0.84$; GA $\alpha = 0.85$; SEP $\alpha = 0.62$; SOC $\alpha = 0.91$). As shown in **Table 2**, the test-retest correlations are high (<0.81) in all dimensions and in the total of the test. The *t*-tests for related samples show that the scores are stable 15 days after the first application.

TABLE 2 | Means, standard deviations, Pearson's r_{xy} and t-student comparing test and re-test results for each of the SCAARED dimensions.

SCAARED	First time		Second time				
	Mean	STD	Mean	STD	r_{xy}	*t*	*p*
Total	54.21	10.79	53.21	12.28	0.92	0.78	0.45
Panic disorder	16.28	7.31	16.21	6.50	0.95	0.11	0.91
Generalized anxiety disorder	22.36	2.95	21.50	3.82	0.87	1.71	0.11
Separation anxiety disorder	5.29	2.64	5.36	2.13	0.81	−0.17	0.86
Social anxiety disorder	10.29	3.29	10.00	3.39	0.86	0.34	0.74

TABLE 3 | Correlations between DASS-21 and BAI tests and SCAARED dimensions ($N = 131$).

SCAARED	DASS-21 stress	DASS-21 anxiety	DASS-21 depression	BAI
TOTAL	0.67	0.73	0.60	0.73
Panic disorder	0.58	0.71	0.50	0.68
Generalized anxiety disorder	0.66	0.57	0.53	0.60
Separation anxiety disorder	0.26	0.34	0.19	0.42
Social anxiety disorder	0.30	0.36	0.39	0.36

p-values ≤ 0.05.

TABLE 4 | The Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve and total fitting rates of the predictive model.

SCAARED	AUC	0.0, 0.10, 0.20, 0.30, 0.40	0.50, 0.60, 0.70, 0.80	0.90	1.0
TOTAL	0.88		0.81		
Panic disorder	0.85		0.75		
Generalized anxiety disorder	0.83		0.73		
Separation anxiety disorder	0.62		0.48		
Social anxiety disorder	0.72		0.59		

Regarding concurrent validity, the DASS-21 and BAI tests were applied simultaneously to the SCAARED. **Table 3** shows the correlation indices between the scores of the SCAARED dimensions and each of the tests used as criteria. Note that all correlations were significant (p-values ≤ 0.05) with the lowest being for Separation Anxiety.

The diagnostic value was assessed by taking as a criterion having reached the cut-off point in the DASS-21 and in the BAI tests. The ROC curve was examined for each of the SCAARED subscales and for the total score. As shown in **Table 4**, the values of the area under the curve and the fitting model. The AUC Index value of the 0.88 total test can be considered very adequate. This is the same as the predicted values of AG (0.83) and TP/S (0.85). However, TAS (0.62) and AS (0.72) constructs are not properly adjusted. When analyzing the ROC curve data, the diagnostic contrast criteria used must be taken into account.

A confirmatory factorial analysis with Mplus version 8.3 (46) was conducted on the original model of SCAARED (29). The values of RMSEA (0.085), the Tucker Lewis Index (TLI = 0.621) and the comparative fit index (CFI = 0.641) are close to the critical values in each case, the SRMR (0.112) although the χ^2 value is significant the final fit to the model ($\chi^2 = 3{,}545.71$; $p = 0.00$), point to an inadequate fit of the data to the original model.

Because the original structure was not confirmed, we carried out an EFA on the same sample to venture a possible dimensionality different from the proposal in the CFA. An EFA was performed on the 44 items and same sample. The Kaise-Meyer-Olkin index of sampling adequacy is 0.76 and Bartlett's sphericity test is significant ($\chi^2 = 1{,}464.1$; df = 861; $p = 0.0$), indicating that although the sample is small, we can proceed with the analysis (47).

The EFA was conducted to verify the four factor structure with Mplus. The values RMSEA (0.046), the TLI (0.92), CFI (0.93), SRMR (0.0866) and $\chi 2$ value (4,529.38; $p = 0.00$) indicate a good fit in the four-factor solution found and shown in **Table 5**. The first factor replicates the construct of Generalized Anxiety; the second factor rebuilds the construct of Panic Disorder. The third factor is defined by the items related to Social Anxiety and the fourth factor is defined by the items of Separation Anxiety.

DISCUSSION

The aims of this study were to translate the SCAARED questionnaire (29) into Spanish and evaluate its psychometric properties in a sample of 131 college students. Overall, the results from the Spanish version of the SCAARED indicated good internal consistency (Cronbach $\alpha > 0.90$), 2-week test-retest reliability (>0.86; $p = 0.001$), and adequate convergent validity with the DASS-21 and BAI. The results of the ROC analysis (AUC 0.88) inform us of excellent predictive value. The lowest correlation between the SCAARED and these instruments was with Separation Anxiety Disorder dimension because this disorder was only recently included in the DSM-5 as an adult anxiety disorder and as expected, except for the SCAARED, other anxiety self-reports do not include symptoms for this disorder.

The results of the Exploratory Factor Analysis showed four-factor structure (Generalized Anxiety, Social Anxiety, Panic Disorder/Significant Somatic Symptoms, and Separation Anxiety), which are consistent with the original SCAARED and correspond to the four factors reported for the instrument to screen for anxiety disorders in youth, the SCAARED (29). Moreover, the results of diagnostic validity, evaluated by means of the AUC indicators of the ROC curve were satisfactory. The above noted findings indicate that the Spanish version of the SCAARED behaves similarly to the English version and therefore appears to be an appropriate instrument for screening anxiety disorders in Spanish speaking adult populations. The fact that there are also Spanish versions of SCARED to screen youth for anxiety disorders (33, 34) with similar factorial structures, allows to use them as tools for evaluation of anxiety symptomatology in parents and their children and longitudinal studies of anxiety symptoms from childhood into adulthood.

Other adult anxiety measures available in Spanish are either dimensional (DASS-21; BAI; STAI, etc.) or specific to only one disorder (GAD-Q-IV, SPIN, LSAS, PDSS, etc.). In contrast, the SCAARED provides information on four types of anxiety disorders described in the DSM-5 and has excellent psychometric properties. In addition, it can be easily administered, is freely accessible, and time-effective (5–10 min). Finally, as noted above, the SCAARED can be crucial in obtaining and contrasting information from the patient/participant throughout life, since it bears similarities with the SCARED scale of which a Spanish version is already available (34).

Several limitations of our study are worth mention including a relatively small sample size, most of which were females and being a non-clinical sample which has impeded the calculation of some psychometric properties, such as inter-rater reliability and discriminatory validity. Consequently, further studies including larger samples and in clinical populations are needed. The present study consisted of the translation and adaptation of the scale and consequently, no qualitative

TABLE 5 | Factor Analysis for the four-factor solution (Saturations below 0.30 have been excluded).

N Item	Factor I Generalized anxiety disorder	Factor II Social phobia disorder	Factor III Panic disorder/significant somatic symptoms	Factor IV Separation anxiety disorder
21 I worry about things working out for me, [Le preocupa cómo le van a salir las cosas]	0.81			
08 It is hard for me to stop worrying, [Le cuesta dejar de preocuparse]	0.73			
23 I am a worrier, [Se preocupa demasiado]	0.68			
35 I worry about what is going to happen in the future, [Le preocupa de lo que vaya a pasar en el futuro]	0.66			
29 People tell me that I worry too much, [La gente le dice que se preocupa demasiado]	0.66			
37 I worry about how well I do things, [Se preocupa saber si está haciendo bien las cosas]	0.62	−0.31		
31 When I worry a lot, I feel restless, [Cuando se preocupa mucho, se siente inquieto(a)]	0.60			
07 I am nervous, [Estoy nervioso(a)]	0.53			
09 People tell me that I look nervous, [La gente me dice que parezco nervioso(a)]	0.50			
39 I worry about things that have already happened, [Me preocupo de las cosas que ya han sucedido]	0.46			
44 When I worry a lot, I feel irritable, [Cuando me preocupo mucho, me siento irritable]	0.46			
05 I worry about people liking me, [Me preocupa gustar a la gente]	0.35	−0.32		
24 When I worry a lot, I have trouble sleeping, [Cuando me preocupo mucho, tengo problemas para dormir]				
22 When I get anxious, I sweat a lot, [Cuando me siento ansioso(a), sudo mucho]				
34 I feel shy with people I don't know well, [Me siento tímido(a) con gente que no conozco bien]		−0.89		
27 It is hard for me to talk with people I don't know well, [Es difícil para mí hablar con gente que no conozco bien]		−0.85		
03 I don't like to be with people I don't know well, [No me gusta estar con personas que no conozco bien]		−0.83		
43 I am shy, [Soy tímido(a)]		−0.81		
10 I feel nervous with people I don't know well, [Me siento nervioso(a) con personas que no conozco bien]		−0.79		
42 I feel nervous when I go to parties, dances, or any place where there will be people that I don't know well, [Me siento nervioso(a) cuando voy a fiestas, bailes o cualquier lugar donde haya gente que no conozco bien]		−0.74		
41 I feel nervous when I am with other people and I have to do something while they watch me (for example: speak, play a sport), [Me siento nervioso(a) cuando estoy con otras personas y tengo que hacer algo mientras me miran (por ejemplo: hablar, hacer un deporte)]		−0.68		
17 I worry about going to work or school, or to public places, [Me preocupa ir al trabajo o a la universidad o instituto o a lugares públicos]		−0.46		
38 I am afraid to go outside or to crowded places by myself, [Tengo miedo de salir o ir a lugares concurridos solo(a)]		−0.44		0.38
14 I worry about being as good as other people, [Me preocupa ser tan bueno(a) como los demás]	0.34	−0.35		
01 When I feel nervous, It is hard for me to breathe, [Cuando me siento nervioso(a), me cuesta respirar]			0.70	
40 When I get anxious, I feel dizzy, [Cuando me pongo ansioso(a), me siento mareado(a)]			0.66	
06 When I get anxious, I feel like passing out, [Cuando me pongo ansioso(a), siento que voy a desmayarme]			0.66	
32 I am afraid of having anxiety (or panic) attacks, [Tengo miedo de tener ataques de ansiedad (o pánico)]			0.60	

(Continued)

TABLE 5 | Continued

N Item	Factor I Generalized anxiety disorder	Factor II Social phobia disorder	Factor III Panic disorder/significant somatic symptoms	Factor IV Separation anxiety disorder
19 I get shaky, [Me pongo tembloroso(a)]			0.56	
18 When I get anxious, my heart beats fast, [Cuando me siento ansioso(a), mi corazón late rápido]			0.55	
28 When I get anxious, I feel like I'm choking, [Cuando me siento ansioso(a), siento que me estoy ahogando]	0.36		0.52	
36 When I get anxious, I feel like throwing up, [Cuando me siento ansioso(a), tengo ganas de vomitar]			0.49	
12 When I get anxious, I feel like I'm going crazy, [Cuando me pongo ansioso(a), siento que me estoy volviendo loco(a)]			0.43	
15 When I get anxious, I feel like things are not real, [Cuando me pongo ansioso(a), siento que las cosas no son reales]			0.43	
02 I get headaches when I am at school, at work or in public places, [Tengo dolores de cabeza cuando estoy en la universidad, instituto, en el trabajo o en lugares públicos]			0.35	
25 I get really frightened for no reason at all, [Me asusto mucho sin ninguna razón]			0.31	
20 I have nightmares about something bad happening to me, [Tengo pesadillas sobre algo malo que me está pasando]				
26 I am afraid to be alone in the house, [Tengo miedo de estar solo(a) en la casa]				0.82
13 I worry about sleeping alone, [Me preocupa dormir solo(a)]				0.79
30 I don't like to be away from my family, [No me gusta estar lejos de mi familia]				0.50
04 I get nervous if I sleep away from home, [Me pongo nervioso(a) si duermo fuera de casa]				0.44
33 I worry that something bad might happen to my family, [Me preocupa que algo malo le pueda pasar a mi familia]				0.41
16 I have nightmares about something bad happening to my family, [Tengo pesadillas sobre algo malo que le pasa a mi familia]				0.40
11 I get stomachaches at school, at work, or in public places, [Me dan dolores del estómago en la universidad, instituto, en el trabajo o en lugares públicos]				0.32

Factor I: Items 5, 7, 8, 9, 14, 21, 23, 28, 29, 31, 35, 37, 39, 44.
Factor II: Items 3, 5, 10, 14, 17, 27, 34, 37, 38, 41, 42.
Factor III: Items 1, 2, 6, 12, 15, 18, 19, 25, 28, 32, 36, 40.
Factor IV: Items 4,11, 13,16, 26, 30, 33.

studies (e.g., discussion groups) have been conducted on the comprehensibility of items in the Spanish. Nevertheless, the authors carefully reread the items in both languages to assess comprehensibility and changes were incorporated when consensus indicated that a change improved the translation. Likewise, during the administration of the questionnaire, special attention was paid to the evaluation of the meaning of each element, without giving rise to significant questions or observations on the part of the people participating in the study.

In summary, similar to the English version of the SCAARED, the Spanish version showed good psychometric properties suggesting that it is a potential tool to screen for DSM-5 anxiety disorders in non-clinical adult populations. Further studies in large samples of clinical populations are necessary to evaluate its sensitivity and specificity as well as cut-off points to screen for anxiety disorders.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

This study was approved by the University of Valencia's Human Research Ethics Committee (H1549280336722). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2021.589422/full#supplementary-material

REFERENCES

1. Kessler RC, Angermeyer M, Anthony JC, de Graaf R, Koen D, Gasquet I, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry.* (2007) 6:168–76.
2. Polanczyk GV, Salum GA, Sagaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J Child Psychol Psychiatry.* (2015) 56:345–65. doi: 10.1111/jcpp.12381
3. Remes O, Brayne C, van der Linde R, Lafortune, L. A systematic review of reviews on the prevalence of anxiety disorders in adult populations. *Brain Behav.* (2016) 6:e00497. doi: 10.1002/brb3.497
4. Lijster JM, Dierckx B, Utens EM, Verhulst FC, Zieldorff C, Dieleman GC, et al. The age of onset of anxiety disorders: a meta-analysis. *Can J Psychiatry.* (2016) 62:237–46. doi: 10.1177/0706743716640757
5. Essau CA, Lewinsohn PM, Lim JX, Ho MR, Rohde P. Incidence, recurrence and comorbidity of anxiety disorders in four majordevelopmental stages. *J Affec Disord.* (2018) 228:248–53. doi: 10.1016/j.jad.2017.12.014
6. LeBlanc NJ, Mackenzie B, Aude, H. Anxiety disorders in emerging adulthood. In: Bui E, Charney M, Baker A, editors. *Clinical Handbook of Anxiety Disorders*. Switzerland: Springer Nature Switzerland AG (2020). p.157–73.
7. Essau CA, Lewinsohn PM, Olaya B, Seeley JR. Anxiety disorders in adolescents and psychosocial outcomes at age 30. *J Affec Disord.* (2014) 163:125–32. doi: 10.1016/j.jad.2013.12.033
8. Rogers AH, Wieman ST, Baker AW. Anxiety comorbidities: mood disorders, substance use disorders, and chronic medical illness. In: Bui E. Charney M, Baker A, editors. *Clinical Handbook of Anxiety Disorders*. Switzerland: Springer Nature Switzerland AG. (2020). p. 77–104.
9. Black JJ, Rofey DL. An overview of common psychiatric problems among adolescent and young adult females: focus on mood and anxiety. *Best Pract Res Clin Obstet and Gynaecol.* (2018) 48:165–73. doi: 10.1016/j.bpobgyn.2017.08.007
10. Boden JM, Fergusson DM, Horwood LJ. Anxiety disorders and suicidal behaviours in adolescence and young adulthood: findings from a longitudinal study. *Psychol Med.* (2007) 37:431–40. doi: 10.1017/S0033291706009147
11. Ahmadpoor J, Mohammadi Y, Soltanian AR, Poorolajal J. Psychiatric disorders and associated risky behaviors among Iranian university students: results from the Iranian PDABs survey. *J Public Health*. (2020), 1–8. doi: 10.1007/s10389-020-01229-8
12. Dennis RE, Boddington SJ, Funnell NJ. Self-report measures of anxiety: are they suitable for older adults? *Aging Mental Health.* (2007) 11:668–77. doi: 10.1080/13607860701529916
13. Iglesias García C, López García P, Ayuso Mateos JL, García JA, Bobes J.Screening for anxiety depression in Primary Care: utility of 2 brief scales adapted to the new ICD-11-PC. *Rev Psiquiatr Salud Ment.* (2020) 30. S1888-9891(20)30014-8. doi: 10.1016/j.rpsm.2019.12.001
14. Balsamo M, Cataldi F, Fairfield B. Assessment of anxiety in older adults: a review of self-report measures. *Clin Interv Aging.* (2018) 13:573–93. doi: 10.2147/CIA.S114100
15. Vasa RA, Keefer A, Reaven J, South M, White SW. Priorities for advancing research on youth with autism spectrum. *J Autism Dev Disord.* (2018) 48:925–34. doi: 10.1007/s10803-017-3320-0
16. Sandin B, Chorot P, Valiente, RM. *TCC de los Trastornos de Ansiedad: Innovaciones en niños y Adolescentes*. Madrid: Klinic (2016).
17. Simon E, Bögels SM. Screening for anxiety disorders in children. *Eur Child Adolesc Psychiatry.* (2009) 18:625–34. doi: 10.1007/s00787-009-0023-x
18. Spielberger C, Gorsuch R, Lushene R, Vagg P, Jacobs G. *Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press (1983).
19. Beck A, Steer R. *Beck Anxiety Inventory Manual*. San Antonio, TX: The Psychological Corporation (1993); Spanish version (2011).
20. Muñiz J, Fernández-Hermida JR. La opinión de los psicólogos españoles sobre el uso de los tests. *Papeles Psicól.* (2010) 31:108–21. Available online at: http://www.papelesdelpsicologo.es/pdf/1801.pdf
21. Sanz J. Recomendaciones para la utilización de la adaptación española del Inventario de Ansiedad de Beck (BAI) en la práctica clínica. *Clin Salud.* (2014) 25:39–48. doi: 10.1016/S1130-5274(14)70025-8
22. Bados A, Solanas A, Andrés R. Psychometric properties of the Spanish version of depression, anxiety and stress scales (DASS). *Psicothema.* (2005) 17:679–83. Available online at: http://www.psicothema.com/pdf/3165.pdf
23. García-Campayo J, Zamorano E, Ruiz MA, Pardo A, Pérez-Páramo M, López-Gómez V Freire, et al. Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. *Health Qual Life Outcomes.* (2010) 8:8. doi: 10.1186/1477-7525-8-8
24. García-López LG, Bermejo RM, HidalgoMD. The social phobia inventory: screening and cross-cultural validation in Spanish adolescents. *Span J Psychol.* (2010) 13:970–80. doi: 10.1017/S1138741600002614
25. Bobes J, Badía X, Luque A, García M, Paz González M, Dal-Ré R.Validación de las versiones en español de los cuestionarios Liebowitz Social Anxiety Scale, Social Anxiety and Distress Scale y Sheehan Disability Inventory para la evaluación de la fobia social. *Med Clin.* (1999) 112:530–8.
26. Santacana M, Fullana MA, Bonillo A, Morales M, Montoro M, Rosado S, et al. Psychometric properties of the Spanish sel-report version of the Panic Disorder Severity Scale. *Compr Psychiatry.* (2014) 55:1467–72. doi: 10.1016/j.comppsych.2014.04.007
27. Kendall P, Compton S, Walkup J, Birmaher B, Albano A, Sherrill J., et al. Clinical characteristics of anxiety disordered youth. *J Anxiety Disord.* (2010) 24:360–65. doi: 10.1016/j.janxdis.2010.01.009
28. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Arlington, VA: Asociación Americana de Psiquiatría (2013).
29. Angulo M, Rooks BT, Gill M, Goldstein T,Sakolsky, D, Goldstein, et al. Psychometrics of the screen for adult anxiety related disorders (SCAARED)-A new scale for the assessment of DSM-5 anxiety disorders. *Psychiatry Res.* (2017) 253:84–90. doi: 10.1016/j.psychres.2017.02.034
30. Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J, et al. The Screen for Child Anxiety Related Emotional Disorders (SCARED): scale construction and psychometric characteristics. *J Am Acad Child Adolesc Psychiatry.* (1997) 36:545–53. doi: 10.1097/00004583-199704000-00018
31. Birmaher B, Brent DA, Chiappetta L, Bridge J, Monga S, Baugher M. Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study. *J Am Acad Child Adolesc Psychiatry.* (1999) 38:1230–36. doi: 10.1097/00004583-199910000-00011
32. Hale WW 3rd., Crocetti E., Raaijmakers QA, Meeus WH. A meta-analysis of the cross-cultural psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED). *J Child Psychol Psychiatry.* (2011) 52:80–90. doi: 10.1111/j.1469-7610.2010.02285.x
33. Canals J, Hernández-Martínez C, Cosi S, Domènech E. Examination of a cutoff score for the Screen for Child Anxiety Related Emotional Disorders (SCARED) in a non-clinical Spanish population. *J Anxiety Disord.* (2012) 26:785–91. doi: 10.1016/j.janxdis.2012.07.008
34. Vigil-Colet A, Canals J, Cosi S, Lorenzo-Seva U, Ferrando PJ, Hernández-Martínez C, et al. The factorial structure of the 41-item version of the Screen for Child Anxiety Related Emotional Disorders (SCARED) in a Spanish population of 8 to 12 years-old. *Int J Clin Health Psychol.* (2009) 9:313–27. Available online at: https://www.redalyc.org/pdf/337/33712028009.pdf
35. Román F, Santibáñez P, Vinet EV. Uso de las Escalas de Depresión Ansiedad Estrés (DASS-21) como Instrumento de Tamizaje en

Jóvenes con Problemas Clínicos. *Acta de Investig Psicol.* (2016) 6:2325–36. doi: 10.1016/S2007-4719(16)30053-9
36. Buela-Casal G, Guillén-Riquelme A, Seisdedos Cubero N. *Cuestionario de Ansiedad Estado-Rasgo*. Octava edición. Madrid: TEA Ediciones (2011).
37. Lovibond P, Lovibond S. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the beck depression and anxiety inventories. *Behav Res Ther.* (1995) 33:335–43. doi: 10.1016/0005-7967(94)00075-U
38. Sheehan D, Janavs J, Baker R, Harnett-Sheehan K, Knapp E, Sheehan M, et al. *MINI International Neuropsychiatric Interview (5,0,0 version)*. Tampa, FL (2006). Available online at: http://www.iiap.es/files/mini.pdf (accessed January 12, 2020).
39. Muñiz J, Elosua P, Hambleton R. Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema.* (2013) 25:151–57. doi: 10.7334/psicothema2013.24
40. Daza P, Novy D, Stanley M, Averill P. The depression anxiety stress scale-21: Spanish translation and validation with a Hispanic sample. *J Psychopathol Behav Assess.* (2002) 24:195–205. doi: 10.1023/A:1016014818163
41. Fonseca-Pedrero E, Paino M, Lemos-Giráldez S, Muñiz, J. Psychometric properties of the Depression Anxiety and Stress Scales-21 (DASS-21) in Spanish college students. *Ansiedad Estrés.* (2010) 16:215–26.
42. Sanz J, Navarro ME. Propiedades psicométricas de una versión española del Inventario de Ansiedad de Beck (BAI) en estudiantes universitarios. *Ansiedad Estrés.* (2003) 9:59–84. Available online at: https://www.researchgate.net/publication/285908290_Propiedades_psicometricas_de_una_version_espanola_del_Inventario_de_Ansiedad_de_Beck_BAI_en_estudiantes_universitarios
43. Magán I, Sanz J, García-Vera MP. Psychometric properties of a Spanish version of the Beck Anxiety Inventory (BAI) in general population. *Span J Psychol.* (2008) 11:626–40. doi: 10.1017/S1138741600004637
44. Sanz J, García-Vera M, Fortún M. El "Inventario de Ansiedad de Berck" (BAI): propiedades psicométricas de la versión española en pacientes con trastornos psicológicos. *Psicol Conductual.* (2012) 20:563–83. Available online at: https://www.behavioralpsycho.com/wp-content/uploads/2019/08/05.Sanz_20-3oa.pdf
45. Vázquez-Morejón A, Vázquez-Morejón R, Bellido-Zanin BG. Beck anxiety inventory: psychometric characteristics in a sample from the clinical Spanish population. *Span J Psychol.* (2014) 17:E761 doi: 10.1017/sjp.2014.76 (accessed November 10, 2020).
46. Muthén LK, Muthén BO. *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén (2017). Available online at: http://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf (accessed November 10, 2020).
47. Cerny BA, Kaiser HF. A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate Behav Res.* (1977) 12:43–7. doi: 10.1207/s15327906mbr1201_3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychiatry

ORIGINAL RESEARCH
published: 26 May 2021
doi: 10.3389/fpsyt.2021.626859

Psychometric Properties of the Chinese Version of the 10-Item Ruminative Response Scale Among Undergraduates and Depressive Patients

Jiayue He[1,2,3†], *Yalin Liu*[1,2,3†], *Chang Cheng*[1,2,3], *Shulin Fang*[1,2,3], *Xiang Wang*[1,2,3*] and *Shuqiao Yao*[1,2,3*]

[1] Medical Psychological Center, The Second Xiangya Hospital, Central South University, Changsha, China, [2] Medical Psychological Institute of Central South University, Central South University, Changsha, China, [3] National Clinical Research Center for Mental Disorders, The Second Xiangya Hospital, Central South University, Changsha, China

OPEN ACCESS

Edited by:
Wissam El-Hage,
Université de Tours, France

Reviewed by:
Laurent Boyer,
AP-HM Assistance Publique des Hôpitaux de Marseille, France
Antoine Yrondi,
Centre Hospitalier Universitaire de Toulouse, France

***Correspondence:**
Shuqiao Yao
shuqiaoyao@csu.edu.cn
Xiang Wang
wang0916xia@gmail.com

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Mood and Anxiety Disorders, a section of the journal Frontiers in Psychiatry

Received: 07 November 2020
Accepted: 19 April 2021
Published: 26 May 2021

Citation:
He J, Liu Y, Cheng C, Fang S, Wang X and Yao S (2021) Psychometric Properties of the Chinese Version of the 10-Item Ruminative Response Scale Among Undergraduates and Depressive Patients. Front. Psychiatry 12:626859. doi: 10.3389/fpsyt.2021.626859

Objective: Rumination is considered as a key process in the mechanism of depression. Assessing rumination is important for both research and clinical practice. The Ruminative Response Scale (RRS) is a widely-used instrument to measure rumination. This study aimed to examine the psychometric properties of the Chinese 10-item Ruminative Response Scale (RRS-10) in a large sample of Chinese undergraduates and depressive patients.

Methods: A total of 1,773 university students and 286 clinical patients with major depressive disorder finished the Chinese version of the RRS10, State Trait Anxiety Inventory (STAI), and Beck Depression Inventory (BDI). A confirmatory factor analysis (CFA) was performed to examine the two-factor structure (reflection and brooding) of the RRS-10. The correlations among RRS-10, STAI, and BDI were explored in two samples. In addition, the measurement invariance of the RRS-10 across gender, time, and groups with and without depressive symptoms were further investigated. The internal consistency and test-retest reliability were also evaluated.

Results: Confirmatory Factor Analysis revealed that the two-factor structure of RRS-10 fitted reasonably both in undergraduates (CFI = 0.933, TLI = 0.905, RMSEA = 0.071, SRMR = 0.035) and depressive patients (CFI = 0.941, TLI = 0.910, RMSEA = 0.077, SRMR = 0.057). The results of the multi-group confirmatory factor analysis supported the full strict invariance across genders and across groups (undergraduates and depressive patients). The full strong invariance over time was also supported by MGCFA. Besides, the RRS-10 showed acceptable internal consistency and good stability.

Conclusions: The RRS-10 has good reliability and validity in different samples and over time, which demonstrated that RRS-10 is a valid measurement instrument to assess rumination.

Keywords: RRS-10, depression, rumination, factor structure, measurement invariance

INTRODUCTION

Rumination, defined as repetitive thoughts focusing on negative feelings and their causes, implications, and consequences is a method of coping with a negative mood. Individuals with a ruminative style of negative mood will focus their attention on their negative emotional state and ruminate about the causes of their depression and the features of its consequences so that they unable engage in some happy or neutral activities to get rid of their depression, thus prolonging the duration of depression. In 1987, Susan Nolen-Hoeksema first put forward the response style theory of depression (1). According to this theory, rumination is an important vulnerable factor for depression, which might aggravate and prolong depressive episodes (1–6). Researchers found that ruminative response could predict the severity of depression among clinical and non-clinical samples after 1 year (7). Longitudinal research has also revealed that even when controlling the most basic level of depression, rumination still has a significant effect on depressive symptoms (8–11). These results suggested that rumination is not inherently depressing, but can prolong an existing depressed mood.

On the basis of the response style theory, Nolen-Hoeksema et al. develop the Response Style Questionnaire (RSQ) to evaluate two different response styles of depression: rumination and distraction (12). The Ruminative Response Scale (RRS) was developed from the rumination subscale of the Response Style Questionnaire (RSQ). It has been revised over the years, leading to the current 22-item RRS. The RRS has shown good reliability (Cronbach $\alpha = 0.74$–0.92, $r_{test-retest} = 0.48$–0.76) and good validity in the USA (13, 14), Japan (8), Korea (15), Dutch (16), Brazil (17), France (18), and Spain (19), and in clinical and non-clinical samples (20). The Chinese version of the RRS has been reported to be useful for assessing rumination in a large undergraduate sample (21).

In a principal component analysis study, Roberts et al. (22) determined that the RRS was composed of three dimensions: symptom-based rumination, introspection/self-isolation, and self-blame. Bagby et al. came up with the two-factor structure of RRS among clinical patients, including symptom-focused rumination and self-focused rumination (23). Treynor and his Colleagues found some of the RRS items overlapped with depression scale constructs, and were thus classified as depressed-symptom rumination items (24). Thus, previous studies have removed 12 depression-related items from the RRS and found that the structure of scale, consisting of the remaining 10 items, had two 5-item factors: *brooding* and *reflection* (24). The brooding dimension of the RRS-10 refers to "mood pondering" (e.g., "Why do I have problems other people don't have?"), whereas the reflection dimension describes cognitive (as opposed to emotive) reassessment of past and present events, feelings, and behaviors (e.g., "Go someplace alone to think about your feelings"). The original RRS10 two-factor model (Factor 1: brooding; Factor 2: reflection) which was investigated by Treynor has been confirmed in several studies (8, 14, 15, 21). However, there still have been some inconsistencies surrounding the two-factor model. For example, Arana et al. explored the fact that the two-factor structure only retained eight of the original items (excluding item 2 and item 9) (25). A confirmatory study found that the original two-factor model was not confirmed among a community sample (17). A number of studies in recent years have confirmed that different types of rumination have different effects on depression: brooding is a risk factor which may prolong or exacerbate depression, while reflection is a protective factor which does not prolong depression (11, 26). Thus, it is important to determine whether the two-factor structure of rumination is consistent among levels of depression (16). But there has been a relative paucity of research examining the factor structure of the RRS10 in depressive patients.

Another issue that needs to be further investigated is whether RRS-10 has the same structure in different groups and whether items have the same meaning for individuals in different groups. In the research of ruminative response, the comparison of the level of rumination between different groups is usually carried out without the test of measurement invariance. However, to make the scale effective and interpretable between group comparisons, it is necessary to prove whether it has measurement invariance (27). That is to say, measurement invariance is a prerequisite for the comparison of differences between groups (28). Measurement invariance is defined as "a given factorial defined construct has the same measurement parameters across two or more samples (i.e., the loadings, intercepts and residual matrix are equal among different groups)" (29). Without evidence of measurement invariance, it cannot be concluded that group difference in rumination reflected true differences between groups, as the difference may be due to the item bias of the scale (30).

According to the ruminative response style theory of depression, women have been shown to be more likely to ruminate about the causes of their mood than men in the face of depression (31), and it has been suggested that this difference in response styles could explain, at least in part, the gender disparity in adult depression (22, 32–34). To ensure such inter-group difference is valid, and not reflecting an artifact of the instrument, it is necessary to confirm consistency of meaning for the scale's items between groups (35). Thus, to compare gender differences in rumination, it is essential that gender invariance of the scale should be established (36). Previous study has demonstrated that the measurement invariance of the RRS-10 was acceptable across genders among an undergraduate sample (14), but the result was not generalized to clinical populations.

Moreover, to make valid score comparisons over time, it is important to assess whether scale items measure the same construct over time, a property known as longitudinal invariance (37). Although changes in rumination over time have been routinely investigated, few studies have explored the longitudinal invariance of the RRS-10 up to now (38). However, without prior testing of longitudinal measurement invariance, it is not possible to determine whether the time changes observed in a structure are due to real changes or to changes in structure or structural measurements over time (27).

In summary, measurement invariance of the RRS10 across gender was supported in a previous study, but the result was not generalized to clinical populations, whereas little research has examined measurement invariance of the

RRS10 over time and between groups with and without depressive symptoms.

Thus, the aims of the present study were 3-fold. First, we tested the reliability of the RRS-10 in undergraduate and depressive patients. Second, we examined the two-factor model of the RRS-10 in the two samples. Third, we explored the measurement invariance of the RRS-10 across gender, time, and groups with and without depressive symptoms.

MATERIALS AND METHODS

Participants and Procedure

Undergraduate participants were recruited from two Chinese universities in Hunan Province. The scale was completed in the classroom and the data were collected by well-trained psychology researchers. Students who had a history of mental disorder, history of neurological disorder, or intellectual disability were excluded. A total of 1,872 university students were surveyed, 10 of which were excluded due to a history of mental disorder and 89 of which were excluded due to missing data. The final student sample includes 1,773 participants (95% completion rate) who returned valid questionnaires. The mean age of the final sample were 18.86 years [standard deviation (SD) = 1.22], including 853 men with a mean age of 18.69 years (SD = 1.14) and 920 women with a mean age of 19.02 years (SD = 1.27). For test-retest reliability analysis and longitudinal invariance, a subgroup of 633 participants (343 men, 54%; and 290 women, 46%) completed the RRS-10 again 2 months later. They had a mean age of 18.39 years (SD = 0.90).

The clinical sample which consisted of depressive patients was recruited from the psychological clinic of Second Xiangya Hospital. Diagnosis was conducted independently by two psychiatrists using the Structured Clinical Interview for the DSM-IV-TR Axis I Disorders-Patient Edition. All the patients met the depression standard of the DSM-IV-TR. Exclusion criteria were (1) prior DSM-IV-TR Axis I disorder (except depressive disorder); (2) history of alcohol/substance abuse; (3) diagnosed neurological disorder; (4) intellectual disability. A total of 286 depressive patients provided complete data, including 126 men (44.1%) and 160 women (55.9%). They had a mean age of 23.25 (SD = 6.62).

Participants were told that the information in these scales would not be disclosed to anyone outside of the research team and all participants provided informed consent. The Ethics Committee of the Second Xiangya Hospital of Central South University (Code: 025) approved the study. The study was unpaid, and all the participants volunteered to complete the study.

Measures

10-item Ruminative Response Scale (RRS-10)

The 10-item RRS, which was part of the larger Response Styles Questionnaire, is a self-rating scale designed to assess thoughts and behaviors when people feel depressed (12). It has two subscales (Brooding and Reflection) and its items are graded on a scale from 1 (never) to 4 (always), with higher scores indicating a greater rumination tendency. The original RRS10 has been demonstrated to have high internal reliability and good test-retest reliability (15, 19, 20, 24). The Chinese version of RRS-10 was translated from English into Chinese by two psychologists, and then it was translated back into English by a bilingual translator with repeated revisions to ensure translation accuracy. No questionnaire item was removed or altered significantly during translation.

State-Trait Anxiety Inventory (STAI)

The widely used self-reported STAI (39) consists of State Anxiety Inventory (SAI) and Trait Anxiety Inventory (TAI) components for measuring the distinct concepts of state and trait anxiety (e.g., I feel nervous; I worry too much over something that really doesn't matter). Each component scale has 20 items answered on a 1–4 scale, with higher score indicating more severe anxiety symptoms. The STAI has high internal consistency (Cronbach's α: state anxiety = 0.89–0.95; trait anxiety = 0.89–0.92) and good test-retest reliability (r ranging from 0.62 to 0.96 for state anxiety and ranging from 0.84 to 0.98 for state anxiety over periods of 2 to 4 weeks) (40–42). The Chinese version of STAI also has good internal reliability (Cronbach's α: state anxiety = 0.91; trait anxiety = 0.92) and test-retest reliability (r: 0.91 for state anxiety and 0.76 for trait anxiety over 2 weeks) (43). In the present study, the STAI had good internal consistency (Cronbach's α: state anxiety = 0.89; trait anxiety = 0.84).

Beck Depression Inventory (BDI)

The BDI is a multiple-choice self-reporting 21-item scale (44) used primarily to assess the presence and severity of depressive symptoms in the prior 2 weeks in clinical and non-clinical populations (e.g., guilty feelings; loss of pleasure). Each question is answered on a 0–3-point scale of intensity. The BDI total score range is from 0 to 63 points, with higher scores indicating more severe symptoms. The Chinese version of the BDI has good reliability since the Cronbach's α was 0.94 for clinical samples and 0.88–0.94 for non-clinical samples (45). The BDI also exhibited good internal consistency (Cronbach's α = 0.85) in the present study. We used the BDI to assess the convergent validity of the RRS-10 with respect to conceptualization of rumination in relation to depressive mood or depressive symptoms.

Data Analysis

To evaluate the reliability of the RRS-10, we calculated Cronbach's α values, mean inter-item correlations (MICs), split-half reliability, and test-retest reliability. Cronbach's $\alpha > 0.70$ (>0.60 in some cases) was considered acceptable. MICs in the range of 0.10–0.40 were considered optimal.

To evaluate validity, we examined STAI and BDI score relationships with RRS-10 scores and its subscale. These analyses were conducted in IBM SPSS 20.0. The starting hypothesis was that there is a strong, positive correlation among RRS-10, BDI, and STAI. Moreover, to examine whether depression and anxiety were predicted by demographic variables (gender and age) and rumination, we performed multiple linear regression both in the undergraduate sample and the clinical sample, with the BDI total score and STAI total score as dependent variables, respectively.

Confirmatory factor analysis (CFA) with Weighted Least Squares Estimation was employed to determine the goodness of fit of the two-factor structure model of the RRS-10 in undergraduates and depressive patients to establish well-fitting baseline model. Model fit was assessed based on the comparative fit index (CFI), Tucker-Lewis index (TLI), standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA) with a 90% confidence interval (CI). The acceptable fit criteria applied were: CFI $\geq$ 0.90, TLI $\geq$ 0.90, SRMR $\leq$ 0.08, and RMSEA $\leq$ 0.08 (28, 46).

Multigroup CFA (MGCFA) was then used to examine the measurement invariance of the RRS-10 across gender, time, and groups. Four aspects of measurement invariance were tested (47, 48). First, configural invariance (Model 1) was examined to test the consistency latent variable constitution across groups. Second, weak invariance (Model 2) was examined to probe inter-group consistency of factor loading. Third, strong invariance (Model 3) was examined to test whether the intercepts of observed variables were equal across groups. Fourth, strict invariance (Model 4) was examined to test whether error variance was consistent across groups. Measurement invariance was inferred from changes in CFI (ΔCFI), TLI (ΔTLI), and RMSEA (ΔRMSEA) with the following acceptability criteria: ΔCFI $\leq$ 0.010, ΔTLI $\leq$ 0.010, and ΔRMSEA $\leq$ 0.015 (49). CFA and MGCFA were completed in Mplus 7.4.

RESULTS

Reliability

Cronbach's α values, mean inter-item correlations, and split-half reliability for the RRS-10 were reported by sample in **Table 1**. In both samples, the Cronbach's α values were >0.8 for the whole scale and >0.7 for each dimension. All mean MICs were between 0.310 and 0.400. Split-half reliability was slightly higher in the clinical sample (0.744–0.763) than in the undergraduate sample (0.706–0.729). Good test-retest reliability for the full scale and each dimension were confirmed in the undergraduate sample (**Table 1**).

Convergent Validity

As shown in **Table 2**, Pearson analyses demonstrated very significant ($p < 0.01$) positive correlation coefficients among RRS-10 total scale, brooding subscale, reflection subscale, BDI, and STAI scores in both undergraduate sample and clinical sample. These direct correlations indicated good convergent validity of the RRS-10 with depression and anxiety scales.

We then examined the joint contribution of RRS-10 and demographic variables to predict depressive and anxiety scores through a series of multiple regression analyses. The results in **Table 3** showed that gender and brooding subscale scores were significant predictors of BDI total score in the undergraduate sample, whereas gender, age, brooding subscale score, and reflection subscale scores were significant predictors of STAI total score in the undergraduate sample. Only brooding subscale score was a significant predictor of BDI total score and of STAI total score in the clinical sample (**Table 3**).

CFA

All fit indices of the two-factor model reached our acceptability criteria in the undergraduate sample and clinical sample (Undergraduate sample: CFI = 0.933, TLI = 0.905, RMSEA = 0.071, SRMR = 0.035; Clinical sample: CFI = 0.941, TLI = 0.910, RMSEA = 0.077, SRMR = 0.057) (**Table 4**). Hence, CFA confirmation of the two-factor structure of the RRS-10 indicated that this model could be used as a baseline model for measurement invariance testing. The factor loadings of each item were shown in **Table 5** (The Chinese items were showed in **Supplementary Material**). All items factor loadings were 0.35 or greater.

TABLE 1 | Cronbach's α values, mean inter-item correlations, split-half reliability, and test-retest reliability of the RRS-10 and its two dimensions by sample.

	Undergraduate sample				Clinical sample		
	Cronbach's α	MIC	Split-half coefficient	Test-retest coefficient	Cronbach's α	MIC	Split-half coefficient
RRS-10	0.819	0.310	0.729	0.895	0.831	0.310	0.763
Brooding	0.719	0.337	0.706	0.660	0.709	0.400	0.747
Reflection	0.743	0.367	0.720	0.776	0.768	0.337	0.744

TABLE 2 | Correlations among STAI, BDI, and RRS-10 scores.

	Undergraduate sample				Clinical sample			
	Brooding	Reflection	RRS-10	STAI	Brooding	Reflection	RRS-10	STAI
Reflection	0.521**				0.593**			
RRS-10	0.849**	0.894**			0.848**	0.711**		
STAI	0.287**	0.112**	0.218**		0.435**	0.217**	0.640**	
BDI	0.399**	0.177**	0.285**	0.610**	0.424**	0.241**	0.638**	0.813**

**$p < 0.01$.

TABLE 3 | Multiple regression analyses with BDI total score (above) and STAI total score (below) as the dependent variable.

	Undergraduate sample				Clinical sample			
	ß	**95%CI**	***t***	***P***	**ß**	**95%CI**	***t***	***P***
BDI								
Gender	0.071	0.387, 1.678	3.137	**0.002**	0.091	−0.572, 5.357	1.589	0.113
Age	0.041	−0.006, 0.153	1.807	0.071	0.074	−0.078, 0.374	1.287	0.199
Brooding	0.325	0.796, 1.097	12.358	**< 0.001**	0.424	1.098, 2.187	5.942	**<0.001**
Reflection	0.006	−0.112, 0.142	0.236	0.813	0.003	−0.544, 0.566	0.040	0.968
	$F = 57.013$, $\boldsymbol{p}$ **<0.01**, $R^2 = 0.116$				$F = 15.190$, $\boldsymbol{p}$ **<0.01**, $R^2 = 0.195$			
STAI								
Gender	0.130	2.739, 5.652	5.651	**<0.001**	0.056	−2.694, 8.067	0.983	0.326
Age	0.074	0.115, 0.473	3.229	**0.001**	0.038	−0.255, 5.151	0.666	0.506
Brooding	0.313	1.652, 2.337	11.411	**<0.001**	0.470	2.302, 4.275	6.566	**<0.001**
Reflection	−0.059	−0.625, 0.030	−2.159	**0.031**	−0.062	−1.420, 0.544	−0.878	0.381
	$F = 50.250$, $\boldsymbol{p}$ **<0.01**, $R^2 = 0.106$				$F = 15.806$, $\boldsymbol{p}$ **< 0.01**, $R^2 = 0.198$			

Significant p-values are bolded.

TABLE 4 | Goodness of fit indexes for the two-factor model of the RRS-10.

	χ^2	df	CFI	TLI	SRMR	RMSEA
Undergraduate sample	4323.7	45	0.933	0.905	0.035	0.071
Clinical sample	77.264	30	0.941	0.910	0.057	0.077

χ^2, *Chi-square; df, degrees of freedom; CFI, Comparative Fit Index; SRMR, Standardized Root Mean Square Residual; RMSEA, Root Mean Square Error of Approximation.*

Measurement Invariance Across Gender

Based on the two-factor model of RRS-10, we proceeded with subsequent measurement invariance testing. Our first model specified configural invariance, meaning that the same factor structure was estimated for women and men; no inter-group constraints were placed on the parameter estimates. All goodness of fit indices obtained met the requirements of configural invariance (**Table 6**). Thus, configural invariance was established and the model was used as a baseline model for weak invariance (Model 6) analysis, in which factor loading was equalized across the groups. All goodness of fit indices met the requirements of weak invariance (ΔCFI = 0.001, ΔTLI = −0.008, and ΔRMSEA = 0.003). Thus, weak invariance was established between males and females, and strong invariance (Model 3) was examined with equal intercepts across genders. All requirements for goodness of fit indices for the strong invariance test were met (ΔCFI = 0.002, ΔTLI = −0.004, and ΔRMSEA = 0.002). Finally, for strict invariance testing (Model 4), the additional constraint of equal error variance across the two groups was added. All criteria for indices of strict invariance were met (ΔCFI = 0.007, ΔTLI = 0.000, and ΔRMSEA = 0.002), as shown in **Table 6**, establishing strict invariance for the undergraduate sample. Together, these results support the configural, metric, scalar, and strict invariance of the re-specified two-factor model of the RRS-10 across genders in our undergraduate sample.

In our clinical sample, the baseline models were deemed suitable for representing the data for males and for females (fit indices reported in **Table 6**), providing evidence in support of configural invariance. Furthermore, the changes observed in CFI (<0.010), TLI (<0.010), and RMSEA (<0.015) supported both weak equivalence and strong equivalence of the RRS-10 (**Table 6**). However, strict invariance was not supported since ΔCFI and ΔTLI >0.01 (ΔCFI = 0.022, ΔTLI = 0.014).

Measurement Invariance Across Time

Model fitting indexes for configural and weak invariance models met our measurement invariance requirements, and the changes in CFI, TLI, and RMSEA values supported weak and strong equivalence of the RRS-10 across the two testing time points (**Table 7**). Thus, the measurement invariance of the RRS-10 across time was established.

Measurement Invariance Across Clinical and Non-clinical Samples

The model fitting indexes for configural and weak invariance models met our measurement invariance requirements, and the changes in TLI, CFI and RMSEA values supported weak, strong, and strict equivalence of the RRS-10 across our non-clinical (undergraduate sample) and clinical samples (depressive patients) shown in **Table 7**. These results indicated that the RRS-10 exhibited good measurement invariance across clinical (i.e., depressive patients) and non-clinical samples.

DISCUSSION

The current study aimed to examine the reliability and validity of the Chinese RRS-10 in clinical and non-clinical samples. The high internal consistency and test-retest reliability values were

TABLE 5 | The factor loadings of each item in RRS-10.

RRS-10 item	Undergraduate sample	Clinical sample
Think "What am I doing to deserve this?"	0.401	0.460
Analyze recent events to try to understand why you are depressed.	0.664	0.686
Think "Why do I always react this way?"	0.601	0.369
Go away by yourself and think about why you feel this way.	0.513	0.729
Write down what you are thinking and analyze it.	0.565	0.723
Think about a recent situation, wishing it had gone better.	0.574	0.585
Think "Why do I have problems other people don't have?"	0.682	0.724
Think "Why can't I handle things better?"	0.503	0.550
Analyze your personality to try to understand why you are depressed.	0.625	0.743
Go someplace alone to think about your feelings.	0.643	0.530

obtained in two samples. Then, a CFA was conducted, supporting a similar two-factor structure as that established in previous studies (8, 14, 15, 21, 24). Measurement invariance of the Chinese RRS-10 were well-established across gender, time, and groups with and without depressive symptoms. To our knowledge, this was the first research to explore the measurement invariance across times and different groups (clinical and non-clinical). The present results showed excellent reliability and validity of the RRS-10 in the clinical and non-clinical groups.

For the reliability analysis of the RRS-10, all Cronbach's α coefficients, in both the undergraduate sample and clinical sample, reached acceptable standards ($\alpha > 0.70$). These results were consistent with previous studies, which reported the internal reliability from 0.74 to 0.92 (8, 13–19). All the mean inter-item coefficients were between 0.10 and 0.40 both in the undergraduate sample and clinical sample and the high test-retest values also indicated good reliability of the RRS-10.

According to the convergent validity, moderate but positive correlations were found between rumination (RRS-10 and its subscales) and psychiatric symptoms (depression and anxiety) in clinical and non-clinical groups, which demonstrate that individuals who had more ruminative thinking seemed to have greater depressive and anxiety symptoms. Previous research has shown a strong link between rumination and psychiatric illness, especially depression (50, 51). Lam et al. (52) found that, in a non-clinical group, RRS scores predicted a more predominant ruminative response style. A clinical research also found a strong association between the RRS score and the duration and severity of depressive episodes in patients with depression.

Furthermore, in multiple regression analyses, we found that the brooding subscale of the RRS-10 was a significant predictor of depression and anxiety symptoms in both undergraduates and clinical samples. The concepts of brooding and reflection in the context of the RRS represent two different types of rumination, with the former encompassing repeated, passive attending to one's own negative emotions and evaluating one's own status and goals harshly. Reflection, on the other hand, involves one's efforts to solve problems and, thereby, alleviate his or her symptoms of depression and anxiety. Brooding is associated with increased negativity bias and negative coping styles, while reflective rumination has a less clear relationship with negative outcomes (53). A meta-analysis indicated that brooding had a moderate effect size for suicidal ideation and history of suicide attempt, but reflection was only associated with suicidal ideation (54). Ricarte et al. also found that anxiety and brooding were positively correlated even after controlling for depression scores (55). Meanwhile, our findings that reflection subscores were associated with STAI scores, but not BDI scores, suggests that reflection may play an important role in anxiety disorders rather than depressive illnesses (56, 57).

Our CFA confirmation of the goodness of fit of the two-factor structure of the Chinese RRS-10 in our undergraduate and clinical samples were consistent with previous RRS factor analysis studies demonstrating a well-stabilized two-factor structure (19). Based on this result, we were comfortable employing the two-factor structure of the RRS-10 as a baseline model for examining measurement invariance.

Importantly, researchers' ability to compare groups in a valid way is dependent upon measurement invariance (30). The present study examined the measurement invariance of the RRS-10 across gender, time, and groups (clinical and non-clinical). Our MGCFA confirmed good morphological, weak, strong, and strict invariance of the Chinese RRS-10 across gender in undergraduate samples, which was consistent with previous studies (14, 21). The configural invariance was supported, which indicated that rumination was conceptualized similarly in women and men which was reflected by two factors measuring brooding and reflection. Besides, there was support for weak invariance, which means that the units of measurement are equal in men and women, that is, the items and potential factors of the scale have the same meaning in men and women (58). Moreover, the present establishment of strong invariance indicated that inter-gender group differences in scores could be interpreted as reflecting true group differences in latent variables, which provided the same reference point between men and women. Intergroup comparisons were meaningful only if the units and reference points are the same. Therefore, it is the premise to compare the latent mean that the weak equivalence and the strong equivalence are satisfied (58). Finally, the strict invariance was supported in women and men which reflected the cross-group difference of latent variable variation. In clinical samples, the strict equivalence was not supported. But the residual equivalence is the most strict equivalence limit and it is not necessary for most research (59). In summary, the results of this study confirmed that the Chinese RRS-10 has strong equivalence, indicating that the scale is effective and interpretable between gender groups.

Regarding measurement invariance over time, our results supported the conclusion that the Chinese RRS-10 had configural, weak, and strong invariance between an initial

TABLE 6 | Measurement invariance of the RRS-10 across gender.

Model	χ^2	df	CFI	TLI	SRMR	RMSEA	ΔCFI	ΔTLI	ΔRMSEA
Undergraduate sample									
Model 1	215.461	60	0.951	0.926	0.037	0.053	–	–	–
Model 2	224.761	68	0.950	0.934	0.039	0.050	0.001	−0.008	0.003
Model 3	241.059	76	0.948	0.938	0.040	0.048	0.002	−0.004	0.002
Model 4	272.072	86	0.941	0.938	0.045	0.048	0.007	0.000	0.002
Clinical sample									
Model 1	88.136	53	0.949	0.913	0.057	0.069	–	–	–
Model 2	97.831	61	0.946	0.921	0.063	0.066	0.003	−0.008	0.003
Model 3	111.083	69	0.939	0.920	0.068	0.066	0.007	0.001	0.000
Model 4	136.978	80	0.917	0.906	0.071	0.072	0.022	0.014	−0.006

Model 1: morphological invariance; Model 2: metric invariance; Model 3: strong invariance; Model 4: strict invariance.

TABLE 7 | Measurement invariance of the RRS-10 across time and across samples with and without depressive symptoms.

Model	χ2	*df*	CFI	TLI	SRMR	RMSEA	ΔCFI	ΔTLI	ΔRMSEA
Measurement invariance across time									
Model 1	226.929	64	0.932	0.904	0.045	0.063	–	–	–
Model 2	255.836	74	0.924	0.907	0.048	0.062	0.008	−0.003	0.001
Model 3	279.870	82	0.917	0.909	0.050	0.062	0.007	−0.002	0.000
Model 4	279.068	90	0.921	0.921	0.052	0.058	−0.004	−0.012	0.004
Measurement invariance across samples with and without depressive symptoms									
Model 1	340.632	64	0.939	0.914	0.041	0.065	–	–	–
Model 2	372.675	72	0.934	0.917	0.049	0.064	0.005	−0.003	0.001
Model 3	392.447	74	0.930	0.915	0.056	0.065	0.004	0.002	−0.001
Model 4	385.503	68	0.930	0.907	0.059	0.068	0.000	0.008	−0.003

Model 1: morphological invariance; Model 2: metric invariance; Model 3: strong invariance; Model 4: strict invariance.

test and a re-test 2 months later, at least for general population individuals. This confirmation of longitudinal invariance indicated that researchers could be confident that changes in RRS-10 scores over time reflect real changes in rumination over time, rather than an artifact produced by composition instability. Because longitudinal invariance was assessed over a relatively short 2-month time interval, it is impossible to draw conclusions about the stability and structural invariance of RRS-10 over much longer intervals, such as several years or decades. Longer-term research is needed to further verify the longitudinal invariance of RRS-10 over longer periods of time.

The present research also supports the conclusion that the RRS-10 has configural, weak, strong, and strict measurement invariance between non-clinical (undergraduates) and clinical (depressive) samples. These results indicate that the form of latent variables in the RRS-10 is consistent between healthy adults and depressive patients, with equivalent factor loading, intercepts, and error variances of each item. This establishment of scale equivalence allows the inferences that the RRS-10 has the same reference point between clinical and non-clinical populations, and that the relationship between the scale's observation indicators and potential individual characteristics have the same meaning between general population and depressive groups.

Several limitations of our study should be acknowledged. First, the data were obtained principally from self-report measures, which are by nature subjective. Second, we did not control for socioeconomic and demographic variables (e.g., family income, religion, social relationships), which are associated with ruminative response and could affect the results of the RRS-10. Third, the samples only included undergraduate and depressive patients, thus limiting the generalizability of the results. Fourth, the level of rumination was different across cultures, thus, the measurement invariance of the RRS-10 across different cultures could be tested in the future.

CONCLUSION

The RRS-10 has good psychometric characteristics and measurement invariance across gender, time, and populations with and without depressive symptoms. The present results support the conclusion that the RRS-10 is a valid and reliable self-reported instrument for examining rumination, in relation to depressed mood, in Chinese adults and in patients with depressive symptoms.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Second Xiangya Hospital of Central South University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

SY supervised the study. JH and YL performed, collected, and wrote the paper. CC and SF contributed to the analysis. XW revised the paper. All co-authors revised and approved the version to be published.

FUNDING

This research was supported by grants from the National Science and Technology Project for Professional Basic Research (Grant Number: 2015FY111600), the National Natural Science Foundation (Grant Number: 81370034), and the Fundamental Research Funds for the Central Universities of Central South University (Grant Number: 2020zzts284).

ACKNOWLEDGMENTS

We would like to express our gratitude to all those who helped us during the writing of this thesis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2021.626859/full#supplementary-material

REFERENCES

1. Nolen-Hoeksema S. Sex differences in unipolar depression: evidence and theory. *Psychol Bull.* (1987) 101:259–82. doi: 10.1037/0033-2909.101.2.259
2. Nolen-Hoeksema S, Wisco BE, Lyubomirsky S. Rethiking rumination. *Perspect Psychol Sci.* (2008) 3:400–24. doi: 10.1111/j.1745-6924.2008.00088.x
3. Watkins ER, Roberts H. Reflecting on rumination: consequences, causes, mechanisms and treatment of rumination. *Behav Res Ther.* (2020) 127:103573. doi: 10.1016/j.brat.2020.103573
4. Nolen-Hoeksema S, Stice E, Wade E, Bohon C. Reciprocal relations between rumination and bulimic, substance abuse and depressive symptoms in female adolescents. *J Abnorm psychol.* (2007) 116:195–207. doi: 10.1037/0021-843X.116.1.198
5. Watkins ER. Constructive and unconstructive repetitive thoughts. *Psychol Bull.* (2008) 134:163–206. doi: 10.1037/0033-2909.134.2.163
6. Abela JRZ, Parkinson C, Stolow D, Starrs C. A test of the integration of the hopelessness and response styles theories of depression in middle adolescence. *J Clin Child Adolesc.* (2009) 38:354–64. doi: 10.1080/15374410902851630
7. Nolen-Hoeksema S. The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *J Abnorm Psychol.* (2000) 109:504–11. doi: 10.1037/0021-843X.109.3.504
8. Hasegawa A, Koda M, Hattori Y, Kondo T, Kawaguchi J. Longitudinal prediction of the brooding and reflection subscales of the Japanese ruminative response sacle for depression. *Psychol Rep.* (2013) 113:566–85. doi: 10.2466/02.15.PR0.113x24z5
9. Hankin BL. Rumination and depression in adolescents: investigating symptom specificity in a multi-wave prospective study. *J Clin Child Adolesc.* (2008) 37:701–13. doi: 10.1080/15374410802359627
10. Abela JRZ, Brozina K, Haigh EP. An examination of the response styles theory of depression in third-and seventh-grade children: a short-term longitudinal study. *J Abnorm Child psych.* (2002) 30:515–27. doi: 10.1023/A:1019873015594
11. Roelofs J, Rood L, Meesters C, te Dorsthorst V, Bogels S, Alloy LB, et al. The influence of rumination and distraction on depressed and anxious mood: a Prospective examination of the response styles theory in children and adolescents. *Eur Child Adolesc Psychiatr.* (2009) 18:635–42. doi: 10.1007/s00787-009-0026-7
12. Nolen-Hoeksema S. Responses to depression and their effects on the duration of depressive episodes. *J Abnorm Psychol.* (1991) 100:569–82. doi: 10.1037/0021-843X.100.4.569
13. Cowdrey FA, Park RJ. Assessing rumination in eating disorders: pricipal component analysis of a minimally modified ruminative response scale. *Eating Behaviors.* (2011) 12:321–4. doi: 10.1016/j.eatbeh.2011.08.001
14. Whisman MA, Miranda R, Fresco DM, Heimberg RG, Jeglic EL, Weinstock LM. Measurement invariance of the ruminative response scale across gender. *Assessment.* (2020) 27:508–17. doi: 10.1177/1073191118774131
15. Lee S, Kim W. Cross-cultural adaptation, reliability, and validity of the revised Korean version of ruminative response scale. *Psychiatry Investig.* (2014) 11:59–64. doi: 10.4306/pi.2014.11.1.59
16. Griffith JW, Raes F. Factor structure of the ruminative response scale A community-sample study. *Eur J Psychol Assess.* (2015) 31:247–53. doi: 10.1027/1015-5759/a000231
17. Lucena-Santos P, Pinto-Gouyeia J, Carvalho SA, Oliveira MD. Is the widely used two-factor structure of the ruminative response scale invariant across different samples of woman? *Psychol Psychother.* (2018) 91:398–416. doi: 10.1111/papt.12168
18. Parola N, Zendjidjian XY, Alessandrini M, Baumstarck K, Loundou A, Fond G, et al. Psychometric properties of the ruminative response scale-short form in a clinical sample of patients with major depressive disorder. *Patient Prefere Adher.* (2017) 11:929–37. doi: 10.2147/PPA.S125730
19. Extremera N, Fernandez-Berrocal P. Validity and reliability of Spanish versions of the ruminative responses scale-short form and the distraction responses scale in a sample of Spanish high school and college students. *Psychol Rep.* (2006) 98:141–50. doi: 10.2466/pr0.98.1.141-150
20. Schoofs H, Hermans D, Raes F. Brooding and reflection as subtypes of rumination: evidence from confirmatory factor analysis in nonclinical samples using the Dutch ruminative response scale. *J Psychopathol Behav Assess.* (2010) 32:609–17. doi: 10.1007/s10862-010-9182-9
21. Lei XX, Zhong MT, Liu Y, Xi C, Zhu XZ, Yao SQ, et al. Psychometric properties of the 10-item ruminative response scale in Chinese university students. *BMC Psychiatry.* (2017) 17:152. doi: 10.1186/s12888-017-1318-y
22. Roberts JE, Gilboa E, Gotlib IH. Ruminative response style and vulnerability to episodes of dysphoria: gender, neuroticism, and episode duration. *Cogn Ther Res.* (1998) 22:401–23. doi: 10.1023/A:1018713313894

23. Bagby RM, Parker JDA. Relation of rumination and distraction with neuroticism and extraversion in a sample of patients with major depression. *Cogn Ther Res.* (2001) 25:91–102. doi: 10.1023/A:1026430900363
24. Treynor W, Gonzalez R, Nolen-Hoeksema. S. Rumination reconsidered: a psychometric analysis. *Cogn Ther Res.* (2003) 27:247–59. doi: 10.1023/A:1023910315561
25. Arana FG, Rice KG. Cross-cultural validity of the ruminative response scale in Argentina and the States. *Assessment.* (2020) 27:309–20. doi: 10.1177/1073191117729204
26. Lo CSL, Ho SMY, Hollon SD. The effeets of rumination and negative cognitive styles on depression: a mediation analysis. *Behav Res Ther.* (2008) 46:487–95. doi: 10.1016/j.brat.2008.01.013
27. Jovanovic V. Measurement invariance of the Serbian version of the satisfaction with life scale across age, gender and time. *Eur J Psychol Assess.* (2019) 35:555–63. doi: 10.1027/1015-5759/a000410
28. He JY, Zhong X, Gao YD, Xiong G, Yao SQ. Psychometric properties of the Chinese version of the childhood trauma questionnaire short form (CTQ-SF) among undergraduates and depressive patients. *Child Abuse Neglect.* (2019) 91:102–8. doi: 10.1016/j.chiabu.2019.03.009
29. Little TD. Mean and covariance structures (MACS) analyses of cross-cultural data: practical and theoretical issues. *Multivar Behav Res.* (1997) 32:53–76. doi: 10.1207/s15327906mbr3201_3
30. Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ Res Methods.* (2000) 3:4–70. doi: 10.1177/109442810031002
31. Nolen-Hoeksema S, Parker L, Larson J. Ruminative coping with depressed mood following loss. *J Personal Soc Psychol.* (1994) 67:92–104. doi: 10.1037/0022-3514.67.1.92
32. Zetsche U, Ehring T, Ehlers A. The effects of rumination on mood and intrusive memories after exposure to traumatic material: an experimental study. *J Behav Ther Exp Psychiatr.* (2009) 40:499–514. doi: 10.1016/j.jbtep.2009.07.001
33. Yoder JD, Lawrence C.L. Are gender differences in self-reported rumination explained by women's stereotyping? *Sex Roles.* (2010) 65:94–101. doi: 10.1007/s11199-010-9913-0
34. Wupperman P, Neumann CS. Depressive symptoms as a function of sex-role, rumination, and neuroticism. *Pers Indiv Differ.* (2006) 40:189–201. doi: 10.1016/j.paid.2005.05.017
35. Byrne BM, Stewart SA. The MACS approach to testing for multigroup invariance of a second-order structure: a walk through the process. *Struct Equ Model.* (2006) 13:287–321. doi: 10.1207/s15328007sem1302_7
36. Liu JD, Chung PK, Chen WP. Constraints of recreational sport participation: measurement invariance and latent mean differences across sex and physical activity status. *Percept Mot Skills.* (2014) 119:363–76. doi: 10.2466/06.03.PMS.119c24z0
37. Van de Schoot R, Lugtig P, Hox J. A checklist for testing measurement invariance. *Eur J Dev Psychol.* (2012) 9:486–92. doi: 10.1080/17405629.2012.686740
38. Lo BCY, Ng TK, So T. Parental demandingness predicts adolescents' rumination and depressive symptoms in a one-year longitudinal study. *J Abnorm Child Psych.* (2020). doi: 10.1007/s10802-020-00710-y. [Epub ahead of print].
39. Bados A, Gomez-Benito J, Balaguer G. The state-trait anxiety inventory, trait version: does it really measure anxiety? *J Pers Assess.* (2010) 92:560–7. doi: 10.1080/00223891.2010.513295
40. Spielberger CD. *Manual for the State-Trait Anxiety Inventory (Form Y)*. Menlo Park, CA: Mind Garden (1983).
41. Kvaal K, Ulstein I, Nordhus IH, Engedal K. The Spielberger State-Trait Anxiety Inventory (STAI): the state scale in detecting mental disorders in geriatric patients. *Int J Geriatr Psych.* (2005) 20:629–34. doi: 10.1002/gps.1330
42. Fountoulakis KN, Papadopoulou M, Kleanthous S, Papadopoulou A, Bizeli V, Nimatoudis L, et al. Reliability and psychometric properties of the Greek translation of the StateTrait Anxiety Inventory form Y: preliminary data. *Ann Gen Psychiatr.* (2006) 5:2. doi: 10.1186/1744-859X-5-2
43. Ma WF, Liu YC, Chen YF, Lane HY, Lai TJ, Huang LC. Evaluation of psychometric properties of the Chinese mandarin version state-trait anxiety inventory Y form in Taiwanese outpatients with anxiety disorders. *J Psychiatr Ment Hlt.* (2013) 20:499–507. doi: 10.1111/j.1365-2850.2012.01945.x
44. Beck AT, Brown GK, Steer RA. Manual of beck depression inventory-II. *Am Univ Washington DC.* (1996) 21:88. doi: 10.1037/t00742-000
45. Wu PC, Huang TW. Gender-related invariance of the Beck Depression Inventory II for Taiwanese adolescent samples. *Assessment.* (2014) 21:218–26. doi: 10.1177/1073191112441243
46. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J.* (1999) 6:1–55. doi: 10.1080/10705519909540118
47. Meredith W. Measurement invariance, factor analysis and factorial invariance. *Psychometrika.* (1993) 58:525–43. doi: 10.1007/BF02294825
48. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc.* (1995) 90:773–95. doi: 10.1080/01621459.1995.10476572
49. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling.* (2002) 9:233–55. doi: 10.1207/S15328007SEM0902_5
50. Sukhodolsky DG, Golub A, Cromwell EN. Development and validation of the anger rumination scale. *Pers Indiv Differ.* (2001) 31:689–700. doi: 10.1016/S0191-8869(00)00171-9
51. Broderick PC, Korteland C. Coping styles and depression in early adolescence: relationships to gender, gender role, implicit beliefs. *Sex Role.* (2002) 46:201–13. doi: 10.1023/A:1019946714220
52. Lam D, Smith N, Checkley S, Rijsdijk F, Sham P. Effect of neuroticism, response style and information processing on depression severity in a clinically depressed sample. *Psychol Med.* (2003) 33:469–79. doi: 10.1017/S0033291702007304
53. Satyshur MD, Layden EA, Gowins JR, Buchanan A, Gollan JK. Functional connectivity of reflective and brooding rumination in depresses and healthy women. *Cogn Affect Behav Ne.* (2018) 18:884–901. doi: 10.3758/s13415-018-0611-7
54. Rogers ML, Joiner TE. Rumination, suicidal ideation, and suicide attempts: a meta-analytic review. *Rev Gen Psychol.* (2017) 21:132–42. doi: 10.1037/gpr0000101
55. Ricarte JJ, Ros L, Latorre JM, Munoz MD, Aguilar MJ, Hernandez JV. Role of anxiety and brooding in specificity of autobiographical recall. *Scand J Psychol.* (2016) 57:495–500. doi: 10.1111/sjop.12323
56. Hankin BL. Development of sex differences in depressive and co-occurring anxious symptoms during adolescence: descriptive trajectories and potential explanations in a multiwave prospective study. *J Clin Child Adolesc.* (2009) 38:460–72. doi: 10.1080/15374410902976288
57. Field AP, Morgan J. Post-event processing and the retrieval of autobiographical memories in socially anxious individuals. *J Affect Disorders.* (2004) 18:647–63. doi: 10.1016/j.janxdis.2003.08.004
58. Schmitt N, Kuljanin G. Measurement invariance: review of practice and implications. *Hum Resour Manage R.* (2008) 18:210–22. doi: 10.1016/j.hrmr.2008.03.003
59. Xi C, Zhong MT, Lei XX, Liu Y, Ling Y, Zhu XZ, et al. Psychometric properties of the Chinese version of the neuroticism subscale of the NEO-PI. *Front Psychol.* (2018) 9:1–10 doi: 10.3389/fpsyg.2018.01454

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychiatry

ORIGINAL RESEARCH
published: 12 August 2021
doi: 10.3389/fpsyt.2021.688154

Nine Forward–Backward Translations of the Hopkins Symptom Checklist-25 With Cultural Checks

Patrice Nabbe[1], Jean Yves Le Reste[1], Morgane Guillou-Landreat[2], Radost Assenova[3], Djurdjica Kasuba Lazic[4], Slawomir Czachowski[5], Stanislava Stojanović-Špehar[4], Melida Hasanagic[6], Heidrun Lingner[7], Ana Clavería[8], María Rodríguez-Barragán[9,10,11], Agnieszka Sowinska[12,13], Stella Argyriadou[14], Charileos Lygidakis[15], Bernard Le Floch[1], Tristan Montier[16,17], Harm Van Marwijk[18] and Paul Van Royen[16,19]*

[1] Department of General Practice, EA 7479 Soins primaires, Santé publique, Registre des cancers de Bretagne Occidentale, Université de Bretagne Occidentale, Brest, France, [2] Department of Addictology, EA 7479 Soins primaires, Santé publique, Registre des cancers de Bretagne Occidentale, Université de Bretagne Occidentale, Brest, France, [3] Department of Urology and General Medicine, Faculty of Medicine, Medical University of Plovdiv, Plovdiv, Bulgaria, [4] Department of Family Medicine "Andrija Stampar," School of Public Health, School of Medicine, University of Zagreb, Zagreb, Croatia, [5] Department of Clinical Psychology and Neuropsychology, Nicolaus Copernicus University, Torun, Poland, [6] Health Care Studies, University "Djemal Bijedic," Mostar, Bosnia and Herzegovina, [7] Centre for Public Health and Healthcare, Hannover Medical School, Hanover, Germany, [8] Xerencia Xestión Integrada de Vigo, Servizo Galego de Saúde, Instituto de Investigación Sanitaria Galicia-Sur, Red de Investigación en Actividades Preventivas y de Promoción de la Salud, Vigo, Spain, [9] Centro de Atención Primaria La Mina, Gerencia Territorial de Atención Primaria de Barcelona, Instituto Catalán de la Salud, Sant Adrià de Besòs, Barcelona, Spain, [10] Fundación Instituto Universitario de Investigación en Atención Primaria de Salud Jordi Gol i Gurina (IDIAP Jordi Gol), Barcelona, Spain, [11] Universidad Autónoma de Barcelona, Barcelona, Spain, [12] Department of Experimental Linguistics, Nicolaus Copernicus University, Torun, Poland, [13] Escuela de Inglés, Universidad Catolica del Norte, Antofagasta, Chile, [14] The Greek Association of General Practitioners (ELEGEIA), Thessaloniki, Greece, [15] Department of Behavioral and Cognitive Sciences, University of Luxembourg, Esch-sur-Alzette, Luxembourg, [16] INSERM, Etablissement Français du Sang, UMR 1078, Génétique, Génomique Fonctionnelle et biotechnologies, Univ Brest, Brest, France, [17] Service de Génétique Médicale et Biologie de la Reproduction, CHRU de Brest, Brest, France, [18] Department of Primary Care and Public Health, Brighton and Sussex Medical School, University of Brighton, Brighton, United Kingdom, [19] Department of Family Medicine, Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium

OPEN ACCESS

Edited by:
Rita Roncone, University of L'Aquila, Italy

Reviewed by:
Joseph Gone, Harvard University, United States
Stefano Barlati, University of Brescia, Italy

***Correspondence:**
Patrice Nabbe
nabbe.patrice@univ-brest.fr

Specialty section:
This article was submitted to Social Psychiatry and Psychiatric Rehabilitation, a section of the journal Frontiers in Psychiatry

Received: *31 March 2021*
Accepted: *07 July 2021*
Published: *12 August 2021*

Citation:
Nabbe P, Le Reste JY, Guillou-Landreat M, Assenova R, Kasuba Lazic D, Czachowski S, Stojanović-Špehar S, Hasanagic M, Lingner H, Clavería A, Rodríguez-Barragán M, Sowinska A, Argyriadou S, Lygidakis C, Le Floch B, Montier T, Van Marwijk H and Van Royen P (2021) Nine Forward–Backward Translations of the Hopkins Symptom Checklist-25 With Cultural Checks. Front. Psychiatry 12:688154. doi: 10.3389/fpsyt.2021.688154

Introduction: The Hopkins Symptom Checklist-25 (HSCL-25) is an effective, reliable, and ergonomic tool that can be used for depression diagnosis and monitoring in daily practice. To allow its broad use by family practice physicians (FPs), it was translated from English into nine European languages (Greek, Polish, Bulgarian, Croatian, Catalan, Galician, Spanish, Italian, and French) and the translation homogeneity was confirmed. This study describes this process.

Methods: First, two translators (an academic translator and an FP researcher) were recruited for the forward translation (FT). A panel of English-speaking FPs that included at least 15 experts (researchers, teachers, and practitioners) was organized in each country to finalize the FT using a Delphi procedure.

Results: One or two Delphi procedure rounds were sufficient for each translation. Then, a different translator, who did not know the original version of the HSCL-25, performed a backward translation in English. An expert panel of linguists compared

the two English versions. Differences were listed and a multicultural consensus group determined whether they were due to linguistic problems or to cultural differences. All versions underwent cultural check.

Conclusion: All nine translations were finalized without altering the original meaning.

Keywords: depression, Hopkins Symptom Checklist-25, depressive disorder, HSCL-25, diagnostic tool

INTRODUCTION

How to manage people with depression in primary care is a growing challenge worldwide. Indeed, Family practice physicians (FPs) are at the frontline, while secondary care services are increasingly under threat (1–4). Depression manifests (for laypersons) itself in various ways: (i) as a syndromic "disorder" in which contextual distress, anxiety, and somatoform disorders overlap; (ii) as a suffering that is difficult to express, acknowledge, and discuss; and (iii) as a long-term condition with subjective and objective features that can be measured (5). Due to these inter-individual variabilities, FPs may experience difficulties in detecting depression and may easily misjudge the symptoms and their intensities, if they do not use formal instruments (6, 7). Moreover, the depression incidence and prevalence rates differ widely in family practice, due to complex contextual variations, differences in healthcare systems, concepts of disorder, objectives, and practices, as well as cultural variations in symptom expression (8, 9). These difficulties may lead to inappropriate care and potential side effects due to drugs' use as well as public health issues (10–12). A short discussion of the results obtained using a relevant questionnaire is often the first step toward an open dialogue with the patient.

Collaborative primary care mental health models can improve the management of patients with depression. To this aim, the European General Practice Research Network (EGPRN) developed a collaborative research agenda (13). Specifically, the EGPRN adopted a standardized methodology in which European FPs experts from different healthcare systems and who speak different languages and have different cultural references set up an established consensus procedure to identify reliable, standardized, efficient, and ergonomic tools for depression assessment that take into account cultural and linguistic differences (14–17). These tools need to be accepted by both FPs and psychiatrists to improve collaboration (18). They must be feasible in the FP's surgery, in primary or psychiatric care, and also suitable for research purposes (19). Finally, they must be validated and reliable.

A handbook was developed to guide the selection of a single tool that would be then translated into different languages, using a forward and backward translation procedure (inspired by Brislin's translation model). This is a consensual procedure that has been used in other cross-cultural studies (20–22). At each step, the key points and purposes were debated and decided by consensus among the involved European experts. First, a systematic literature review, according to the PRISMA criteria, allowed the identification of seven tools that had been validated against a psychiatric examination using the DSM-IV or DSM-5 major depression criteria (23). Then, a consensus procedure (RAND/UCLA Appropriateness Method) led to the selection of one tool on the basis of its effectiveness, reliability, and ergonomics (24): the self-report Hopkins Symptom Checklist-25 (HSCL-25) (23–26). This is a validated, reliable diagnostic tool to assess (27, 28) the presence and severity of anxiety and depression symptoms during the previous week (29, 30). Its specificity compared with clinical interview is robust: between 0.78 to 0.88, the reliability (Alpha de Cronbach) is between 0.87 to 0.97 (31). The HSCL-25 short length self-administered format is perfectly suited for use in busy primary care settings with many competing demands. It may represent a practical instrument to alert FPs to potentially depressive or anxious symptomatology.

A qualitative procedure with the FP's involvement was necessary to obtained that were linguistically and culturally equivalent to the original version, ecologically embedded in primary care.

The objective of the present study was to translate the HSCL-25 into the languages of the different team members, without losing homogeneity, and in a language suitable to the primary care context (22, 32).

MATERIALS AND METHODS

This three-step standardized study included: (i) forward translation (FT), (ii) backward translation (BT), and (iii) cultural check (8, 33, 34) (**Figure 1**).

The FT was carried out with an incorporated Delphi consensus procedure (35–37). This is a systematic, interactive method that involves a panel of experts using iterative procedures (38) and that allows reaching consensus in a rigorous way (39–41). This process requires:

- Anonymity of participants to ensure response reliability and avoid contamination,
- Iteration, which allows participants to refine their views in the light of the group work progress,
- Feedback control under the investigator's responsibility,
- Statistical aggregation of the group's responses to allow a quantitative and qualitative analysis of data (42–45).

The EGPRN French team ensured that this protocol was followed throughout the process. The FT of the different HSCL-25 items had to be validated daily by the expert panel, composed of EGPRN members, all actively involved in the process.

Abbreviations: BT, backward translation; CE, cultural effect; DSM, diagnostic and statistical manual of mental disorders; EGPRN, European general practice research network; FPs, family practice physicians; FT, forward translation; PRISMA, preferred reporting items for systematic reviews and meta-analyses; RAND, research and development; RAND/UCLA, research and development/University of California Los Angeles.

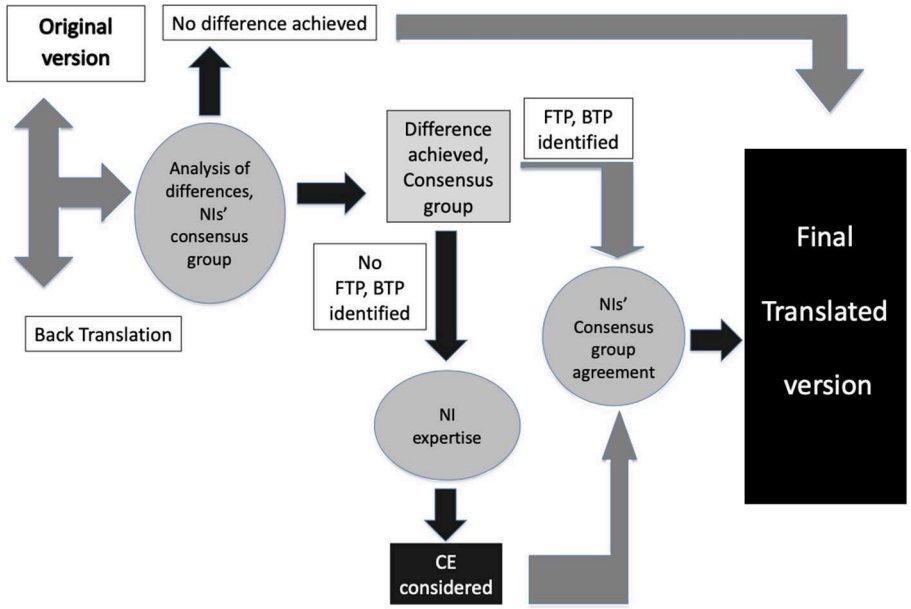


FIGURE 1 | The translation procedure. CE, cultural effect; BTP, backward translation problems; FTP, forward translation problems.

Briefly, for each language, the National Investigators (NI) selected translators knowledgeable about healthcare terminology to organize two translation (FT and BT) teams who were blind to the other team's work. The FT team included one member of the FP research group and one official translator for each country. The BT team involved one (or two) FPs and one official translator (22).

The NIs also recruited a panel of FP experts in their own countries, anonymized the experts' responses, and allocated an identification number for later identification (42). Initially, 20 to 30 experts were recruited per country to secure the presence of at least 15 participants till the project end. The FP experts were selected using the following inclusion criteria: native of their country of residence and speaking their native language, and fluent in English (32). At least half of them had to be involved in teaching and/or research activities. To assess the panel representativeness of their country FPs, the experts provided the following information: sex, practice type, years of practice, and publication record (46).

According to the Brislin's Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures, once the FT was completed, a BT was performed with two goals: (i) to ensure the identification of language issues and (ii) to detect translation problems linked to cultural adaptation issues. Indeed, as translation biases related to cultural aspects of each country were possible, a cultural check was required to ensure homogeneity (17, 20, 33, 34, 47). To this aim, in each country, an FP researcher and a linguist analyzed all BT propositions and compared them with the original HSCL-25 version to establish whether there was any significant difference in terms of meaning. Their report was submitted to a consensus group whose task was to clarify the nature of each FT-BT discrepancy from three problem areas: (i) BT problems were eliminated if the difference was explained by an incorrect BT; (ii) FT problems were defined as an anomaly in transcribing the original English (semantic/idiomatic differences relative to the original English version); and (iii) cultural effects (CE) were considered validated if there was no linguistic problem with the translation, but the item needed to be modified to be understood by the patients in their own "everyday" language (**Figure 1**).

This led to a linguistically stable, definitive translation that maintained the HSCL-25 meaning (i.e., structure and question order and method of use) for each involved country.

Ethical request: The EGPRN French team was in charge of checking the volunteering process and confirming the absence of potential conflicts of interest for all participants. The Ethics Committee of the approved the whole process.

The EGPRN French team recruited all NIs and obtained their consent, managed the voluntary participation in the study and produced an absence of conflict-of-interest statement.

Each NI asked participants to sign the informed consent.

RESULTS

NI Panel Description

The NI panel included 11 NIs (including $n = 8$ women) from eight European countries. They were all FPs, EGPRN members, and fluent in English. Ten NIs practiced in urban areas of more than 5,000 inhabitants and one worked in an urban area with 2,000–5,000 inhabitants. Eight had also teaching duties in addition to being researchers (total number of publications by the panel members: 152). The mean number of years of practice and of research were 21.3 and 12.4 years, respectively. In the panel, two NIs were from two distinct cultural regions of coastal

TABLE 1 | National investigators' panel.

Experts	Gender	Country	Academic Status	Number of inhabitants	Practice type	Number of international publications	Years of practice	Years of research
9	F	Bulgaria	Teacher/Researcher	>5,000	FP group practice	9	14	12
7	F	Croatia	Teacher/Researcher	>5,000	Alone	6	20	12
8	F	Croatia	Teacher/Researcher	>5,000	FP group practice	18	30	20
11	M	France	Teacher/Researcher	>5,000	FP group practice	11	20	5
5	F	Germany	Researcher	2,000–5,000	Ceased practicing 2 years previously	19	23	5
10	F	Germany	Researcher	>5,000	FP group practice	4	18	7
3	F	Greece	Teacher/Researcher	>5,000	FP and paramedic group practice	14	30	18
4	M	Italy	Researcher	>5,000	FP group practice	23	7	6
6	M	Poland	Teacher/Researcher	>5,000	FP group practice	20	30	12
2	F	Spain (Cataluña)	Teacher/Researcher	>5,000	FP group practice	13	22	25
1	F	Spain (Galicia)	Teacher/Researcher	>5,000	FP group practice	15	20	14

F, female; M, male; FPs, family practice physicians.

Spain (Catalonia and Galicia), and two were Croats. The other countries were each represented by a single NI (**Table 1**).

Forward Translation

For the Delphi consensus procedure, 14 (Germany) to 31 experts (Spain) were recruited. In compliance with the selection criteria, they were all FPs and fluent in English. The expert panel included 215 FPs (111 men and 104 women). Among them, 20 worked in a city of <2,000 inhabitants, 36 in a city with 2,000–5,000 inhabitants, and 159 in a city with >5,000 inhabitants. Their clinical experience was analyzed according to years of practice (mean: 16.4 years of experience) (**Table 2**).

In Poland, Bulgaria, Germany, Spain, and the Catalonia region of Spain, there was only one Delphi round, and two rounds in the other countries. Almost all translation proposals for each item of the HSCL-25 questionnaire were accepted in one round (273/320: 85.3%) (**Table 3**). The other proposals for which consensus was not reached went through a second round. The NI and the forward official translator synthesized the experts' comments to produce a new translation proposition for the second round.

Some Translation Issues Required a Second Proposal and Another Delphi Round

In Croatian, eleven proposals were rejected in the first round. For example, for item #17 ("Feeling blue"), the first proposal was "Bili ste tužni," which was considered to be too focused on melancholia, and was modified to "Bili ste sjetni," closer to the concept of sadness. All new proposals were accepted during the second round.

As a German version of the HCL-25 was already available, the German NIs proposed that their expert panel would discuss this version, without producing a new FT. All items were accepted in the first Delphi round. At this step, the German NIs stopped the procedure. No cultural check was performed.

Nine Greek proposals were rejected in the first round. For example, for item #1 ("Being scared for no reason"): the first proposal "Είμαι τρομοκρατημένος χωρίς αιτία" was considered too strong. Consensus was reached on the second proposal: "Είμαι τρομαγμένος χωρίς αιτία." All new proposals were accepted during the second round.

In the French translation, consensus was not reached on 18 proposals in the first round and needed further specification in the second round. For example, for item #25 ("Sleep disturbance"), the first proposal was "Vous n'arrivez pas à dormir" that was modified to "Votre sommeil était perturbé," closer to the English word: "disturbance." All new proposals were accepted during the second round.

In the Italian translation, consensus was not reached on five proposals during the first round. For example, for item #5 ("Heart racing"), the first proposal "Avere tachicardia" was considered too focused on clinical symptoms and was modified to "Sentire il cuore battere veloce," which was more familiar according to the reviewers. All new proposals were accepted during the second round.

In the Spanish Galician translation, consensus was not reached on three proposals in the first round. For example, for item #6 ("Trembling"), the first proposal was "Trema," the present indicative of the verb "Tremar." The second proposal was "Ten tremores" and was accepted in the second round. All new proposals were accepted during the second round.

Backward Translation and Cultural Check

The initial instructions, the 25 items, the quotation and the explanatory sentences were all back-translated into English by the BT team. In total, 36 propositions were analyzed. All BTs were compared linguistically to the original. Differences were noted for

TABLE 2 | Characteristics of each country expert panel.

	N (women)	Practice (mean years)	Number of inhabitants in the practice area			Academic researcher and/or teacher		Number of publications	Participants in the second Delphi round
			<2,000	2,000–5,000	>5,000	Number	Experience (mean, years)		
Bulgaria	22 (13)	20.5	1	5	16	5	5.4	8	No second round
Catalonia	22 (9)	15.7	0	2	20	20	10.5	22	No second round
Croatia	16 (13)	19.2	1	1	14	16	11.5	15	15
France	16 (7)	12.5	1	7	8	15	6.3	11	15
Galicia	20 (6)	22.3	0	0	20	17	13.1	19	20
Germany	14 (8)	16.7	0	3	11	9	10	6	No second round
Greece	26 (13)	10.9	10	9	7	24	5.1	26	15
Italy	18 (6)	17.2	3	2	13	13	14	12	No second round
Poland	30 (18)	11.9	4	6	20	26	13.1	10	No second round
Spain	31 (11)	19.5	0	1	30	27	12	30	No second round
Total	215 (104)	15.55	20	36	178	172	10.1	159	4 Second round

submission to the NIs and the consensus group. Three consensus group meetings were necessary with national feedback between each. The main adaptations, produced as a result of national feedback and the consensus resulting from the cultural check, are described below.

By Languages and Language Groups

Croatia: 8 items were different (2 were BT problems, and 8 required a cultural adaptation).

The main cultural aspect was the use of the present perfect, which is a tense of state and not of action, commonly employed in daily life. Therefore, in items #2, 7, 9, and 10, "feeling" was replaced by "you have been." Only one item seemed to be stronger than in the original version. Indeed, "Faintness," was replaced by "Weakness," but in Croatian this is equivalent to faintness.

Bulgaria: 3 items were different (2 were BT problems, and 1 required a cultural adaptation).

"Feeling low in energy" became "A sense of low energy." Overall, the Bulgarian translation was the most stable among the three Slavic languages.

Poland: 13 items were different (7 were BT problems, and 6 required a cultural adaptation).

Most problems resulted from a conceptual issue. For instance, in Polish, "Heart racing" became "Palpitations," "Trembling" became "Tremors," and "An effort" was translated into "A burden." "Headache" was translated into "Headaches" in Polish for grammatical reasons.

In all three slavic languages (Croatian, Bulgarian, and Polish), "Feeling restless" was translated into "Anxiety" because there is no equivalent word to express these ideas. A word-by-word translation, in that case, was impossible.

For the Greek language, the translation was mainly based on an adaptation according to gender. The experts concluded that there was a general CE affecting all parts of the scale. However, no real difference in meaning was detected, and the Greek HSCL-25 scale remained stable relative to the original.

France: 5 items were different (4 were BT problems, and 1 required a cultural adaptation).

For the French scale, the present tense is normally used in everyday language. However, the past tense was used in the FT. In everyday life French, the past tense is considered an older, upper-class language style. Therefore, all tenses were modified. For instance, "Tout était un effort pour vous" became "Tout est un effort pour vous" in the final version.

Italy: 7 items were different (6 were BT problems, and 1 required a cultural adaptation).

In the Italian scale, the male plural form was used because this is the usual way of speaking/writing; the translation had to be modified according to gender.

Spain: 6 items were different (1 was a BT problem, and 5 required a cultural adaptation).

"Feeling no interest" was translated in "No siente interes por nada" in standard Spanish, and "Worthless feeling" became "Feeling useless." However, in Standard Spanish, "inutil" means also "worthless."

Catalonia: 7 items were different (4 were BT problems, and 3 required a cultural adaptation).

Galicia: 5 items were different (1 was a BT problem, and 4 required a cultural adaptation).

In the Galician scale, item #14 "Losing sexual interest," was translated into "Loss of sexual interest" that expresses a state, and not an action (the original English version); however, the local experts considered it a normal way of speaking/writing in that language.

In the Galician and Catalan translations, "Blame oneself" turned into "Blame yourself" in the BT because the term "oneself" is not commonly employed.

For the Hispanic languages, the translation had to be modified according to gender. The item "Faintness" was translated into "Weakness" (e.g., "Debilidad," "Debilitat," and "Debilidade" in standard Spanish, Catalan and Galician respectively). Similarly, the item "Heart racing" was translated into "Palpitations" (i.e., "Palpitaciones" and "Palpitacions" in the standard Spanish and Galician versions).

TABLE 3 | Results of the first Delphi round.

Item/Country	Galicia	Castile	Catalonia	France	Italy	Bulgaria	Croatia	Greece	Germany	Poland
1 Being scared for no reason	C	C	C	C	C	C	C	NC	C	C
2 Feeling fearful	C	C	C	C	C	C	NC	C	C	C
3 Faintness	C	C	C	NC	NC	C	NC	NC	C	C
4 Nervousness	C	C	C	C	C	C	C	C	C	C
5 Heart racing	C	C	C	NC	C	C	C	C	C	C
6 Trembling	NC	C	C	NC	NC	C	C	C	C	C
7 Feeling tense	C	C	C	C	C	C	C	C	C	C
8 Headache	C	C	C	C	C	C	C	C	C	C
9 Feeling panic	C	C	C	NC	C	C	NC	C	C	C
10 Feeling restless	NC	C	C	NC	C	C	NC	C	C	C
11 Feeling low in energy	C	C	C	C	C	C	NC	NC	C	C
12 Blaming oneself	C	C	C	NC	NC	C	C	C	C	C
13 Crying easily	C	C	C	C	C	C	C	NC	C	C
14 Losing sexual interest	C	C	C	NC	C	C	NC	C	C	C
15 Feeling lonely	C	C	C	NC	C	C	NC	C	C	C
16 Feeling hopeless	C	C	C	C	C	C	NC	C	C	C
17 Feeling blue	C	C	C	NC	C	C	NC	C	C	C
18 Thinking of ending one's life	C	C	C	C	C	C	C	NC	C	C
19 Feeling trapped	C	C	C	NC	C	C	C	C	C	C
20 Worrying too much	C	C	C	NC	C	C	NC	NC	C	C
21 Feeling no interest	C	C	C	NC	C	C	NC	NC	C	C
22 Feeling that everything is an effort	C	C	C	C	C	C	C	C	C	C
23 Feelings of worthlessness	C	C	C	NC	C	C	C	NC	C	C
24 Poor appetite	C	C	C	C	C	C	C	NC	C	C
25 Sleep disturbance	NC	C	C	NC	C	C	C	C	C	C
26 Choose the best answer for how you felt over the past week	C	C	C	NC	C	C	C	C	C	C
27 Not at all	C	C	C	C	NC	C	C	C	C	C
28 A little	C	C	C	NC	C	C	C	C	C	C
29 Quite a bit	C	C	C	C	C	C	C	C	C	C
30 Extremely	C	C	C	C	C	C	C	C	C	C
31 The HSCL-25 score is calculated by dividing the total score (sum score of items) by the number of items answered (ranging between 1.00 and 4.00). It is often used as the measure of distress.	C	C	C	NC	NC	C	C	C	C	C
The patient is considered as a "probable psychiatric case" if the mean rating on the HSCL-25 is ≥1.55.										
32 A cut-off value of ≥1.75 is generally used for diagnosis of major depression defined as "a case in need of treatment." This cut-off point is recommended as a valid predictor of mental disorder as assessed independently by clinical interview, somewhat depending on diagnosis and gender.	C	C	C	NC	C	C	C	C	C	C
The administration time of HSCL 25 is 5–10 min										

C, consensus; NC, no consensus.

For All of Languages

Item #17 "Feeling Blue" generated a CE in six of the nine languages. A word-by-word rendition was impossible and required a cultural adaptation.

Items #15 "Feeling lonely," #18 "Thinking of ending one's life," #19 "Feeling trapped" and #25 "Sleep disturbance" remained stable after the BT.

Concerning the scale instructions and the quotation question, the BT was different from the original version in nine items, except the explanation concerning the time required to fill in the scale. Many translation problems were related to "cultural" effects. For example: in French, some terms were replaced by typical expressions commonly employed in questionnaires: e.g., "pencil-and-paper" was

TABLE 4 | Final translation of the HSCL-25 in nine European languages: items 1–25.

HSCL-25 Original version	Greece	Poland	Bulgaria	Croatia	Castile	Catalonia	Galicia	Italy	France
Choose the best answer for how you felt over the past week	Επιλέξτε την καλύτερη απάντηση για το πώς αισθανθήκατε την τελευταία εβδομάδα	Wybierz najlepsza odpowiedz	Изберете отговора, който най-добре описва как сте се чувствали през изминалата седмица	Izaberite jedan odgovor koji najbolje opisuje kako ste se osjećali tijekom prošlog tjedna:	Elija la respuesta que mejor describa cómo se ha sentido durante la semana pasada	Triï la millor resposta per indicar com s'ha sentit en la darrera setmana	Escolla a resposta que mellor describa como se sentiu durante a semana pasada	Scegliere la risposta più adatta su come ti sei sentito/a nell'ultima settimana	Veuillez choisir la réponse qui décrit le mieux comment globalement vous vous sentiez toute la semaine dernière
Being scared for no reason	Είμαι τρομαγμένος/η χωρίς αιτία	Bać sie bez powodu	Чувство за уплаха без причина	Bili ste bezrazložno uplašeni	Se asusta sin motivo	Estar espantat/ espantada sense motiu aparent	Asústase sen motivo	Avere paura senza motivo	Vous avez peur sans raison
Feeling fearful	Αισθάνομαι φοβισμένος /η	Poczucie strachu	Чувство за страх	Bojali ste se	Siente miedo	Sentir por	Ten medo	Sentirsi impauriti	Vous vous sentez effrayé
Faintness	Αίσθημα λιποθυμιάς	Omdlenia	Отпадналост	Bili ste slabi	Debilidad	Debilitat	Debilidade	Sensazione di mancamento	Vous avez une sensation d'étourdissement
Nervousness	Νευρικότητα	Nerwowość	Нервност	Bili ste nervozni	Nerviosismo	Nerviosisme	Nerviosismo	Esseri nervosi	Vous vous sentez nerveux
Heart racing	Ταχυπαλμία	Kołatanie serca	Сърцебиене	Ubrzano vam je lupalo srce	Palpitaciones	Cor accelerat	Palpitacións	Sentire il cuore battere veloce	Vous avez l'impression que votre cœur bat anormalement vite
Trembling	Τρεμούλα	Drzenia	Треперене	Drhtali ste	Tiembla	Tremola	Ten tremores	Tremore	Vous avez la sensation de trembler
Feeling tense	Αισθάνομαι υπερένταση	Poczucie napiecia	Чувство за напрежение	Bili ste napeti	Se siente tenso/a	Sentir-se tens/a	Séntese tenso/a	Sensazione di tensione	Vous vous sentez tendu
Headache	Πονοκέφαλος	Bóle głowy	Главоболие	Boljela vas glava	Dolor de cabeza	Mal de cap	Dor de cabeza	Avere mal di testa	Vous avez des maux de tête
Feeling panic	Αισθάνομαι πανικό	Uczucie paniki	Чувство за паника	Bili ste u panici	Siente pánico	Sensació de pànic	Sente pánico	Sensazione di panico	Vous vous sentez paniqué
Feeling restless	Αισθάνομαι ταραχή	Uczucie niepokoju	Чувство на безпокойство	Bili ste uznemireni	Siente inquietud	Sensació d'inquietud	Séntese inquedo/a	Sensazione di irrequietezza	Vous vous sentez agité

(Continued)

TABLE 4 | Continued

HSCL-25 Original version	Greece	Poland	Bulgaria	Croatia	Castile	Catalonia	Galicia	Italy	France
Feeling low in energy	Αισθάνομαι ότι δεν έχω ενέργεια	Poczucie braku energii	Усещане за понижена енергия	Niste imali dovoljno energije	Siente que le falta energía	Sensació de manca d'energia	Sente que lle falta enerxía	Sentirsi senza energia	Vous manquez d'énergie
Blaming oneself	Κατηγορώ τον εαυτό μου	Obwinianie samego siebie	Самообвинение	Okrivljavali ste se	Se culpa a sí mismo/a	Culpar-se un/a mateix/a	Cúlpase a si mesmo/a	Avere sensi di colpa	Vous ressentez une sensation de culpabilité
Crying easily	Εύκολο κλάμα	Płaczliwość	Плачливост	Bili ste plačljivi	Llora con facilidad	Plora fàcilment	Chora con facilidade	Piangere facilmente	Vous pleurez facilement
Losing sexual interest	Απώλεια σεξουαλικού ενδιαφέροντος	Utrata zainteresowań sfera seksualna	Загубата на сексуален интерес	Niste bili zainteresirani za spolni odnos	Pierde el interés sexual	Pèrdua de l'interès sexual	Perda do interese sexual	Perdere l'interesse sessuale	Vous ressentez un désintérêt pour la vie sexuelle
Feeling lonely	Αισθάνομαι μοναξιά	Poczucie osamotnienia	Чувство за самотност	Bili ste usamljem	Se siente solo/a	Sentir-se sol/a	Séntese só/soa	Sentirsi soli	Vous avez une sensation de solitude
Feeling hopeless	Αισθάνομαι απελπισμένος/η	Poczucie beznadziejności	Чувство за безнадежност	Osjećali ste sebeznadno	Se siente sin esperanza	Sentiment de desesperança	Séntese sen esperanza	Sentirsi senza speranza	Vous vous sentez désespéré
Feeling blue	Νοιώθω πεσμένος/η	Poczucie przygnebienia	Чувстам се нещастен	Bili ste sjetni	Se siente triste	Sentir-se trist/a	Séntese triste	Sentirsi tristi	Vous avez le cafard
Thinking of ending one's life	Σκέφτομαι να δώσω τέλος στη ζωή	Myśli samobójcze	Мисли за самоубийство	Razmišljali ste da si oduzmete Život	Piensa en acabar con su vida	Pensa en treure's la vida	Pensa en acabar coa súa vida	Avere pensieri di togliersi la vita	Vous avez pensé à mettre fin à votre vie
Feeling trapped	Αισθάνομαι παγιδευμένος /η	Poczucie uwiezienia	Чувстам се като в капан	Osjećali ste sekao da ste u klopci	Se siente atrapado/a	Sentir-se atrapat/atrapada	Séntese atrapado/a	Sentirsi intrappolati	Vous vous sentez pris au piège
Worrying too much	Ανησυχώ υπερβολικά	Zamartwianie sie	Притеснявам се твърде много	Bili ste previše zabrinuti	Se preocupa en exceso	Preocupar-se en excés	Preocúpase en exceso	Preoccuparsi troppo	Vous vous inquiétez trop
Feeling no interest	Αισθάνομαι ότι τίποτε δεν είναι ενδιαφέρον	Poczucie braku zainteresowań	Чувство за загуба на интерест	Bez interesa za bilo što	No siente interés por nada	Sentiment de manca d'interès	Non sente interese por nada	Non avere alcun interesse	Plus rien ne vous intéresse
Feeling that everything is an effort	Αισθάνομαι ότι για το καθε τί χρειάζεται να κάνω προσπάθεια	Poczucie, ze wszystko jest ciezarem	Чувство, че всичко изисква усилие	Sve vam je bilo naporno	Siente que todo le cuesta un esfuerzo	Sentir que tot és un esforç	Sente que todo lle supón un esforzo	Sentire che tutto è uno sforzo	Tout est un effort pour vous
Feelings of Worthlessness	Αισθάνομαι ότι δεν αξίζω τίποτε	Poczucie bezwartościowości	Чувство за безполезност	Osjećali ste se bezvrijedno	Se siente inútil	Sentir-se inútil	Séntese inútil	Sentirsi inutili	Vous avez le sentiment d'être bon à rien
Poor appetite	Μείωση της όρεξης	Słaby apetyt	Лош апетит	Imali ste slab apetit	poco apetito	Pèrdua de la gana	Poco apetito	Avere poco appetito	Vous avez perdu l'appétit
Sleep disturbance	Διαταραχές ύπνου	Zaburzenia snu	Нарушения на съня	Imali ste problema sa spavanjem	Problemas para dormir	Alteració de la son	Alteracións do sono	Disturbi del sonno	Votre sommeil est perturbé

TABLE 5 | Final translation of the HSCL-25 in nine European languages: scale instructions.

Scale instructions original version	Greece	Poland	Bulgaria	Croatia	Spain	Catalonia	Galicia	Italy	France
The HSCL-25 score is based on pencil-and-paper self-report of 25 questions about the presence and intensity of anxiety and depression symptoms over the last week.	Η βαθμολογία του ΗΣΨΛ-25 βασίζεται σε γραπτό ερωτηματολόγιο αυτοαξιολόγησης 25 ερωτήσεων σχετικά με την παρουσία και την ένταση των συμπτωμάτων άγχους και κατάθλιψης κατά την τελευταία εβδομάδα. Οι συμμετέχοντες απαντούν σε μία από τις τέσσερις κατηγορίες για κάθε ερώτημα σε μια κλίμακα εύρους τεσσάρων βαθμών με τιμές από 1 μέχρι 4.	Ocena testu HSCL-25 oparta jest na kwestionariuszu 25 pytań, w którym zakreśla sie na papierze obecność i nasilenie objawów leku i depresji w ciagu ostatniego tygodnia.	Резултатът от HSCL-25 се основава на самостоятелно попълнен инструмент на хартиен носител, включващ 25 въпроса за наличието и интензивността на симптоми на тревожност и депресия през последната седмица.	HSCL-25 skor sastoji se od 25 pitanja koja se rješavaju jednostavno olovkom i papirom, a temelji se na samoprocjeni prisutnosti i intenzitetu ansksioznih i depresivnih simptoma tijekom prošlog tjedna.	La puntuación HSCL-25 se basa en un cuestionario auto cumplimentado con lápiz y papel, de 25 preguntas sobre la presencia y la intensidad de ansiedad y síntomas depresivos en la última semana.	L'escala HSCL-25 es basa en un qüestionari auto administrat de 25 preguntes, sobre la presència i la intensitat de símptomes d'ansietat i depressió en la darrera setmana.	A puntuación HSCL-25 baséase nun cuestionario cumprimentado con lapis e papel, de 25 preguntas sobre a presenza e a intensidade de ansiedade e síntomas depresivos na última semana.	Il punteggio dell' HSCL-25 si basa sulla compilazione di un questionario di autovalutazione in cartaceo ("carta/penna") di 25 domande sulla presenza e intensità di sintomi di ansia e depressione nel corso dell'ultima settimana.	La HSCL-25 est un auto-questionnaire en 25 questions relatives à la présence et à l'intensité des symptômes d'anxiété et de dépression durant toute la semaine dernière.
Participants answer to one of four categories for each item on a four-point scale ranging from 1 to 4		Badani odpowiadaja na jedno z czterech mozliwych kategorii na skali mierzacej wartości od 1 do 4.	Участниците избират една от категориите за всяка позиция по скала от четири точки от 1.00 до 4.00.	Ispitanici odgovaraju jednom od četiri kategorija za svako pitanje na skali od 1-4.	Los/ las participantes responden una de cuatro categorías para cada ítem, en una escala de cuatro puntos que van desde 1 a 4.	Els/les participants responen a una de les quatre categories per a cada ítem en una escala de quatre punts que va de l'1 al 4.	Os participantes responden unha de catro categorías para cada ítem, nunha escala de catro puntos que van desde 1 a 4.	I partecipanti rispondono a una delle quattro categorie per ciascun sintomo su una scala di punteggio che va da 1 a 4.	Les participants cotent chaque proposition, sur une échelle en quatre points, cotée de 1 à 4.
1."Not at all"	Καθόλου	Wcale	Съвсем не	Nimalo	En absoluto	Gens	En absoluto	Per niente	Pas du tout d'accord
2."A little"	Λίγο	Troche	Незначително	Malo	Un poco	Una mica	Un pouco	Poco	Un peu d'accord
3."Quite a bit"	Αρκετά	Znacznie	Съвсем малко	Dosta	Bastante	Bastant	Bastante	Abbastanza	Plutôt d'accord
4."Extremely"	Πάρα πολύ	Bardzo mocno	Извънредно	Jako	Mucho	Molt	Moito	Moltissimo	Complètement d'accord

TABLE 6 | Final translation of the HSCL-25 in nine European languages: general instructions.

Scale instructions original version	Greece	Poland	Bulgaria	Croatia	Spain	Catalonia	Galicia	Italy	France
The HSCL-25 score is calculated by dividing the total score (sum score of items) by the number of items answered (ranging between 1.00 and 4.00). It is often used as the measure of distress.	Η βαθμολογία του ΗΣΨ-25 υπολογίζεται διαιρώντας τη συνολική βαθμολογία (αθροιστική βαθμολογία των ερωτημάτων), διά του αριθμού των ερωτημάτων που απαντήθηκαν (κυμαινόμενο μεταξύ του 1,00 έως 4,00). Συχνά χρησιμοποιείται για τη μέτρηση της δυσφορίας.	Wynik testu HSCL-25 jest obliczany poprzez podzielenie całkowitej liczby punktów (suma punktów z każdej pozycji testu) przez liczbę pozycji na które udzielono odpowiedzi (w skali od 1 do 4). Często służy on do pomiaru dystres.	HSCL-25 резултатът се изчислява, като се раздели общият брой точки (сбор точки по критерий) на броя на отговорените критерии (вариращи между 1,00 и 4,00). Той често се използва като мярка за страдание.	Skor HSCL-25 se izračunava dijeljenjem ukupnog zbroja (zbroj skora pojedinih pitanja) s brojem odgovorenih pitanje (raspon od 1,00 do 4,00). Obično se koristi za mjerenje distresa.	La puntuación del HSCL-25 se calcula dividiendo la puntuación total (sumando la puntuación de todos las preguntas) entre el número de respuestas (varía entre 1,00 y 4,00). Se usa habitualmente para medir el malestar psicológico.	La puntuació total del HSCL-25 es calcula dividint la suma de la puntuació dels diferents ítems pel número d'ítems contestats. El resultat total oscil·la entre 1,00 i 4,00. Aquesta escala sovint s'utilitza com a mesura del malestar psicològic.	A puntuación do HSCL-25 calcúlase dividindo a puntuación total (a suma de todas as preguntas) entre o número de respostas (cuxa puntuación oscila entre 1,00 e 4,00). Úsase de forma habitual para medir o nivel del malestar psicológico.	Il punteggio dell' HSCL-25 si calcola dividendo il punteggio totale (somma dei punteggi degli elementi) con il numero di elementi risposti (che variano da 1,00 a 4,00). Spesso si usa come misura di ansietà.	Le score du HSCL- 25 se calcule en divisant la somme des cotations des propositions par le nombre de réponses reçues. Le résultat final est compris entre 1,00 à 4,00. Il est couramment utilisé pour mesurer la souffrance psychologique.
The patient is considered as a "*probable psychiatric case*" if the mean rating on the HSCL-25 is ≥1.55.	Ο ασθενής θεωρείται σαν "πιθανό ψυχιατρικό περιστατικό" εάν η μέση βαθμολογία του ΗΣΨ-25 είναι '=1,55.	Pacjenta uważamy za "prawdopodobny przypadek psychiatryczny" jeśli średnia ocena w teście HSCL-25 jest >/ (wieksza lub równa) 1,55.	Пациентът се приема като "вероятно психиатричен случай," ако средната оценка по HSCL-25 е ^{3} 1,55.	Pacijent se smatra ≪ vjerojatno psihijatrijskim slučajem ≫ ako je srednja vrijednost na HSCL-25 ≥ 1,55.	El/la paciente se considera un "probable caso psiquiátrico" si el valor medio del HSCL-25 es ≥1,55.	El/la pacient és considerat/considerada com a " probable cas psiquiàtric " si la qualificació mitjana del HSCL-25 és ≥ 1,55.	Considérase que o/a paciente é un "caso psiquiátrico probable" se o valor medio do HSCL-25 é ≥ 1,55.	Il paziente è considerato come un "probabile caso psichiatrico" se il punteggio medio dell'HSCL-25 è ≥1,55.	Le patient est considéré comme ≪ probablement atteint d'un trouble psychiatrique ≫ si le score moyen du HSCL-25 est supérieur ou égal à 1,55.

(Continued)

TABLE 6 | Continued

Scale instructions original version	Greece	Poland	Bulgaria	Croatia	Spain	Catalonia	Galicia	Italy	France
A cut-off value of ≥1.75 is generally used for diagnosis of major depression defined as "*a case, in need of treatment*." This cut-off point is recommended as a valid predictor of mental disorder as assessed independently by clinical interview, somewhat depending on diagnosis and gender.	Το όριο του ˙= 1,75 γενικώς χρησιμοποιείται για τη διάγνωση της μείζονος κατάθλιψης που ορίζεται ως "περίπτωση που χρήζει θεραπείας." Αυτό το όριο συνίσταται σαν ένας έγκυρος προγνωστικός δείκτης ψυχικής διαταραχής, όπως εκτιμάται ανεξάρτητα από την κλινική εικόνα, η οποία εξαρτάται κάπως από τη διάγνωση και το φύλο.	Wartość graniczna>/ (wieksza lub równa) 1,75 ogólnie przyjmuje sie w diagnozowaniu ciezkiej depresji, definiowanej jako „przypadek wymagajacy leczenia." Wartość ta jest zalecana jako istotny czynnik w przewidywaniu obecności choroby psychicznej, wymagajacej jednak niezaleznego wywiadu klinicznego i w pewnym sensie zalezy od rozpoznania i płci.	Гранична стойност от [3] 1,75 обикновено се използва за диагностициране на тежка депресия и определя случая като "случай, нуждаещ се от лечение". Тази гранична стойност, получена независимо от клиничното интервю и зависеща до определена степен от диагнозата и пола, се препоръчва като валиден предиктор за психично разстройство.	Razdjelna točka (cut-off) ≥1,75 se koristi za dijagnozu velikog depresivnog poremećaja i to kao slučaj koji zahtjeva liječenje." Razdjelna točka se preporuča kao validni prediktor mentalnog poremećaja podjednako kao i sama procjena neovisnim kliničkim intervjuom, dijelom ovisan o dijagnozi i spolu.	Por lo general se usa un valor de corte de ≥1,75 para el diagnóstico de depresión mayor, definida como "un caso que necesita tratamiento ." Este valor de corte se considera un predictor válido de un trastorno mental, evaluado de forma independiente mediante entrevista clínica, aunque depende en parte del diagnóstico y el género.	Generalment s'utilitza un punt de tall ≥1,75 per al diagnòstic de la depressió major i es defineix com " cas que precisa de tractament." Es recomana aquest punt de tall com un predictor vàlid de trastorn mental com ho seria l'avaluació independent per entrevista clínica, depenent en part del diagnòstic i del gènere.	Polo xeral, úsase un valor de corte ≥ 1,75 para diagnosticar a depresión maior, definida como "un caso que precisa tratamento ." Este valor de corte recoméndase como un predictor válido dun trastorno mental, avaliado independentemente por medio de entrevistas clínicas, aínda que depende en parte do diagnóstico e do xénero.	Un cut-off che sia >=1,75 è normalmente usato per la diagnosi di depressione maggiore definita come "un caso che necessita di trattamento." Questo cut-off è raccomandato come un valido predittore di disordine mentale come valutato in modo indipendente da un colloquio clinico, dipendente in qualche modo dalla diagnosi e dal genere	Un score supérieur ou égal à 1,75 diagnostique généralement une dépression caractérisée et définit ≪ un patient nécessitant un traitement ≫. Ce seuil est considéré comme un score prédictif validé des troubles mentaux. Il a été évalué de manière indépendante par des études cliniques. Il varie peu quelles que soient les situations diagnostiques et le sexe.
The administration time of HSCL-25 is 5 to 10 minutes.	Ο χρόνος χορήγησης του ΗΣ˜Λ 25 είναι 5 έως 10 λεπτά.	Czas na wykonanie testu HSCL 25 wynosi od 5 do 10 minut.	Времето за провеждане HSCL-25 е от 5 до 10 минути.	Vrijeme za ispunjavanje HSCL-25 je 5-10 minuta.	El tiempo de administración del HSCL-25 es de 5 a 10. minutos.	El temps d'administració del HSCL 25 és de 5 a 10 minuts.	O tempo de realización do HSCL-25 é de 5 a 10 minutos.	Il tempo di somministrazione dell'HSCL-25 è da 5 a 10 minuti.	Remplir le questionnaire HSCL-25 prend entre 5 et 10 minutes.

translated into "auto questionnaire" and "Not at all" by "Pas du tout d'accord."

Interestingly, there were translation similarities (often with stronger meanings or medical connotations) not only among languages belonging to the same linguistic group, but also among languages from different groups. The best example concerns item #3 "Faintness" that was translated into "Weakness" in Catalan, Standard Spanish, Galician, and also in Croatian, a term with a more prosaic than medical connotation.

At the end of the cultural analysis, the consensus group finally concluded that the meaning was not changed, and the translation was finalized in all nine languages (see **Tables 4–6**).

DISCUSSION

Using a three-step qualitative procedure, ecologically embedded in primary care, nine consensual translations of the HSCL-25 were obtained that were linguistically and culturally equivalent to the original version, in three language families (Hellenic, Slavic, and Romance). A German version already existed. The aim of this procedure was to meticulously track inconsistencies between local translations that could lead to misinterpretation. This methodical and transcultural validation ensured the transfer of the same content from one language to another and its reliability (17, 47).

The Greek translation remained the most stable, followed by Bulgarian. Item #17, "Feeling blue" was the most challenging to translate, followed by item #3 "Faintness" and item #5 "Heart racing." Some scales needed adaptations in terms of tense (French, Croatian) and in terms of gender (Greek, Italian, and Hispanic languages).

Research and Teaching Implications

Translation remains the most crucial step in the adoption of an instrument developed in another nation using a different language. Errors in translation may distort the original intent of the instrument, thus compromising its validity and reliability (48). Semantic issues might affect comparability in international studies because the same word is interpreted differently across countries and cultures (49, 50). Moreover, some terms and concepts may not exist in other languages or may have additional connotations that backward translations do not always reveal. Challenges arise not only because of the word-to-word literal translation, but also because of the linguistic form of the language, such as tone and syntax (51).

These nine translations of the HSCL-25 are now linguistically similar, in terms of meaning, compared to the original version. However, they need further testing because this first step is not sufficient to complete the task of translating them and supporting their cross-cultural validity. The external and internal validity of each version has to be evaluated to ensure that their reliability is comparable with that of the original version. This will be achieved through quantitative studies in primary care daily practices (52).

In most European countries, FPs can now use this tool for family practice research studies and for assessing depression severity in their patients. The use of such a shared tool may have a great impact on the feasibility of future research on depression in primary care. It will facilitate data comparison among European countries and consequently it will allow statistical reviews on depression epidemiology and symptoms throughout Europe. The use of the same instrument can support the conceptualization of the studied phenomenon across different studies, and the findings can then be compared (21).

LIMITATIONS

A key point of this study was the FPs' involvement in the translation to reduce the selection bias and to ensure the sample quality nevertheless as in all formalized expert consensus procedure a selection bias of the experts remained possible. Our experts' sample was constructed purposively and if we did our best to avoid a selection bias it remained possible. As described by many translators when discussing scientific translation work, a "specialist" in the field (e.g., primary care daily practice in this case) should take a last look at the translation (20, 53, 54) and become the main arbiter of the quality of the final translation (55). Thus, specific attention was paid in choosing FP researchers and certified bilingual translators with sufficient knowledge of healthcare terminology a selection bias was still possible.

The cultural control check was as consistent as possible. It involved a careful step-by-step analysis to prevent confusion bias and linguistic problems. The formalized consensus method allowed the gradual evaluation of each item to strengthen the accuracy of the validated translations and designing the end-result. Nevertheless, an information or a confusion bias remained possible. Our results should be interpretated in the light of these limitations.

CONCLUSION

A translation of the HSCL-25 in which homogeneity is ensured is now available for Spain and its culturally distinct regions of Galicia and Catalonia, and also for France, Greece, Italy, Poland, Bulgaria, and Croatia. It is now ready to be tested in actual and representative primary care populations to further validate its test-parameters.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by CPP (Protection of Persons Committee) of the University Hospital of Brest. Reference CPP: CPP Ouest VI 872; Study ID RCB: n°2014-A01790-47. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

PN designed the study, collected data, led meetings, drafted the article, and submitted it for publication. JL designed the study, collected data, attended meetings, and reviewed the article. MG-L and BL reviewed the article. RA, DK, SC, MH, HL, AC, MR-B, AS, SA, and CL participated as national investigator. SS-S participated as co-national investigator. TM reviewed the article and gave final approval for the version to be published. HV and PV designed the study, reviewed the article, and gave final approval for the version to be published. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Institute for Health Research (NIHR) Applied Research Collaboration Kent, Surrey, Sussex.

ACKNOWLEDGMENTS

We would like to thank all FPs who participated in the research process throughout Europe and all trainees in Family Medicine from Brest University who participated in the research process and our proof-reader Mrs. Alex Gillman for her accurate translations.

REFERENCES

1. King M, Nazareth I, Levy G, Walker C, Morris R, Weich S, et al. Prevalence of common mental disorders in general practice attendees across Europe. *Br J Psychiatry.* (2008) 192:362–7. doi: 10.1192/bjp.bp.107.039966
2. Ayuso-Mateos JL, Vázques-Barquero JL, Dowrick C, Lehtinen V, Dalgard OS, Casey P, et al. Depressive disorders in Europe: prevalence figures from the ODIN study. *Br J Psychiatry.* (2001) 179:308–16. doi: 10.1192/bjp.179.4.308
3. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl.* (2004) 109:21–7. doi: 10.1111/j.1600-0047.2004.00327.x
4. Torzsa P, Szeifert L, Dunai K, Kalabay L, Novák M. Diagnosis and therapy of depression in family practice}. *Orv Hetil.* (2009). 150:1684–93. doi: 10.1556/oh.2009.28675
5. Jorm AF. Mental health literacy. Public knowledge and beliefs about mental disorders. *Br J Psychiatry.* (2000) 177:396–401. doi: 10.1192/bjp.177.5.396
6. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet.* (2009) 374:609–19. doi: 10.1016/S0140-6736(09)60879-5
7. Ani C, Bazargan M, Hindman D, Bell D, Farooq MA, Akhanjee L, et al. Depression symptomatology and diagnosis: discordance between patients and physicians in primary care settings. *BMC Fam Pract.* (2008) 9:1. doi: 10.1186/1471-2296-9-1
8. Cuéllar I, Paniagua FA. *Handbook of Multicultural Mental Health : Assessment and Treatment of Diverse Populations*. San Diego, CA: Academic Press (2000).
9. Marsella AJ, Yamada AM. Culture and mental health: An introduction and overview of foundations, concepts, and issues. In: Cuéllar I, Paniagua FA, editors. Handbook of Multicultural Mental Health. San Diego, CA: Academic Press (2000). p. 3–24.
10. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, et al. Psychotropic drug utilization in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl.* (2004) 109:38–46. doi: 10.1111/j.1600-0047.2004.00331.x
11. Wells KB. Caring for depression in primary care: defining and illustrating the policy context. *J Clin Psychiatry.* (1997) 58(Suppl 1):24–7.
12. Schoenbaum M, Unützer J, McCaffrey D, Duan N, Sherbourne C, Wells KB. The effects of primary care depression treatment on patients' clinical status and employment. *Health Serv Res.* (2002) 37:1145–58. doi: 10.1111/1475-6773.01086
13. Woltmann E, Grogan-Kaylor A, Perron B, Georges H, Kilbourne AM, Bauer MS. Comparative effectiveness of collaborative chronic care models for mental health conditions across primary, specialty, and behavioral health care settings: systematic review and meta-analysis. *Am J Psychiatry.* (2012) 169:790–804. doi: 10.1176/appi.ajp.2012.11111616
14. Lehti A, Hammarström A, Mattsson B. Recognition of depression in people of different cultures: a qualitative study. *BMC Fam Pract.* (2009) 10:53. doi: 10.1186/1471-2296-10-53
15. Kirmayer LJ, Robbins JM, Dworkind M, Yaffe MJ. Somatization and the recognition of depression and anxiety in primary care. *Am J Psychiatry.* (1993). 150:734–41. doi: 10.1176/ajp.150.5.734
16. Steinert C, Hofmann M, Kruse J, Leichsenring F. The prospective long-term course of adult depression in general practice and the community. A systematic literature review. *J Affect Disord.* (2014) 152, 65–75. doi: 10.1016/j.jad.2013.10.017
17. Bullinger M, Anderson R, Cella D, Aaronson N. Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. *Qual Life Res.* (1993) 2:451–9. doi: 10.1007/BF00422219
18. Zhang J, Patel VL, Johnson TR, Shortliffe EH. A cognitive taxonomy of medical errors. *J Biomed Inform.* (2004) 37:193–204. doi: 10.1016/j.jbi.2004.04.004
19. Mitchell AJ. Clinical utility of screening for clinical depression and bipolar disorder. *Curr Opin Psychiatry.* (2012) 25:24–31. doi: 10.1097/YCO.0b013e32834de45b
20. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* (2000) 25:3186–91. doi: 10.1097/00007632-200012150-00014
21. Brislin RW. Comparative research methodology: cross-cultural studies. *Int J Psychol.* (1976) 11:215–29. doi: 10.1080/00207597608247359
22. Jones PS, Lee JW, Phillips LR, Zhang XE, Jaceldo KB. An adaptation of Brislin's translation model for cross-cultural research. *Nurs Res.* (2001) 50:300–4. doi: 10.1097/00006199-200109000-00008
23. Nabbe P, Le Reste JY, Guillou-Landreat M, Munoz Perez MA, Argyriadou S, Claveria A, et al. Which DSM validated tools for diagnosing depression are usable in primary care research? A systematic literature review. *Eur Psychiatry.* (2017) 39:99–105. doi: 10.1016/j.eurpsy.2016.08.004
24. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. *The RAND/UCLA Appropriateness Method User's Manual* (2001).
25. Fröjdh K, Håkansson A, Karlsson I, Frojdh K, Hakansson A. The Hopkins Symptom Checklist-25 is a sensitive case-finder of clinically important depressive states in elderly people in primary care. *Int J Geriatr Psychiatry.* (2004) 19:386–90. doi: 10.1002/gps.1102
26. Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behav Sci.* (1974) 19:1–15. doi: 10.1002/bs.3830190102
27. Nettelbladt P, Hansson L, Stefansson CG, Borgquist L, Nordström G. Test characteristics of the Hopkins Symptom Check List-25 (HSCL-25) in Sweden, using the Present State Examination (PSE-9) as a caseness criterion. *Soc Psychiatry Psychiatr Epidemiol.* (1993) 28:130–3. doi: 10.1007/BF00801743
28. Sandanger I, Moum T, Ingebrigtsen G, Dalgard OS, Sørensen T, Bruusgaard D. Concordance between symptom screening and diagnostic procedure: the Hopkins symptom checklist-25 and the composite international diagnostic interview I. *Soc Psychiatry Psychiatr Epidemiol.* (1998) 33:345–54. doi: 10.1007/s001270050064
29. Ventevogel P, De Vries G, Scholte WF, Shinwari NR, Faiz H, Nassery R, et al. Properties of the Hopkins symptom checklist-25 (HSCL-25) and the Self-Reporting Questionnaire (SRQ-20) as screening instruments used

in primary care in Afghanistan. *Soc Psychiatry Psychiatr Epidemiol.* (2007) 42:328–35. doi: 10.1007/s00127-007-0161-8
30. Moum T. Mode of administration and interviewer effects in self-reported symptoms of anxiety and depression. *Soc Indic Res.* (1998) 45:279–318. doi: 10.1023/A:1006958100504
31. Lipman RS, Covi' L, Shapiro AK. THE HOPKINS SYMPTOM CHECKLIST (HSCL) factors derived from the HSCL-90. *J Affect Disord.* (1979) 1:9–24. doi: 10.1016/0165-0327(79)90021-1
32. Spooner D, Pachana N. Ecological validity in neuropsychological assessment: a case for greater consideration in research with neurologically intact populations. *Arch Clin Neuropsychol.* (2006) 21:327–37. doi: 10.1016/j.acn.2006.04.004
33. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract.* (2011) 17:268–74. doi: 10.1111/j.1365-2753.2010.01434.x
34. Maneesriwongul W, Dixon JK. Instrument translation process: a methods review. *J Adv Nurs.* (2004) 48:175–86. doi: 10.1111/j.1365-2648.2004.03185.x
35. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ Clin Res Ed.* (1995) 311:376–80. doi: 10.1136/bmj.311.7001.376
36. Haute Autorité Santé. *Bases Méthodologiques Pour l'élaboration de Recommandations Professionnelles Par Consensus Formalisé*. HAS (2006).
37. Linstone HA, Turoff M. *The Delphi Method : Techniques And Applications*. Advanced Book Program. Reading, MA: Addison-Wesley Pub. Co. (1975).
38. Hasson F, Keeney S, McKenna H. Research guidelines for the delphi survey technique. *J Adv Nurs.* (2000) 32:1008–15. doi: 10.1046/j.1365-2648.2000.t01-1-01567.x
39. Graham B, Regehr G, Wright JG. Delphi as a method to establish consensus for diagnostic criteria. *J Clin Epidemiol.* (2003) 56:1150–6. doi: 10.1016/S0895-4356(03)00211-7
40. Hassan T, Barnett D. Delphi type methodology to develop consensus on the future design of EMS systems in the United Kingdom. *Emerg Med J.* (2002) 19:155–9. doi: 10.1136/emj.19.2.155
41. De Villiers MR, De Villiers PJT, Kent AP. The delphi technique in health sciences education research. *Med Teach.* (2005) 27:639–43. doi: 10.1080/13611260500069947
42. Bourrée F, Michel P, Salmi LR. Consensus methods: review of original methods and their main alternatives used in public health. *Rev Epidemiol Sante Publiq.* (2008) 56:e13–21. doi: 10.1016/j.respe.2008.10.005
43. Powell C. The Delphi technique: myths and realities. *J Adv Nurs.* (2003) 41:376–82. doi: 10.1046/j.1365-2648.2003.02537.x
44. Romm FJ, Hulka BS. Developing criteria for quality of assessment: effect of the Delphi technique. *Health Serv Res.* (1979) 14:309–12.
45. Jamieson S. Likert scales: how to (ab)use them. *Med Educ.* (2004) 38:1217–8. doi: 10.1111/j.1365-2929.2004.02012.x
46. Anadón M, Guillemette F. La recherche qualitative est-elle nécessairement inductive? *Rech Qual.* (2007) 5:26–37.
47. Herdman M, Fox-Rushby J, Badia X. "Equivalence" and the translation and adaptation of health-related quality of life questionnaires. *Qual Life Res.* (1997) 6:237–47.
48. Yu DSF, Lee DTF, Woo J. Issues and challenges of instrument translation. *West J Nurs Res.* (2004) 26:307–320. doi: 10.1177/0193945903260554
49. Daugherty JC, Puente AE, Fasfous AF, Hidalgo-Ruzzante N, Pérez-Garcia M. Diagnostic mistakes of culturally diverse individuals when using North American neuropsychological tests. *Appl Neuropsychol.* (2017) 24:16–22. doi: 10.1080/23279095.2015.1036992
50. Schnohr CW, Gobina I, Santos T, Mazur J, Alikasifuglu M, Välimaa R, et al. Semantics bias in cross-national comparative analyses: is it good or bad to have "fair" health? *Health Qual Life Outcomes.* (2016) 14:70. doi: 10.1186/s12955-016-0469-8
51. Hanrahan D, Sexton P, Hui K, Teitcher J, Sugarman J, London AJ, et al. Linguistic and cultural challenges in communication and translation in ussponsored HIV Prevention research in emerging economies. *PLoS ONE.* (2015). 10:e0133394. doi: 10.1371/journal.pone.0133394
52. Nabbe P, Le Reste JY, Guillou-Landreat M, Gatineau F, Le Floch B, Montier T, et al. The French version of the HSCL-25 has now been validated for use in primary care. *PLoS ONE.* (2019) 14:e0214804. doi: 10.1371/journal.pone.0214804
53. Skulmoski GJ, Hartman FT, Krahn J. The delphi method for graduate research. *J Inf Technol Educ.* (2007) 6:1. doi: 10.28945/199
54. Vesga O, Agudelo M, Salazar BE, Rodriguez CA, Zuluaga AF. Generic vancomycin products fail *in vivo* despite being pharmaceutical equivalents of the innovator. *Antimicrob Agents Chemother.* (2010) 54:3271–9. doi: 10.1128/AAC.01044-09
55. Balliu C. L'enseignement de la traduction médicale: pour une nouvelle pragmatique. *Meta J Traduct.* (1994) 39:15–25. doi: 10.7202/001964ar

Author Disclaimer: The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychiatry

ORIGINAL RESEARCH
published: 20 August 2021
doi: 10.3389/fpsyt.2021.678916

Validation of the French Version of the Child Posttraumatic Stress Checklist in French School-Aged Children

Morgane Gindt[1,2,3], *Aurelien Richez*[1,2,3], *Michèle Battista*[1,3], *Roxane Fabre*[2,4], *Susanne Thümmler*[1,2], *Arnaud Fernandez*[1,2,3*†] *and Florence Askenazy*[1,2,3†]

[1] Service Universitaire de Psychiatrie de l'Enfant et de l'Adolescent, Hôpitaux Pédiatriques de Nice CHU-Lenval, Nice, France, [2] Université Côte d'Azur, CoBTek, FRIS, Nice, France, [3] Centre Expert du Psychotrauma PACA Corse, Nice, France, [4] Département de Sante Publique, Centre Hospitalier Universitaire de Nice, Université Cote d'Azur, Nice, France

OPEN ACCESS

Edited by:
David Cohen,
Université Pierre et Marie Curie, France

Reviewed by:
Bertrand Olliac,
University of Limoges, France
Alexis Revet,
Centre Hospitalier Universitaire de Toulouse, France

***Correspondence:**
Arnaud Fernandez
fernandez.a@pediatrie-chulenval-nice.fr

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Child and Adolescent Psychiatry, a section of the journal Frontiers in Psychiatry

Received: 10 March 2021
Accepted: 07 June 2021
Published: 20 August 2021

Citation:
Gindt M, Richez A, Battista M, Fabre R, Thümmler S, Fernandez A and Askenazy F (2021) Validation of the French Version of the Child Posttraumatic Stress Checklist in French School-Aged Children. Front. Psychiatry 12:678916. doi: 10.3389/fpsyt.2021.678916

Background: The child posttraumatic stress disorder checklist (CPC) updated to DSM-5 is a questionnaire aimed to assess posttraumatic stress disorder (PTSD) symptoms in children. It is available in both parents and child versions. The back-translation method has been used for the French translation of the CPC. It has not been yet validated in French-speaking populations. The aim of this study was to assess the psychometric properties and the validity of the CPC in a sample of French-speaking schoolchildren and their parents.

Methods: The sample was composed by 176 children outpatients implicated in the Nice terrorist attack (14 July 2016) aged 7–17 (mean = 11.68 years, SD = 2.63 months) and 122 parents. Cronbach's alpha was used to test CPC internal consistency. The Spearman-correlation coefficient was performed between the French version of the CPC and the Kiddie Schedule for Affective Disorders and Schizophrenia Present and Lifetime version (K-SADS-PL) to assess the convergent validity. An ROC curve was constructed to verify the validity of the cutoff scores. An evaluation of the sensitivity and specificity of each score and a comparison with the diagnosis of the K-SADS-PL were made. Finally, a principal component analysis with varimax rotation was computed to analyze the structure of the French version of the CPC.

Results: Cronbach's alpha coefficient was 0.90 for child version and 0.91 for parent version of the CPC. There was a statistical correlation between the K-SADS-PL for PTSD and the total score of CPC for the child version ($r = 0.62$; $p < 0.001$) and for the parent version ($r = 0.55$; $p < 0.001$). The sensitivity and specificity of the children version with a threshold of >20 were 73.1 and 84.7%, respectively, using the K-SADS-PL as the diagnostic reference for PTSD. Concerning the parent version, using the same recommended cutoff score, the sensitivity, and specificity were 77 and 80.5%, respectively.

Conclusions: The psychometric properties of the French CPC are good. This questionnaire appears to be valid and should be used in French-speaking children.

Keywords: CPC, french, validation, school-aged children, PTSD, psychiatry, psychometric properties

INTRODUCTION

Since the consideration of the specific problem of posttraumatic stress disorder (PTSD) in children in the DSM-III-R (1), the vision of the consequences of a psychotrauma in children has changed. Many studies have shown that traumatic experience in childhood affects the overall development of the child (social development, emotional development, or cognitive development) (2, 3).

In DSM-5 (4), specific clusters were defined according to the age of the child, in particular for young children. The American Academy of Child and Adolescent Psychiatry (5, 6) recommends emphasizing the use of different sources of information (children, parents, and/or caregiver) in order to establish the diagnosis of this disorder. Indeed, children may not reveal their traumatic experience to their parents (7) or parents may minimize it.

Pediatric PTSD includes four main categories of symptoms: revival of the event; avoidance behaviors; impaired cognition and mood; and neuro-vegetative overactivation. These are the same groups of symptoms as those seen in adults. However, their clinical expressions tend to be different: traumatic games, return of developmental fears previously extinguished, or even bedwetting and encopresis (8, 9).

An early diagnosis of PTSD in children allows the practitioner to offer specific and rapid treatment, in order to avoid the chronicity of this disorder and the associated generalization mechanisms. If the symptoms of PTSD are not rapidly treated, developmental damages may appear: self-esteem disorder, personality traits, or cognitive impairments.

In order to help diagnose PTSD, in addition to clinical interviews, several tools can be used. To assess the presence and intensity of posttraumatic stress symptoms, semi-structured interviews or self- or hetero-questionnaires can be offered. These different techniques have some advantages and disadvantages (10). Self-administered questionnaires are quick and easy to administer. Conversely, semi-structured interviews are relatively long and require training. Nevertheless, they allow a more detailed assessment of symptoms and give better information (10).

Scheeringa (11), in agreement with the DSM-5, developed the child PTSD checklist (CPC). This questionnaire assesses the symptoms of PTSD, according to the DSM-5, as well as their frequency in children from 7 to 18 years, following a traumatic event. The lowest age of 7 years has been carried over from the original validation of the English version of the scale (11). Before this age, taking a self-administered questionnaire individually turns out to be impossible because of the lack of reading skills.

It has a child version and a parent version. These versions are built on the same model: a first section evaluating the presence of traumatic events in the child's life (direct or indirect exposure); if an event is checked, the following two sections are then proposed. A second section assesses the frequency of the symptoms (21 questions) and the last one the functional impairment (6 questions). Children and parents are asked to answer each question using a Likert scale of 0 (Never) to 4 (Daily). The cutoffs are 20 for the intensity of the symptoms and 4 for the functional discomfort. The completion time is 15–20 min. It is generally recommended to pass it with a clinician (psychologist or child and adolescent psychiatrist).

Currently, the only specific scale validated to assess the pediatric PTSD in French-speaking children is the Child Post-Traumatic Stress Reaction Index (CPTS-RI) (12). This questionnaire was validated in 2014 by Olliac et al. It helps to highlight the presence of PTSD symptoms and to indicate the intensity of these symptoms. However, this questionnaire is based on the DSM-IV and therefore does not take into account the changes brought about by the DSM-5.

The aim of this article is to validate and examine the psychometric properties of the CPC French version, using the data collected in the "14-7" Program, conducted with children exposed to the Nice (France) terrorist attack, in 2016 (13).

METHODS

Participants and Procedures

The data used for the validation of the French CPC were obtained from a study carried out in the aftermath of the terrorist attack of July 14, 2016, in Nice, France, which resulted in 86 deaths and ∼30,000 people exposed to the attack. A total of 176 children aged 7–17 (mean = 11.68 years, SD = 2.63 months) were recruited (CPC child version). All of them were exposed to a DSM-5 type 1 traumatic event. Among them, 86 were girls (49%). A total of 122 parents were also included to evaluate the psychometric properties of the CPC parent/caregiver version.

The French Consultative Committee for the Protection of Individuals in Biomedical Research (national ethics committee) approved all procedures of the present study (number 2017-A02212-51). An informed consent was signed by the parents and the child.

Measures

The team of the pediatric psychotrauma center of Nice (France), using the Back Translation Method, carried out the French translation (14).

The K-SADS-PL (Kiddie—Schedule for Affective Disorders and Schizophrenia, Present, and Lifetime version) is a semi-structured diagnostic interview for children aged 6–18 (15), in agreement with the DSM-5. It is carried out by questioning the parent(s) and child, in order to integrate them into a summary note, which includes the report of the parent(s), the child's report, and the clinical observations during the interview. The interview covers both current issues (including why the family is seeking an assessment), as well as the latest episodes of the disorder. Most articles use a rating scale with three levels of severity (not present, subliminal, and threshold, which combines both moderate and severe presentations).

The use of the K-SADS-PL makes it possible to take into account the absence of redundancy between the questions due to an oral evaluation vs. written evaluation and the comparison between the oral responses of children and parents to their specific versions of the CPC. In addition, the K-SADS-PL is one of the few clinical instruments available in the French language evaluating pediatric PTSD according to DSM-5.

TABLE 1 | PCA for children and parents. Matrix of items.

	Children		Parents			
	Factor 1	**Factor 2**	**Factor1**	**Factor 2**	**Factor 3**	
Item 1	0.720	0.116	0.579	0.105	0.196	Repetitive memories
Item 2	0.679	0.145	0.553	0.358	0.229	Nightmares
Item 3	0.587	0.111	0.480	0.329	0.201	Derealization
Item 4	0.591	0.106	0.402	0.415	0.315	Freezing
Item 5	0.756	–	0.667	0.368	–	Emotional trouble
Item 6	0.727	–0.105	0.536	0.485	0.278	Physical disturbance
Item 7	0.802	0.177	0.354	0.596	0.333	Negative emotions
Item 8	0.555	–	0.394	0.373	0.242	Avoidance of conversations
Item 9	0.523	–	0.692	0.170	0.137	Avoidance of places or objects
Item 10	–	0.167	0.208	0.115	–	Difficulty remembering
Item 11	0.559	0.191	0.340	0.506	0.154	Negative beliefs
Item 12	0.447	0.105	0.150	0.430	0.232	False thoughts
Item 13	0.553	–	0.543	0.123	–	Anhedonia
Item 14	0.233	–	–	0.243	0.532	Distance from relatives
Item 15	0.493	0.174	0.222	0.118	0.965	Positive emotional difficulties
Item 16	0.648	0.115	0.289	0.509	0.441	Irritability
Item 17	0.309	0.949	–	0.646	–	Imprudence
Item 18	0.489	–	0.850	–	0.169	State of emergency
Item 19	0.621	–	0.685	0.302	0.177	Startle reaction
Item 20	0.571	0.241	0.297	0.387	0.418	Concentration difficulties
Item 21	0.615	–	0.451	0.364	0.400	Sleep disturbances
Eigenvalue	7.78	1.28	8.47	1.69	1.23	
% of variance	37.1	6.1	40.3	8.1	5.9	

Statistical Analyses

Principal component analysis was first carried out. The numbers of dimensions selected were evaluated looking at the eigenvalue diagram. A factorial analysis with an orthogonal rotation (Varimax) was performed (16). The rates of variance explained by the dimensions selected were determined. The loading values were checked in the case of each dimension. Only items that were substantially loaded ($>|0.40|$) on a single factor were selected.

Each dimension that emerged from the principal component analysis was used to define a subscale. The score obtained on each subscale was computed by summing up the answers to the items comprising the subscale. Items were scored from 0 to 4. The floor and ceiling effects were evaluated.

Internal consistency was tested by Cronbach's alpha. A coefficient higher than 0.60 was considered as good (16).

The Spearman-correlation coefficient evaluated concurrent validity between the K-SADS-PL for PTSD and the CPC score.

To prove the validity of the cutoff scores, an ROC curve was constructed which evaluated the sensitivity and specificity of each score compared to the diagnosis with the K-SADS-PL. A total severity cutoff score of three points with the K-SADS-PL was chosen, as it corresponds to the cutoff for clinical diagnosis. Then, we analyzed the ROC curve with the CPC cutoff >20 as recommended by Scheeringa (11).

To determine the link between the different scores and sex and gender in child version, Pearson correlation and Student *t*-test were used.

All analyses were performed using child version in the first time and parent version in the second time.

All statistical analyses were conducted using R software version 3.6.1 (The R Foundation for Statistical Computing, Vienna, Austria) with a statistical threshold for significance set to 0.05 (two-tailed).

RESULTS

Factor Validity

For child version, PCA of the 21 items explained 43% of the variance with four factors (**Table 1**). The score obtained by summing up the 21 items ranged from 0 to 84, and the mean score was 23.6 (SD = 16.5). No floor or ceiling effects beyond the 15% threshold were observed.

The first factor consisted of the four DSM-5 symptoms: revival of the event; avoidance behaviors; impaired cognition and mood; and neuro-vegetative overactivation (items 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 15, 16, 18, 19, 20, and 21). It explained 37% of the variance. The score obtained by summing up the 18 items ranged

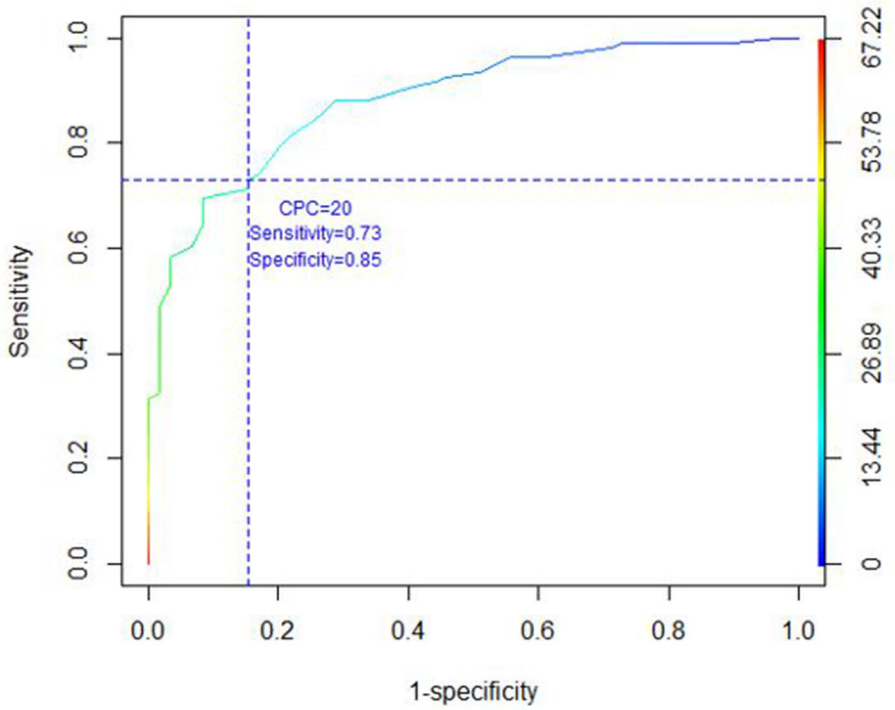


FIGURE 1 | ROC CURVE child version of CPC.

from 0 to 72, and the mean score was 21.4 (SD = 15.6). No floor or ceiling effects were observed.

The second factor was represented only by item 17 which explained 6% of the variance. The mean score was 0.62 (SD = 1.12) and ranged from 0 to 4. One hundred twenty children (68.6%) responded the lower response.

Items 10 and 14 were not included in a factor because the factor loading was low on the two dimensions that emerged.

For parent/caregiver version, PCA of the 21 items explained 54% of the variance with three factors (**Table 1**). The score obtained by summing up the 21 items ranged from 0 to 84, and the mean score was 23.2 (SD = 16.6). No floor or ceiling effects were observed (**Table 1**).

The first factor explained 40% of the variance and consists of items 1, 2, 3, 5, 6, 8, 9, 13, 18, 19, and 21. This factor reflected symptoms of reexperiencing, avoidance, emotional and physical disturbance, anhedonia, nightmares, and sleep disturbances. The score obtained by summing up the 11 items ranged from 0 to 44, and the mean score was 14.0 (SD = 10.3). No floor or ceiling effects were observed.

The second factor explained 8% of the variance with items 4, 7, 11, 12, 16, and 17. They corresponded to symptoms of irritability and negative emotional state, negative beliefs and false thoughts, and freezing. The score obtained by summing up the six items ranged from 0 to 24, and the mean score was 5.7 (SD = 4.0). No floor or ceiling effects were observed.

The third factor consisted of items 14, 15, and 20, explained 6% of the variance, and was related to the distance from relatives (family and friends), attentional difficulties, and positive emotional difficulties. The score obtained by summing up the three items ranged from 0 to 12, and the mean score was 2.9 (SD = 2.8). Thirty-three parents (27.3%) had the lower score possible.

Item 10 was not included in a factor because the factor loading was low on the three dimensions that emerged.

Internal Consistency

Cronbach's alpha for the total CPC was high and homogeneous for the child version (0.90). The first factor had also a good internal consistency (Cronbach's alpha = 0.91). For parents' version, the Cronbach's alphas were 0.92 for the total scale, 0.90 for the first factor, 0.80 for the second factor, and 0.67 for the third factor.

Concurrent Validity

A positive correlation between the K-SADS-PL for PTSD and the total score of CPC was found for the child version ($r = 62$; $p < 0.001$) and for the parent/caregiver version ($r = 0.55$; $p < 0.001$).

For child version, the score total and the first factor were not associated with age and gender. The second factor was not associated with age, but was associated with gender with a lower

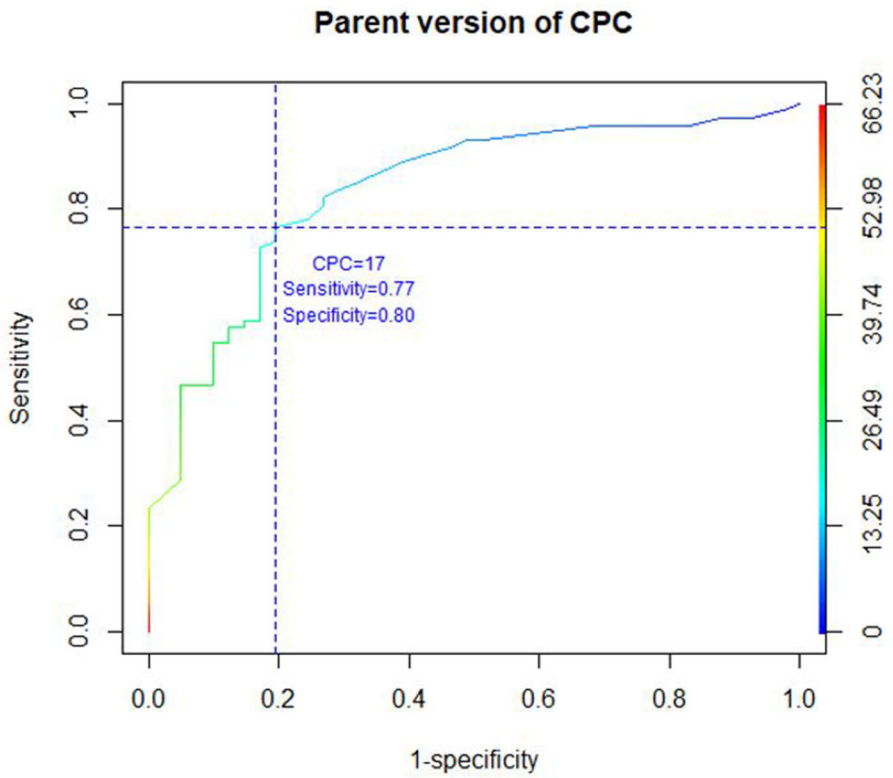


FIGURE 2 | ROC CURVE parent version of CPC.

score for girls (mean = 0.8, SD = 1.3 vs. mean = 0.4, SD = 0.9, $p = 0.010$).

Receiver Operating Characteristic (ROC) Curve

Taking the K-SADS-PL as the diagnostic reference, with a diagnostic cutoff of ≥20 for child version as recommended by Scheeringa (11), the sensitivity and specificity of the child version at that threshold were 73.1 and 84.7%, respectively (**Figures 1**, **2**). Concerning the parent/caregiver version, using the cutoff of ≥17, the sensitivity and specificity were 76.7 and 80.5%, respectively. The sensitivity and specificity for both versions at various cutoff scores can be calculated from the ROC curve coordinates (figX_1 & X_2). The area under the curve was 0.88 for the child version and 0.84 for the parent/caregiver version.

DISCUSSION

Results suggest that the CPC exhibits good psychometric properties (internal consistence, concurrent validity, and factorial validity) in French-speaking school-aged children.

In this French version, the internal consistency was good (Cronbach's alpha 0.90 for the child version and 0.91 for the parent version). In the Olliac study, Cronbach's alpha was 0.87 for the CPTS-RI, with variations between 0.91 and 0.68 depending on the samples (12). However, the internal consistency of the CPC appears to be as good as that of the CPTS-RI.

For the child version, the two-factor structure of our French version explained 37 and 6% of the variance, respectively. The first factor consists of the four main symptoms of PTSD: items exploring reexperiencing of the event, avoidance, alteration of cognition and mood, and overactivation. It explains 37% of the variance. The second factor included the symptoms of imprudence and represented 6% of the variance. With an observed variance of 43% explained by the four factors, the French CPC seems to be a valid tool (17).

Item 11 (concerning negative beliefs), item 12 (unwanted false thoughts concerning the traumatic event), and item 13 (anhedonia) have been added in the DSM-5. It seems to be central symptoms of the pediatric PTSD.

Negative beliefs and false thoughts, in the DSM-5, refer to a change in the child's belief for himself, the world, or other people (18): "Persistent and exaggerated negative expectations about one's self, others, or the world (e.g., 'I am bad,' 'no one can be trusted,' 'I've lost my soul forever,' 'my whole nervous system is permanently ruined,' 'the world is completely dangerous')" (2). Studies have shown that this symptom correlates positively with the presence of PTSD in children (19). In addition, it appears that these negative beliefs refer to maladaptive responses to the psychotrauma experienced and could be involved in the development of internalized symptoms (20).

Anhedonia refers to a loss of interest in previously enjoyed activities and a decrease in the ability to

experience pleasure (21). Recent studies suggest that anhedonia is a transdiagnostic construct (22–25). It is also frequently seen in other neuropsychiatric disorders with which depression is commonly comorbid, such as for example obsessive-compulsive disorder (26) or PTSD (27). Anhedonia appears to be more prevalent in girls than in boys (28). Cumulative traumatic experiences increase anhedonia in PTSD (29). There are also strong associations between anhedonia and dissociative symptoms in children (30).

The results obtained for the PCA indicate that CPC explains as well the observed variance of the CPTS-RI (43% with two factors vs. 44.8% with three factors). The main symptoms are found globally in factor 1 and explain a significant percentage of variance for the two scales. On the other hand, the CPC is more refined for the other factors, due to the inclusion of child-specific symptoms (e.g., negative beliefs or false thoughts) that have been added in the DSM-5.

The main limitation of this study concerns the sample. Indeed, this research was offered to children who lived on July 14, 2016, with or without PTSD. As a result, other studies will have to be carried out in order to test the psychometric properties of this scale, in particular on repeated trauma (e.g., maltreatment or witnessing domestic violence). Furthermore, the number of subjects analyzed is less than the number of subjects needed according to Garson (31). Nevertheless, the results seem robust and statistically significant. We also limited the heterogeneity by analyzing the scores of patients with the same traumatic event.

CONCLUSION

The investigation of PTSD according to DSM-5 may be challenging in children and adolescents. The French version of the CPC is quickly and easily administrated and scored. Its psychometric properties make it a valuable self-administered tool for clinicians and researchers to assess PTSD symptoms in the pediatric population.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: gindt.m@pediatrie-chulenval-nice.fr.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The National Ethics Committee NORD OUEST III (number 2017-A02212-51). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

MG, MB, ST, and FA: conceived and designed the experiments. MG, AR, and MB: performed the experiments. MG, AR, and RF: analyzed the data. MG, AR, AF, and FA: wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

REFERENCES

1. Association AP, DSM-III. *APAWGtR. Diagnostic and Statistical Manual of Mental Disorders*. DSM-III-R: American Psychiatric Association (1987).
2. Matte-Landry A, Collin-Vézina D. Cognitive outcomes of children who have experienced complex trauma: a systematic review protocol. *JBI Evidence Synthesis*. (2020) 18:543–52. doi: 10.11124/JBISRIR-D-19-00036
3. Irigaray TQ, Pacheco JB, Grassi-Oliveira R, Fonseca RP, Leite JCdC, Kristensen CH. Child maltreatment and later cognitive functioning: a systematic review. *Psicol Reflexão e Crítica*. (2013) 26:376–87. doi: 10.1590/S0102-79722013000200018
4. Association AP. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. Washington, DC: American Psychiatric Pub (2013).
5. Action AO. Practice Parameters. Practice parameters for the assessment and treatment of children and adolescents with posttraumatic stress disorder. *J Am Acad Child Adolesc Psychiatry*. (1998) 37:997–1001. doi: 10.1097/00004583-199809000-00025
6. Jackson Y, McGuire A, Tunno AM, Makanui PK. A reasonably large review of operationalization in child maltreatment research: assessment approaches and sources of information in youth samples. *Child Abuse Neglect*. (2019) 87:5–17. doi: 10.1016/j.chiabu.2018.09.016
7. Saltzman WR, Pynoos RS, Layne CM, Steinberg AM, Aisenberg E. Trauma-and grief-focused intervention for adolescents exposed to community violence: results of a school-based screening and group treatment protocol. *Group Dynam Theory Res Practice*. (2001) 5:291. doi: 10.1037/1089-2699.5.4.291
8. Terr LC. Childhood traumas: an outline and overview. *Am J Psychiatry*. (1991) 148:10–20. doi: 10.1176/ajp.148.1.10
9. Joinson C, Sullivan S, von Gontard A, Heron J. Stressful events in early childhood and developmental trajectories of bedwetting at school age. *J Pediatric Psychol*. (2016) 41:1002–10. doi: 10.1093/jpepsy/jsw025
10. Oh DL, Jerman P, Boparai SKP, Koita K, Briner S, Bucci M, et al. Review of tools for measuring exposure to adversity in children and adolescents. *J Pediatric Health Care*. (2018) 32:564–83. doi: 10.1016/j.pedhc.2018.04.021
11. Scheeringa, M. *Child PTSD checklist–Child version and parent version[Measurement instrument]*. New Orleans, LA: Tulane Department of Psychiatry & Behavioral Sciences (2014).
12. Olliac B, Birmes P, Bui E, Allenou C, Brunet A, Claudet I, et al. Validation of the French version of the child post-traumatic stress reaction index: psychometric properties in French speaking school-aged children. *PLoS ONE*. (2014) 9:e112603. doi: 10.1371/journal.pone.0112603
13. Gindt M, Thümmler S, Soubelet A, Guenolé F, Battista M, Askenazy F. Methodology of "14–7" program: a longitudinal follow-up study of the pediatric population and their families exposed to the terrorist attack of nice on July 14th, 2016. *Front Psychiatry*. (2019) 10:629. doi: 10.3389/fpsyt.2019.00629
14. Thümmler S, Gindt M, Maria F, Nachon O, Battista M, Askenazy F. Child *PTSD Checklist (CPC) (French Version 18/05/17 of CPC DSM-5, Scheeringa 23/05/14)*. Nice.
15. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolescent Psychiatry*. (1997) 36:980–8. doi: 10.1097/00004583-199707000-00021
16. Bullinger M. Creating and evaluation cross-cultural instruments. In: *Qual Life Pharmacoecon Clin Trials*. (1996). p. 659–68.

17. Bernstein IH, Garbin CP, Teng GK. Exploratory factor analysis. In: *Applied Multivariate Analysis*. New York, NY: Springer (1988). p. 157–97. doi: 10.1007/978-1-4613-8740-4_6
18. Cox KS, Resnick HS, Kilpatrick DG. Prevalence and correlates of posttrauma distorted beliefs: Evaluating DSM-5 PTSD expanded cognitive symptoms in a national sample. *J Traumatic Stress.* (2014) 27:299–306. doi: 10.1002/jts.21925
19. Lee SH, Kim EJ, Noh J-W, Chae J-H. Factors associated with post-traumatic stress symptoms in students who survived 20 months after the Sewol ferry disaster in Korea. *J Korean Med Sci.* (2018) 33:e90. doi: 10.3346/jkms.2018.33.e90
20. Berman IS, Petretric P, Bridges AJ. Beyond child maltreatment: the incremental value of household dysfunction in the prediction of negative beliefs and internalizing symptoms in women. *J Am Coll Health*. (2019) 69:537–45. doi: 10.1080/07448481.2019.1687483
21. Pushkarskaya H, Sobowale K, Henick D, Tolin DF, Anticevic A, Pearlson GD, et al. Contrasting contributions of anhedonia to obsessive-compulsive, hoarding, and post-traumatic stress disorders. *J Psychiatric Res.* (2019) 109:202–13. doi: 10.1016/j.jpsychires.2018.11.029
22. Abramovitch A, Pizzagalli DA, Reuman L, Wilhelm S. Anhedonia in obsessive-compulsive disorder: beyond comorbid depression. *Psychiatry Res.* (2014) 216:223–9. doi: 10.1016/j.psychres.2014.02.002
23. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am Psychiatric Assoc.* (2010) 167:748–51. doi: 10.1176/appi.ajp.2010.09091379
24. Insel TR, Cuthbert BN. Brain disorders? Precisely. *Science.* (2015) 348:499–500. doi: 10.1126/science.aab2358
25. Weinberg A, Liu H, Hajcak G, Shankman SA. Blunted neural response to rewards as a vulnerability factor for depression: results from a family study. *J Abnormal Psychol.* (2015) 124:878. doi: 10.1037/abn0000081
26. Overbeek T, Schruers K, Griez E. Comorbidity of obsessive-compulsive disorder and depression: prevalence, symptom severity, and treatment effect. *J Clin Psychiatry.* (2002) 63:1106–12. doi: 10.4088/JCP.v63n1204
27. Campbell DG, Felker BL, Liu C-F, Yano EM, Kirchner JE, Chan D, et al. Prevalence of depression–PTSD comorbidity: implications for clinical practice guidelines and primary care-based interventions. *J General Int Med.* (2007) 22:711–8. doi: 10.1007/s11606-006-0101-4
28. Reebye P, Moretti MM, Wiebe VJ, Lessard JC. Symptoms of posttraumatic stress disorder in adolescents with conduct disorder: sex differences and onset patterns. *Canad J Psychiatry.* (2000) 45:746–51. doi: 10.1177/070674370004500808
29. Wechsler-Zimring A, Kearney CA. Posttraumatic stress and related symptoms among neglected and physically and sexually maltreated adolescents. *J Traumatic Stress.* (2011) 24:601–4. doi: 10.1002/jts.20683
30. Laor N, Wolmer L, Kora M, Yucel D, Spirman S, Yazgan Y. Posttraumatic, dissociative and grief symptoms in Turkish children exposed to the 1999 earthquakes. *J Nervous Mental Dis.* (2002) 190:824–32. doi: 10.1097/00005053-200212000-00004
31. Garson GD. *Testing Statistical Assumptions*. Asheboro, NC: Statistical Associates Publishing (2012).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychiatry

ORIGINAL RESEARCH
published: 26 August 2021
doi: 10.3389/fpsyt.2021.704319

The Transplant Evaluation Rating Scale Predicts Clinical Outcomes 1 Year After Lung Transplantation: A Prospective Longitudinal Study

Mariel Nöhre[1,2†], *Martina de Zwaan*[1,2*†], *Maximilian Bauer-Hohmann*[1], *Fabio Ius*[2,3], *Christina Valtin*[4] *and Jens Gottlieb*[2,4]

[1] Department of Psychosomatic Medicine and Psychotherapy, Hannover Medical School, Hannover, Germany, [2] Member of the German Center for Lung Research (DZL), Biomedical Research in End-stage and Obstructive Lung Disease Hannover, Hannover, Germany, [3] Department of Cardiac, Thoracic, Transplant, and Vascular Surgery, Hannover Medical School, Hannover, Germany, [4] Department of Respiratory Medicine, Hannover Medical School, Hannover, Germany

OPEN ACCESS

Edited by:
Michael Noll-Hussong,
Saarland University Hospital, Germany

Reviewed by:
Mitchell Henry,
Ohio State University Hospital, United States
Philipp Lepper,
Saarland University Hospital, Germany
Ulrik Fredrik Malt,
University of Oslo, Norway
Frank Langer,
Saarland University Hospital, Germany

***Correspondence:**
Martina de Zwaan
dezwaan.martina@mh-hannover.de

[†]These authors share first authorship

Specialty section:
This article was submitted to Psychosomatic Medicine, a section of the journal Frontiers in Psychiatry

Received: 02 May 2021
Accepted: 22 July 2021
Published: 26 August 2021

Citation:
Nöhre M, de Zwaan M, Bauer-Hohmann M, Ius F, Valtin C and Gottlieb J (2021) The Transplant Evaluation Rating Scale Predicts Clinical Outcomes 1 Year After Lung Transplantation: A Prospective Longitudinal Study. Front. Psychiatry 12:704319. doi: 10.3389/fpsyt.2021.704319

Objectives: It has been recommended that all candidates for lung transplantation undergo pre-transplant psychosocial evaluation for risk assessment. However, psychosocial issues are only important if they correlate with outcomes after transplantation.

Methods: In this prospective study patients who were referred for lung transplantation from 2016 to 2018 (n = 352) at Hannover Medical School were evaluated using the Transplant Evaluation Rating Scale (TERS). Clinical outcomes included listing, and post-transplant outcomes including mortality, medical aspects such as lung allograft dysfunction, hospitalizations, and renal function, behavioral aspects such as BMI and adherence, and mental issues such as levels of depression, anxiety, and quality of life. TERS scores were divided into tertiles and, in addition, the impact of the two subscale scores—"defiance" and "emotional sensitivity"—was investigated.

Results: Of the patients who were transplanted (n = 271) and were still alive (n = 251), 240 had already reached their 1-year assessment at the end of 2020 and were evaluated 1 year after the operation. A subgroup of 143 received an extended mental assessment. BMI, adherence scores, levels of anxiety, depression, and quality of life 1 year post-transplantation differed significantly between TERS tertiles with higher TERS scores predicting less favorable outcomes. The TERS subscale "defiance" was predictive of BMI and adherence whereas the TERS subscale "emotional sensitivity" was predictive of symptoms of anxiety and depression, and quality of life 1 year after transplantation. Patients in the lowest TERS tertile were more likely to having been listed and—as a trend—to having survived the first year after transplantation

Conclusions: Our findings show that psychosocial factors as measured by TERS score are predictors of behavioral and mental outcomes 1 year after lung transplantation. The TERS allows us to focus on psychosocial risk factors that can be treated or minimized before or after transplantation.

Keywords: lung transplantation, Transplant Evaluation Rating Scale, psychosocial functioning, quality of life, psychosomatic medicine

INTRODUCTION

Lung transplantation is an accepted treatment option for patients with irreversible chronic lung disease, with more than 4,500 procedures performed per year worldwide and 350 in Germany (1, 2). In patients awaiting lung transplantation, symptoms of depression and anxiety and poor pre-transplant quality of life are highly prevalent (3–5) and may be associated with worse post-transplant outcomes, including increased mortality (6, 7). All patients who are considered lung transplant candidates usually undergo a transplant evaluation that includes both medical and psychosocial aspects. The objective of pre-surgical psychosocial evaluations is to identify patients at risk for medical, behavioral and emotional complications during and after organ transplantation. Thus, this evaluation is supposed to judge suitability for transplantation and to guide proactive interventions before and after transplantation (8).

In patients after lung transplantation, the literature on psychosocial predictors on a wider range of outcome measures is relatively scarce. One study did not find a predictive value of the Psychosocial Assessment of Candidates for Transplant (PACT) on 1-year survival (9) while a more recent study found a significant predictive value of this instrument on longer-term (12 year) survival in lung transplant recipients (10). Others found an association between specific pre-transplant psychosocial factors (executive functioning, memory performance, quality of well-being) and mortality following lung transplantation (11, 12). However, these studies did not use a structured psychosocial risk scale that accounts for all psychosocial factors.

The guideline of the German Medical Association concerning lung transplantation dictates that lung transplant candidates should be evaluated by a mental health professional before transplantation (13). Currently, the TERS is the most frequently used instrument (14, 15). The TERS has demonstrated efficacy in predicting peri- and post-transplant outcomes in patients receiving heart, lung, and liver transplants, but also bone marrow or stem cell transplantation as well as left ventricular assist device implantation (15–20). In previous studies of our group in patients prior to lung transplantation, we evaluated the level of psychosocial functioning using the TERS and validated the TERS and its subscales specifically in patients awaiting lung transplantation (21). However, we did not perform detailed follow-up analyses.

Also, even though survival is by far the most relevant outcome, it is not the only outcome. In addition to survival and transplant rates, success in lung transplantation should also be defined by patient-centered outcomes such as levels of depression and quality of life (22–24). For the evaluation of treatment effectiveness quality of life has become a meaningful clinical endpoint (25). This recognizes that the perspectives of patients are unique and may differ from those of clinicians. Additionally, prediction of adherence is crucial because non- or hypo-adherence to immunosuppressive medication and necessary medical recommendations is closely associated with a less favorable outcome also after lung transplantation (26, 27).

Up to now, no prospective studies have examined the predictive value of TERS scores with regard to a large number of peri- and 1-year post-transplant outcomes in patients awaiting lung transplantation. This would potentially drive attention to and help address specific needs in patients with specific characteristics. Thus, the objective of this single-institution study was to assess the impact of psychosocial factors as measured by TERS score on medical, behavioral, and psychosocial outcomes 1 year after lung transplantation. More specifically, in our study peri- and post-transplant outcomes included listing, mortality, prevalence of chronic lung allograft dysfunction, hospitalizations, renal function, weight, adherence, levels of depression and anxiety, and quality of life. In line with the results in the literature, we expected higher pre-transplantation TERS scores to be associated with poorer medical outcomes, poorer adherence, higher levels of depression and anxiety, and lower quality of life 1 year post-transplantation.

METHODS

Patients and Procedures

Trained residents and master-level psychologists conducted a TERS interview according to a structured protocol during routine psychosocial clinical assessment prior to enlistment for lung transplantation. The structured protocol contains the modules for affective disorders, anxiety disorders, adjustment disorders, substance use disorders, and somatoform disorders of the Structured Clinical Interview for DSM IV disorders (28). Patients ($n = 352$) presenting for psychosocial evaluation prior to lung transplantation in 2016, 2017, and 2018 participated. All lung transplant recipients received scheduled follow-up care at the transplant center. The ethics committee of Hannover Medical School approved the study and all participants gave written informed consent before study entry.

Instruments

Pre-transplantation

The Transplant Evaluation Rating Scale (TERS) (13, 14) is an expert interview for the assessment of psychosocial functioning prior to organ transplantation with satisfying inter-rater reliability scores (kappa between 0.8 and 0.9) (29). The German version has been validated in patients awaiting lung transplantation (21). It covers 10 distinct domains of psychosocial functioning considered relevant for adjustment to transplantation and its consequences: (a) current or past mental disorders (axis 1 according to DSM-IV), (b) personality disorders (axis 2 according to DSM-IV), (c) substance use/abuse, (d) compliance, (e) health behaviors, (f) quality of family and social support, (g) history of coping, (h) current coping with disease and treatment, (i) quality of affect and, (j) mental/cognitive status (past and present). Each of the 10 domains is rated by a clinician on a three-point scale based on the level of presence of symptoms within each domain (1 = minimal/mild, 2 = moderate, 3 = severe impairment). Reflecting the importance of the respective domain for the overall level of psychosocial functioning, each item rating is multiplied by a priori assigned weight (ranging from 1 to 4) and the items are added up to calculate the total (weighted) score (range 26.5-79.5). Higher scores represent greater impairment in the levels of psychosocial functioning.

Several research groups have detected a two-factor structure of the TERS in different transplant sample named "defiance" and "emotional sensitivity" which showed differential convergent and predictive validity (21). "Defiance" is a clearly demarcated behavioral factor comprised of a history of difficulties with substance abuse/use, health self-care, non-compliance, family support, personality disorders, and general coping. "Emotional sensitivity" is composed of items tapping quality of affect, adjustment to illness, mental status, and mental disorders. On the basis of a patient's weighted total score, patients were divided into three tertile groups. The tertile method has been recommended since it does not cause inflation of *p*-values compared with outcome dependent cut points (17, 19). For the two subscales we used the median as a cutoff. Even though the TERS was not developed as a scalable instrument we calculated Cronbach's alpha ($\alpha = 0.647$).

Other clinical variables included demographic information and pulmonary diagnosis. Patients were asked to report their age, sex, years of completed education, and partnership status. Patients were classified into four categories depending on their underlying disease (2): category A, obstructive airway diseases (e.g., chronic obstructive pulmonary disease [COPD]); category B, diseases of the pulmonary circulation (e.g., idiopathic pulmonary arterial hypertension); category C, infectious lung diseases (e.g., cystic fibrosis [CF]); and category D, restrictive lung diseases (e.g., pulmonary fibrosis).

Post-transplantation

Medical Outcomes

Enlistment, patient survival, prevalence of chronic lung allograft dysfunction (FEV1 < 80%), number and duration of hospitalizations, renal function (eGFR), and overall comorbidity (Charlson Comorbidity Index, CCI) (30) were taken from our comprehensive institutional database.

Behavioral Outcomes

To assess adherence, five domains were evaluated using a three-level Likert scale (26) with an overall adherence rating between 0 and 100%. The five domains include: (1) health perception (e.g., inconsistent medication knowledge, tobacco/drug abuse, poor diabetic control, use of sunbeds), (2) home spirometry frequency, (3) contact (e.g., missed appointments), (4) nutrition, exercise (e.g., regular exercise, normal-weight), and (5) trough levels in target range. Adherence ratings were completed at each post-operative visit. Scores were assigned by transplant coordinators and discussed with physicians during daily team meetings. The mental health professional was not involved in the rating of the five adherence domains. Mean adherence scores including all available post-operative ratings up to 1 year were calculated. In a recent study from our center including patients from 2010 to 2013 the median adherence score was 86% in the first 3 years after transplantation. After 5 years, patients below and above this cutoff differed significantly with regard to allograft and patient survival and chronic allograft dysfunction (26). Thus, we used the cutoff of 86% to differentiate between good and suboptimal adherence in our sample.

To estimate the immunosuppressive drug adherence we used the four-item interview version of the Basel Assessment of Adherence to Immunosuppressive Medication Scale (BAASIS©) (31). Participants were asked about how often, over the last 4 weeks, they (1) had not taken their drugs (taking dimension), (2) had taken their medication more than 2 h before or after their prescribed taking time (timing dimension), (3) had skipped at least two consecutive doses of their drugs (drug holidays), and/or (4) had reduced the prescribed amount of their medication (dose reduction). Responses were given on a six-point scale ranging from 0 (never) to 5 (every day). Non-adherence was dichotomously defined as any self-reported non-adherence on any of the four items.

Psychological Outcomes

Depression and Anxiety. All patients filled out the four-item Patient Health Questionnaire (PHQ-4) (32), an ultra-brief self-report questionnaire that consists of a two-item depression scale (PHQ-2) and a two-item anxiety scale (GAD-2). Replies are rated on a four-point Likert scale (0 = not at all to 3 = nearly every day). Thus, the total score of the scale ranges between 0 and 12 points. In the current study, the Cronbach's α for the overall score was 0.842. PHQ-4 scores of 6 or above are considered indicative for the presence of a depressive or anxiety disorder. For the PHQ-2 and the GAD-2, scale scores of ≥3 were suggested as cut-off points between the normal range and probable cases of depression or anxiety, respectively.

The subgroup of 143 patients who participated in a more detailed psychosocial assessment also completed the Patient Health Questionnaire-9 (PHQ-9) (33, 34), a self-report instrument screening for symptoms of depression over the last 2 weeks. Nine items are rated on a four-point Likert scale (0 = not at all to 3 = nearly every day) (Cronbach's $\alpha = 0.811$). They also completed the Generalized Anxiety Disorder-7 (GAD-7) (35, 36), a self-report instrument screening for symptoms of generalized anxiety during the last 2 weeks. Seven items are rated on a four-point Likert scale (0 = not at all to 3 = nearly every day) (Cronbach's $\alpha = 0.895$). In both scales, all scores are summed up into a total score, with higher scores representing higher levels of depressive and anxiety symptoms, respectively. For both scale values from 5 to 9 represent mild, from 10 to 14 moderate, and ≥15 severe symptom severity.

Quality of Life. Self-rated levels of Quality of Life (QoL) were assessed during the clinical interviews with a visual analog scale by asking patients: "on a scale of 0-10, with 10 meaning perfectly satisfied, how satisfied are you with your current quality of life?" (QoL VAS).

The subgroup of 143 patients also completed the Pulmonary-specific Quality-of-Life Scale (PQLS), a self-report questionnaire assessing quality of life specifically in patients with end-stage lung diseases (25, 37). The scale consists of 25 items which are rated on a five-point-Likert-scale ranging from 1 ("not at all") to 5 ("most of the time"). A total score between 25 and 125 can be reached with higher values indicating lower quality of life (Cronbach's $\alpha = 0.871$). Three subscales ("task interference,"

"psychological," and "physical") were identified in the original English version of the PQLS (25). The subscale "task interference" (eight items) (Cronbach's $\alpha = 0.801$) focuses on occupational and social functioning, the subscale "psychological" (seven items) (Cronbach's $\alpha = 0.833$) assesses mental and psychological aspects, and the subscale "physical" (four items) (Cronbach's $\alpha = 0.884$) evaluates physical functioning. Six items do not load on any factors; thus, the total scale is also reported.

They also completed the SF-8, a short form of the SF-36 Health Survey, which is used for generic assessment of physical and mental aspects of health-related quality of life (HRQoL) (38) In the SF-8 each of the 8 SF-36 dimensions is represented by a single item to be assessed over the last 4 weeks (Cronbach's $\alpha = 0.867$). The values of these eight dimensions were aggregated to a physical component summary (PCS) value and a mental component summary (MCS) value which were converted to a standardized T score. The T score is a metric with a mean of 50 and standard deviation of 10 that has been normalized to the US general population. German reference values are available, allowing a comparison between the T scores of our sample and German norms (39).

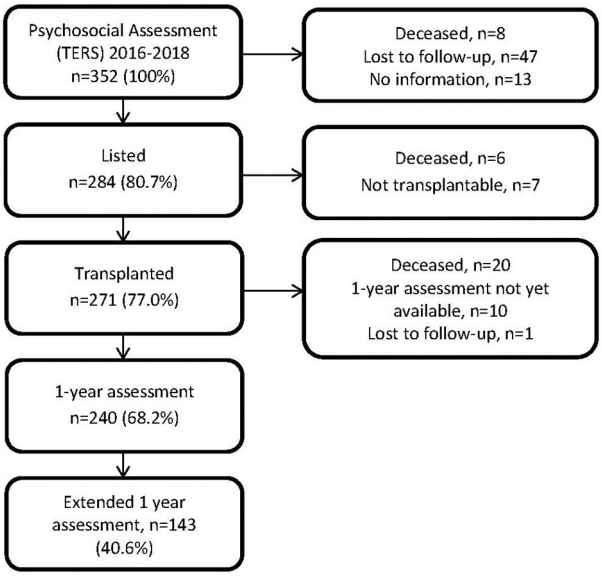


FIGURE 1 | Flowchart of participating patients.

Statistical Analyses

Statistical analyses were performed using SPSS (IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.). Categorical variables are presented as numbers (n) and percentages (%), continuous variables as median and range. Post-transplant outcomes were compared between the TERS tertile groups and between the median split subscale scores "defiance" and "emotional sensitivity" using Kruskal-Wallis *H*-tests and Mann-Whitney *U*-tests for continuous variables and chi-square tests for categorical variables. In addition to significance testing, we calculated Cramer V as effect size (40) for chi-square tests: 0.1 indicates a small effect, 0.3 a medium effect, and 0.5 a large effect and eta squared (η^2) as effect size for non-parametric tests: 0.01 indicates a small effect, 0.06 a medium, and 0.14 a large effect. Binary logistic regression analyses were conducted with significant outcomes as the dependent variable (adherence, BMI, PHQ-4) and TERS tertiles and the two subscale scores, respectively, as the main independent variable controlling for the baseline variables age, sex, educational level, and pulmonary diagnosis. The level of significance was set at ≤ 0.05.

RESULTS

Sample

Of the patients who were transplanted ($n = 271$) and were still alive ($n = 251$), 240 had already reached their 1-year assessment at the end of 2020 and were evaluated (**Figure 1**). Overall, 34 patients had died, 14 before transplantation and 20 (7.4%) of transplanted patients during the first year after transplantation. The median age of our patient sample 1 year after transplantation ($n = 240$) was 55.7 years (range 20-71), 114 (47.5%) were women (**Table 1**). Most patients underwent bilateral lung transplantation ($n = 237$), 10 patients underwent single lung transplantation, 1 patient combined heart-lung transplantation, 2 patients combined lung-liver transplantation, and 21 patients had a double lung re-transplantation due to bronchiolitis obliterans syndrome (BOS). The median LAS score was 34.5 (range 30.6-77.9) with 12 (5%) patients reaching a final pre-transplant lung allocation score (LAS) of 50 or above, which is considered "high" (2). Seventy-two (30%) of the patients met criteria for a lifetime mental disorder and 46 (19.2%) for a current mental disorder. Seventy patients (29.2%) reported experience with psychological/psychiatric treatments and 49 (20.4%) with psychopharmacological treatment. The most frequent diagnoses were affective and anxiety disorders. Sixty patients (25%) had the minimal score on the TERS of 26.5 and 18 (7.5%) scored in the high risk group (≥ 37.5) as defined by Hoodin and Kalbfleisch (20). The total population was stratified according to their TERS scores into tertiles. The three TERS tertiles did not differ with regard to age, sex, and partnership status; however, patients in the highest tertile were significantly less educated and were more often diagnosed with an obstructive lung disease (**Table 2**). These differences were mainly due to differences in the "defiance" subscale (**Table 3**).

Standard maintenance immunosuppression consisted of a triple drug regimen including a calcineurin inhibitor (CNI), prednisolone and mycophenolate mofetil.

Baseline characteristics of all patients who received the TERS ($n = 352$), of the listed patients ($n = 284$), of the transplanted patients ($n = 271$), of the patients with 1-year assessments ($n = 240$), and of patients who participated in the extended 1-year assessment ($n = 143$) are summarized in **Table 1**. No major differences between samples could be detected.

Prediction of Outcome

Medical Outcomes

As of 31th December 2020, 284 patients of the entire sample of 352 psychologically assessed patients had been listed (80.7%). The

TABLE 1 | Baseline characteristics of different samples.

	TERS pre tx available *N* = 352	Listed *N* = 284	Transplanted *N* = 271	1-year post tx assessment *N* = 240	Extended 1-year post tx assessment, *N* = 143
Sex, *n* (%)					
Female	165 (46.9)	134 (47.2)	129 (47.6)	114 (47.5)	66 (46.2)
Male	187 (53.1)	150 (52.8)	142 (52.4)	126 (52.1)	77 (53.8)
Age at TERS assessment, median (range)	53.3 (18-70)	53.5 (18-70)	54 (18-70)	54.3 (18-70)	54.5 (18-70)
Educational level, *n* (%)	*N* = 350	*N* = 283	*N* = 270	*N* = 239	*N* = 142
< 12 years	248 (70.9)	197 (69.6)	187 (69)	165 (69)	99 (69.7)
≥ 12 years	102 (29.1)	86 (30.4)	83 (31)	74 (31)	43 (30.3)
Partnership, *n* (%)					
Yes	280 (79.5)	237 (83.5)	225 (83)	201 (83.8)	120 (83.9)
No	72 (20.5)	47 (16.5)	46 (17)	39 (16.3)	23 (16.1)
LAS category, *n* (%)					
Category A	—	91 (32)	89 (32.8)	79 (32.9)	45 (31.5)
Category B	—	18 (6.3)	16 (5.9)	14 (5.8)	7 (4.9)
Category C	—	55 (19.4)	50 (18.5)	45 (18.8)	35 (24.5)
Category D	—	120 (42.3)	116 (42.8)	102 (42.5)	56 (39.2)
Last pre tx LAS score, median (range)			*N* = 270		
	—	—	34.5 (30.6-94.2)	34.5 (30.6-77.9)	34.5 (30.6-77.9)
Pre tx BMI, kg/m^2, median (range)	*N* = 347	*N* = 263	*N* = 263		
	22.0 (14-34.3)	22.9 (14.1-32.5)	22.4 (14-34.3)	22.4 (14-32.5)	22.5 (14-32.5)
TERS weighted score, median (range)	30.5 (26.5-57.0)	30.8 (26.5-48.5)	30.0 (26.4-48.5)	30.0 (26.5-48.5)	30.5 (26.5-48.5)

BMI, Body Mass Index; LAS Category, pulmonary diagnosis according to the Lung Allocation Score (Category A, obstructive airway diseases; category B, diseases of the pulmonary circulation, category C, infectious lung diseases, category D, restrictive lung diseases); TERS, Transplant Evaluation Rating Scale; pre tx, pre-transplantation.

TABLE 2 | Comparison of baseline characteristics divided by TERS tertiles.

Variable	*N*	TERS ≤ 28 *N* = 80	TERS 29-31.5 *N* = 81	TERS ≥ 32 *N* = 79	χ^2 or *H*, *p*-value
TERS scores, *n* (%) or median (range)					
Age at 1 year post tx, median (range)	240	53 (21-71)	54 (20-66)	57 (22-67)	$H = 2.742$, $p = 0.254$
Sex, *n* (%)	240				
female	114	35 (43.8)	39 (48.1)	40 (50.6)	$\chi^2 = 0.776$ (df = 2), $p = 0.679$
male	126	45 (56.3)	42 (51.9)	39 (49.4)	
Educational level, *n* (%)	239				
<12 years	165	44 (55.0)	57 (71.3)	64 (81.0)	$\chi^2 = 12.858$ (df = 2), $p = 0.002$
≥12 years	74	36 (45.0)	23 (28.7)	15 (19.0)	
Partnership, *n* (%)	240				
yes	201	66 (82.5)	68 (84.0)	67 (84.8)	$\chi^2 = 0.159$ (df = 2), $p = 0.923$
No	39	14 (17.5)	13 (16.0)	12 (15.2)	
LAS-Category, *n* (%)	240				
Category A	79	10 (12.5)	34 (42.0)	35 (44.3)	$\chi^2 = 27.436$ (df = 6), $p < 0.001$
Category B	14	4 (5.0)	5 (6.2)	5 (6.3)	
Category C	45	20 (25.0)	17 (21.0)	8 (10.1)	
Category D	102	46 (57.5)	25 (30.9)	31 (39.2)	

Univariate analyses (Kruskal-Wallis H-tests, Chi square tests).
LAS Category, pulmonary diagnosis according to the Lung Allocation Score (Category A, obstructive airway diseases; category B, diseases of the pulmonary circulation, category C, infectious lung diseases, category D, restrictive lung diseases); TERS, Transplant Evaluation Rating Scale; post tx, post-transplantation.

percentage of patients listed in the low, intermediate and high TERS tertile were 86.1, 82.1, and 72.4% which was significantly different [$\chi^2 = 8.131$ (df = 2) $p = 0.017$; Cramer-V = 0.152]. Of those who died during the first year after transplantation ($n = 20$), 15% were in the low, 60% in the intermediate, and 25% in the high TERS tertile. This difference approached statistical

TABLE 3 | Comparison of baseline characteristics divided by the TERS subscales "defiance" and "emotional sensitivity" (median split).

Variable	*N*	Defiance ≤ 18.75 *N* = 120	Defiance > 18.75 *N* = 120	χ^2 or Z, *p*-value, effect size	Emotional ≤ 10 *N* =132	Emotional > 10 *N* = 108	χ^2 or *Z*, *p*-value, effect size
TERS subscales median split; *n* (%) or median (range)							
Age	240	51 (20-71)	58 (22-67)	$Z = -3.325$, $p = 0.001$	55 (21-71)	56 (20-67)	$Z = -0.481$, $p = 0.630$
Sex	240						
female		58 (48.3)	56 (46.7)	$\chi^2 = 0.067$ (df = 1), $p = 0.796$	58 (43.9)	56 (51.9)	1.491 (df = 1), $p = 0.222$
male		62 (51.7)	64 (52.5)		74 (56.1)	52 (48.1)	
Educational level	239						
<12 years	165	66 (55.5)	99 (82.5)	$\chi^2 = 20.434$ (df = 1), $p < 0.001$, $V = 0.292$	88 (66.7)	77 (72.0)	$\chi^2 = 0.775$ (df = 1); $p = 0.379$
≥12 years	74	53 (44.5)	21 (17.5)		44 (33.3)	30 (28.0)	
LAS-Category, *n* (%)	240						
Category A	79	18 (15)	61 (50.8)	$\chi^2 = 41.812$ (df = 3), $p < 0.001$, $V = 0.417$	39 (29.5)	40 (37.0)	$\chi^2 = 2.154$ (df = 3), $p = 0.541$
Category B	14	8 (6.7)	6 (5.0)		9 (6.8)	5 (4.6)	
Category C	45	36 (30.0)	9 (7.5)		24 (18.2)	21 (19.4)	
Category D	102	58 (48.3)	44 (36.7)		60 (45.5)	42 (38.9)	

Univariate analyses (Mann-Whitney tests, Chi square tests).

LAS Category, pulmonary diagnosis according to the Lung Allocation Score (Category A, obstructive airway diseases; category B, diseases of the pulmonary circulation, category C, infectious lung diseases, category D, restrictive lung diseases); TERS, Transplant Evaluation Rating Scale.

significance with a small effect size [$\chi^2 = 5.858$ (df = 2) $p = 0.053$; Cramer-V = 0.150].

One-year renal function (eGFR), forced expiratory volume in 1 s (FEV1) <80% in relation to the post-transplant baseline FEV1, number of hospitalizations during the first year, and the CCI were not different between TERS groups (**Table 2**). This was also true for the two TERS subscales (data not shown).

Behavioral Outcomes

Overall, 5% ($n = 12$) of the patients were obese (BMI ≥ 30 kg/m^2), 33.2% ($n = 79$) were overweight and 6.7% ($n = 16$) were underweight 1 year after transplantation. Most patients were in the normal-weight range (55%, $n = 131$). Patients in the higher TERS tertiles were more often obese (11.4%) and overweight (39.2%) (**Table 4**).

Overall, 45.8% exhibited an adherence score of <87% indicating suboptimal adherence to components of the medical regimen and transplant program recommendations. More patients in the highest TERS tertile were rated with an adherence score of <87% (58.2%) (**Table 4**).

Both associations (TERS with BMI and adherence, respectively) were mainly due to differences in the "defiance" subscale categories and not the "emotional sensitivity" subscale categories (**Table 5**). Logistic regression analysis adjusted for baseline variables confirmed these significant associations (**Supplementary Tables 1**, **2**). Looking at the individual components of non-adherence, low health perception (e.g., inconsistent medication knowledge, recommendations regarding substance abuse not met, poor diabetic control, use of sunbeds) and missed appointments with the transplant center were predicted by the TERS but not home spirometry frequency, nutrition and exercise, or trough levels outside the target range.

In patients with extended 1-year assessment, we also conducted a BAASIS interview. Even though the difference between TERS tertiles concerning adherent and non-adherent patients according to the BAASIS did not reach statistical significance, the effect size (Cramer-V = 0.189) was comparable to the effect size found for the differences with regard to our comprehensive adherence assessment (Cramer-V = 0.175) (**Table 6**).

Psychological Outcomes

One year after transplantation, only two patients exhibited a PHQ-4 score of 6 or above which is considered indicative for the presence of a depressive or anxiety disorder. Four patients scored 3 or above in the two-item depression subscale and two patients in the two-item anxiety subscale. 143 patients (60%) did not report any symptoms on the PHQ-4 with a total score of 0. PHQ-4 did not differ between TERS groups (**Table 4**); however, there were differences in the "emotional sensitivity" subscale, with patients with scores above the median exhibiting higher PHQ-4 scores (**Table 5**). This was confirmed by a logistic regression analysis controlling for baseline variables (**Supplementary Table 3**).

The QoL VAS exhibited median values around 8 and did not differ between TERS groups (**Table 4**).

Subsample With Extended 1-Year Assessment

143 patients received a more detailed assessment including the BAASIS interview (see above) and were asked to complete additional questionnaires including the PHQ-9, GAD-7, SF-8, and PQLS to complement the minimal psychosocial assessment with the PHQ-4 and the QoL VAS that are routinely completed by all patients.

This subgroup was fairly evenly distributed between the three TERS tertiles, with 44, 47, and 53 patients, respectively. Importantly, this subsample did not differ from the entire sample

TABLE 4 | Comparison of 1-year post-transplant outcomes divided by TERS tertiles ($n = 240$).

Variable	*N*	TERS ≤ 28 *N* = 80	TERS 29-31.5 *N* = 81	TERS ≥ 32 *N* = 79	χ^2 or *H*, *p*-value, effect size
TERS scores, *n* (%) or median (range)					
BMI kg/m^2, *n* (%)	238				
<18.5 (underweight)	16	7 (8.9)	4 (5.0)	5 (6.3)	$\chi^2 = 15.958$ (df = 6), $p = 0.014$, V = 0.183
18.5-24.9	131	49 (62.0)	48 (60.0)	34 (43.0)	
25-29.9 (overweight)	79	23 (29.1)	25 (31.1)	31 (39.2)	
≥30 (obesity)	12	0 (0)	3 (3.8)	9 (11.4)	
eGFR quartiles	240				
≤38.5	60	17 (21.3)	17 (21.0)	26 (32.9)	$\chi^2 = 7.798$ (df = 6), $p = 0.253$
38.6-54.5	60	18 (22.5)	25 (30.9)	17 (21.5)	
54.6-68.9	60	19 (23.8)	20 (24.7)	21 (26.6)	
≥69	60	26 (32.5)	19 (23.5)	15 (19.0)	
CCI	240				
CCI = 0	144	49 (61.3)	48 (59.3)	47 (59.5)	$\chi^2 = 0.079$ (df = 2), $p = 0.961$
CCI > 0	96	31 (38.8)	33 (40.7)	32 (40.5)	
No. of hospitalizations during first year after tx	240				
0	115	39 (48.8)	40 (49.4)	36 (45.6)	$\chi^2 = 0.266$ (df = 2), $p = 0.875$
≥1	125	41 (51.2)	41 (50.6)	43 (54.4)	
FEV1 %	240				
<80%	99	37 (46.3)	33 (40.7)	29 (36.7)	$\chi^2 = 1.506$ (df = 2) $p = 0.471$
≥80%	141	43 (53.8)	48 (59.3)	50 (63.3)	
Adherence score (mean during first year after tx)	240				
<87%	110	31 (38.8)	33 (40.7)	46 (58.2)	$\chi^2 = 7.351$ (df = 2), $p = 0.025$, $V = 0.175$
≥87%	130	49 (61.3)	48 (59.3)	33 (41.8)	
Health perception	240	0 (0-1)	0.1 (0-1)	0.3 (0-1.3)	$H = 14.936$, $p = 0.001$ $\eta^2 = 0.055$
Home spirometry frequency	240	0 (0-1.4)	0 (0-1)	0 (0-1.7)	$H = 3.261$, $p = 0.196$
Contact	240	0 (0-0.6)	0 (0-0.7)	0 (0-0.9)	$H = 4.439$, $p = 0.109$
Nutrition, exercise	240	0.3 (0-1.7)	0.2 (0-1.4)	0.4 (0-1.7)	$H = 5.019$, $p = 0.081$
Trough levels	240	0.6 (0-1.3)	0.6 (0-1.5)	0.6 (0-2)	$H = 0.94$, $p = 0.625$
Total score	240	1.2 (0-4.6)	1.2 (0-4.3)	1.4 (0.2-5.7)	$H = 3.865$, $p = 0.145$
Percentage	240	88.5 (53.8-100)	88.3 (56.7-100)	86 (43.3-98.3)	$H = 3.865$, $p = 0.145$
QoL VAS, median (range)	235	8 (3.5-10)	8.5 (1-10)	8 (2-10)	$H = 0.224$, $p = 0.894$
PHQ-4 total, median (range)	235	0 (0-5)	0 (0-8)	0 (0-12)	$H = 3.376$, $p = 0.185$
PHQ 4 median split	235				
<1	143	54 (68.4)	49 (62.0)	40 (51.9)	$\chi^2 = 4.475$ (df = 2), $p = 0.107$
≥1	92	25 (31.6)	30 (38.0)	37 (48.1)	

Univariate analyses (Kruskal-Wallis H-tests, Chi square tests).

BMI, Body Mass Index; CCI, Charlson Comorbidity Index; eGFR, estimated glomerular filtration rate; FEV1, forced expiratory volume in 1 s (% of first post tx value); PHQ-4, Patient Health Questionnaire-ultrashort version; QoL VAS, quality of life visual analog scale; TERS, Transplant Evaluation Rating Scale; tx, transplantation.

of 240 patients who completed the 1-year follow-up in any of the baseline data or outcomes (**Table 1**); thus, the subsample was most likely representative of the entire follow-up sample.

We found differences in PHQ-9 and GAD-7 scores between the TERS tertiles which were mainly based on differences in the "emotional sensitivity" subscale. The average levels of depression and anxiety tended to be low, with very few patients reporting scores of 10 points or above. A clear association of the TERS tertiles was found with the lung specific quality of life scale PQLS which was predicted by both subscales (**Tables 6**, **7**).

The SF-8 composite summary scores did not differ between TERS tertiles and were almost identical to reference values from the German general population (39). The median of the PCS for the entire sample ($n = 136$) was 52.8 (German population 53.6) and the median of the MCS was 57.2 (German population 57.3) (**Tables 6**, **7**).

TABLE 5 | Comparison of 1-year follow-up outcomes divided by the TERS subscales "defiance" and "emotional sensitivity" (median split).

Variable	*N*	Defiance ≤ 18.75, *N* = 120	Defiance > 18.75, *N* = 120	χ^2 or *Z*, *p*-value, effect size	Emotional ≤ 10, *N* = 132	Emotional > 10, *N* = 108	χ^2 or *Z*, *p*-value, effect size
TERS subscales median split; *n* (%) or median (range)							
BMI kg/m^2, *n* (%)	238						
<18.5	16	9 (7.6)	7 (5.9)	$X^2 = 17.013$ (df = 3), $p = 0.001$, $V = 0.267$	10 (7.7)	6 (5.6)	$\chi^2 = 4.834$ (df = 3), $p = 0.184$
18.5-24.9	131	78 (65.5)	53 (44.5)		74 (56.9)	57 (52.8)	
25-29.9	79	31 (26.1)	48 (40.3)		43 (33.1)	36 (33.3)	
≥30	12	1 (0.8)	11 (9.2)		3 (2.3)	9 (8.3)	
Adherence score (mean during first year after tx)	240						
<87%	110	42 (35.0)	68 (56.7)	$\chi^2 = 11.345$ (df = 1), $p = 0.001$, $V = -0.217$	60 (45.5)	50 (46.3)	$\chi^2 = 0.017$ (df = 1), $p = 0.896$
≥87%	130	78 (65.0)	52 (43.3)		72 (54.5)	58 (53.7)	
Adherence subscale scores (mean during first year after tx), median (range)	240						
Health perception	240	0 (0-1)	0.2 (0-1.3)	$Z = -3.102$, $p = 0.002$	0.1 (0-1)	0.2 (0-1.3)	$Z = -2.042$, $p = 0.041$
Home spirometry frequency	240	0 (0-1.4)	0 (0-1.7)	$Z = -1.297$, $p = 0.195$	0 (0-1.4)	0 (0-1.7)	$Z = -0.099$, $p = 0.921$
Contact	240	0 (0-0.7)	0 (0-0.9)	$Z = -2.490$, $p = 0.013$	0 (0-0.9)	0 (0-0.8)	$Z = -0.088$, $p = 0.930$
Nutrition, exercise	240	0.3 (0-1.7)	0.3 (0-1.7)	$Z = -1.322$, $p = 0.186$	0.3 (0-1.7)	0.3 (0-1.7)	$Z = -1.009$, $p = 0.313$
Trough levels	240	0.6 (0-5)	0.6 (0-1.4)	$Z = -0.862$, $p = 0.389$	0.6 (0-1.4)	0.7 (0-2)	$Z = -0.252$, $p = 0.801$
Total score	240	1 (0-5)	1.4 (0-5.7)	$Z = -1.869$, $p = 0.062$	1.3 (0-5.4)	1.2 (0.2-5.7)	$Z = -0.719$, $p = 0.472$
Percentage	240	90 (50-100)	86 (43.3-100)	$Z = -1.869$, $p = 0.062$	87.5 (45.7-100)	88 (43.3-98.3)	$Z = -0.719$, $p = 0.472$
Qol VAS, median (range)	235	8 (1-10)	8 (2-10)	$Z = -0.275$, $p = 0.784$	8 (2-10)	8 (1-10)	$Z = -0.176$, $p = 0.860$
PHQ 4 total (0-12), median (range)	235	0 (0-8)	0 (0-12)	$Z = -0.240$, $p = 0.911$	0 (0-5)	0 (0-12)	$Z = -2.477$, $p = 0.013$
PHQ 4 median split	235						
<1	143	74 (62.7)	69 (59.0)	$\chi^2 = 0.344$ (df = 1), $p = 0.557$	89 (67.9)	54 (51.9)	$\chi^2 = 6.242$ (df = 1), $p = 0.012$, V = 0.163
≥1	92	44 (37.3)	48 (41.0)		42 (32.1)	50 (48.1)	

Univariate analyses (Mann-Whitney tests, Chi square tests).

BMI, Body Mass Index; CCI, Charlson Comorbidity Index; eGFR, estimated glomerular filtration rate; FEV1, forced expiratory volume in 1 second; PHQ-4, Patient Health Questionnaire-ultrashort version; QoL VAS, quality of life visual analog scale; TERS, Transplant Evaluation Rating Scale; tx, transplantation; V, Cramer V (effect size).

TABLE 6 | Comparison of extended 1-year post-transplant outcomes according to TERS tertiles ($n = 143$).

Variable	*N*	TERS ≤ 28 *N* = 80	TERS 29-31,5 *N* = 81	TERS ≥ 32 *N* =79	χ^2 or *H*, *p*-value, effect size
TERS scores, *n* (%) or median (range)					
BAASIS at 1 year	143				
Adherent	125	38 (86.4)	45 (95.7)	42 (80.8)	$\chi^2 = 5.095$ (df = 2), $p = 0.078$, V = 0.189
Not adherent	18	6 (13.6)	2 (4.3)	10 (19.2)	
SF-8 PCS, median (range)	136	55.7 (20.3-61.1)	52.4 (23.7-61.8)	51 (25.5-58.6)	$H = 4.358$, $p = 0.113$
SF8 MCS, median (range)	136	57.5 (35-62.8)	57.2 (30.8-61.9)	57.2 (23.2-66.9)	$H = 0.775$, $p = 0.679$
PHQ 9 total, median (range)	137	1 (0-12)	3 (0-16)	4 (0-16)	H=7.674, p=.022, η^2=.042
PHQ-9 cutoffs	137				
0-4	99	35 (81.4)	35 (77.8)	29 (59.2)	$\chi^2 = 8.159$ (df = 6), $p = 0.227$
5-9 (mild)	30	6 (14.0)	8 (17.8)	16 (32.7)	
10-14 (moderate)	5	2 (4.7)	1 (2.2)	2 (4.1)	
15-27 (severe)	3	0 (0)	1 (2.2)	2 (4.1)	
GAD 7 total, median (range)	136	0 (0-19)	1 (0-16)	1 (0-16)	$H = 2.868$, $p = 0.238$
GAD 7 cutoffs	136				
0-4	116	39 (90.7)	37 (82.2)	40 (83.3)	$\chi^2 = 5.658$ (df = 6), $p = 0.463$
5-9 (mild)	16	4 (9.3)	7 (15.6)	5 (10.4)	
10-14 (moderate)	2	0	0	2 (4.2)	
15-21 (severe)	2	0	1 (2.2)	1 (2.1)	
PQLS total, median (range)*	114	34 (25-74)	40.5 (25-82)	50 (26-83)	$H = 12.018$, $p = 0.002$, $\eta^2 = 0.09$
Task interference	115	11.2 (8-28)	14 (8-28)	20 (8-38)	$H = 14.040$, $p = 0.001$, $\eta^2 = 0.108$
Psychological functioning	127	9 (7-21)	10 (7-35)	9 (7-25)	$H = 0.448$, $p = 0.799$
Physical functioning	134	4 (4-18)	5 (4-20)	7 (4-20)	$H = 12.621$, $p = 0.002$ $\eta^2 = 0.081$

Univariate analyses (Kruskal-Wallis H-tests, Chi square tests).

BAASIS, Basel Assessment of Adherence to Immunosuppressive Medication Scale; GAD-7, Generalized Anxiety Disorder Screener; PHQ-9, Patient Health Questionnaire-Depression; PQLS, Pulmonary-Specific Quality-of-Life Scale; SF-8 MCS, Short Form 8 Mental Component Summary; SF-8 PCS, Short Form 8 Physical Component Summary; TERS, Transplant Evaluation Rating Scale.

**Part of the PQLS data were published previously in manuscripts assessing the validity of the TERS and of the German version of the PQLS (21, 37).*

DISCUSSION

In this large prospective analysis it could be demonstrated that psychosocial factors as measured by TERS score are predictive of 1-year transplantation outcomes. Patients with lower psychosocial risk were more likely to be listed. The TERS was predictive of behavioral outcomes such as the BMI, adherence, and psychological outcomes such as levels of depression and anxiety, and lung-specific quality of life at 1-year follow-up. The TERS subscales "defiance" and "emotional sensitivity" showed differential predictive validity. While the "defiance" scale score was associated with behavioral outcomes, the "emotional sensitivity" subscale score was predictive for psychological outcomes. Thus, our results support the assumption put forward by Hoodin and Kalbleisch (20) that the TERS is actually a multifaceted construct composed of two subordinate constructs. While related to each other empirically and logically, the two subscales can and should be distinguished conceptually and measured separately.

Medical Outcomes

Even though the prediction of mortality during the first year after transplantation approached significance, this result should not be over interpreted. Mortality rate during the first year after transplantation was low with 20 patients (7.4%) and chronic lung allograft dysfunction is generally rare during the first year. The few studies that reported the association of TERS scores with mortality and graft functioning included markedly longer follow up periods (10, 16). In longer follow-up examinations, mortality should be used as a time dependent variable instead as a binary outcome. Some patients with high-risk TERS scores who were considered unfit for transplantation may not have been offered transplantation. Differences in TERS scores between listed and not listed patients support this. Unfortunately, the data for such patients are not captured in our database.

None of the other medical outcomes were predicted by TERS tertile scores. Most likely, medical issues during the first year after transplantation are predominantly influenced by the transplant

TABLE 7 | Comparison of extended 1-year follow-up outcomes divided by the TERS subscales "defiance" and "emotional sensitivity" (median split).

Variable	*N*	Defiance ≤ 18.75, *N* = 120	Defiance > 18.75, *N* = 120	χ^2 or *Z*, *p*-value, effect size	Emotional ≤ 10, *N* = 132	Emotional > 10, *N* = 108	χ^2 or *Z*, *p*-value, effect size
TERS subscales median split; *n* (%) or median (range)							
BAASIS at one year	143						
Adherent	125	62 (91.2)	63 (84.0)	1.669 (df = 1), $p = 0.196$, $V = 0.108$	65 (86.7)	60 (88.2)	0.080 (df = 1), $p = 0.778$
Not adherent	18	6 (8.8)	12 (16.0)		10 (13.3)	8 (11.8)	
SF-8 PCS, median (range)	136	53.5 (20.3-61.8)	52.4 (25.5-58.8)	$Z = -1.519$, $p = 0.129$	54.4 (20.3-61.1)	51.9 (23.4-61.8)	$Z = -1.229$, $p = 0.219$
SF-8 MCS, median (range)	136	56.7 (30.8-62.8)	57.4 (23.3-66.9)	$Z = -0.508$, $p = 0.612$	57.5 (28.9-62.8)	57.0 (66.9-43.7)	$Z = -1.412$, $p = 0.158$
PHQ 9 total, median (range)	137	2 (0-16)	3 (0-16)	$Z = -0.839$, $p = 0.402$	2 (0-13.5)	3.5 (0-16)	$Z = -3.119$, $p = 0.002$ $\eta^2 = 0.069$
PHQ-9 cutoff	137						
0-4	99	50 (75.8)	49 (69.0)	$\chi^2 = 1.563$ (df = 3), $p = 0.668$, $V = 0.107$	59 (80.8)	40 (62.5)	$\chi^2 = 8.425$ (df = 3), $p = 0.038$, $V = 0.248$
5-9 (mild)	30	12 (18.2)	18 (25.4)		11 (15.1)	19 (29.7)	
10-14 (moderate)	5	3 (4.5)	2 (2.8)		3 (4.1)	2 (3.1)	
15-27 (severe)	3	1 (1.5)	2 (2.8)		0 (0)	3 (4.7)	
GAD 7 total, median (range)	136	1 (0-16)	1 (0-16)	$Z = -0.236$, $p = 0.814$	0 (0-11)	1 (0-16)	$Z = -2.157$, $p = 0.031$ $\eta^2 = 0.03$
GAD 7 cutoff	136						
0-4	116	54 (81.8)	62 (88.6)	$\chi^2 = 4.688$ (df = 3), $p = 0.196$, $V = 0.186$	66 (90.4)	50 (79.4)	$\chi^2 = 4.496$ (df = 3), $p = 0.213$, $V = 0.182$
5-9 (mild)	16	11 (16.7)	5 (7.1)		6 (8.2)	10 (15.9)	
10-14 (moderate)	2	0 (0)	2 (2.9)		1 (1.4)	1 (1.6)	
15-21 (severe)	2	1 (1.5)	1 (1.4)		0 (0)	2 (3.2)	
PQLS total, median (range)	114	38 (25.0-80.4)	46 (25.0-83.0)	$Z = -2.276$ $p = 0.023$ $\eta^2 = 0.045$	37 (25-83)	48 (26-80.4)	$Z = -3.136$, $p = 0.002$ $\eta^2 = 0.086$
Task interference	115	14 (8-33)	17.6 (8-38)	$Z = -2.026$, $p = 0.043$ $\eta^2 = 0.035$	12.5 (8-30)	18 (8-38)	$Z = -3.467$. $p = 0.001$ $\eta^2 = 0.104$
Psychological functioning	127	9 (7-22)	9 (7-35)	$Z = -0.816$, $p = 0.415$	8.6 (7-35)	9.3 (7-25)	$Z = -1.037$, $p = 0.300$
Physical functioning	134	4 (4-20)	7 (4-20)	$Z = -3.554$, $p < 0.001$ $\eta^2 = 0.09$	5 (4-20)	6 (4-20)	$Z = -1.228$, $p = 0.219$

Univariate analyses (Mann-Whitney tests, Chi square tests).

BAASIS, Basel Assessment of Adherence to Immunosuppressive Medication Scale; GAD-7, Generalized Anxiety Disorder Screener; PHQ-9, Patient Health Questionnaire-Depression; PQLS, Pulmonary-Specific Quality-of-Life Scale; SF-8 MCS, Short Form 8 Mental Component Summary; SF-8 PCS, Short Form 8 Physical Component Summary; TERS, Transplant Evaluation Rating Scale; V, Cramer V (effect size).

process itself rather than by psychosocial issues. This might change, however, in the long run.

Behavioral Outcomes

Non-adherence to the medical regimen after transplantation can contribute to poor clinical outcomes (26, 41, 42). Adherence is not only important regarding medication-taking but after lung transplantation also with regard to the regular use of spirometry and other clinical care requirements, regular visits to the transplant center, and lifestyle activities such as nutrition and exercise. Thus, we used a composite adherence measure that has shown to predict mortality and graft loss in a large sample of lung transplant patients (26). The TERS, specifically the "defiance" subscale, was predictive of suboptimal adherence during the first post-transplant year in our sample. Looking at the individual components of non-adherence, especially health perception and regular contacts with the transplant center were predicted by the TERS but not home spirometry frequency, nutrition and exercise or trough levels. Looking at the BAASIS, which was used as an interview and focuses exclusively on adherence to immunosuppressive medication during the last 4 weeks, we found no differences between TERS groups. However, the non-adherence rate was low (12.6%) and the BAASIS covers only a short time period of 4 weeks. Additionally, it has to be kept in mind that suboptimal adherence increases with increasing time since transplantation. In a large sample of lung transplant patients Drick et al. (27) could demonstrate that 37% of all non-adherent patients were transplanted ≥8 years prior to BAASIS assessment. Thus, during the first year adherence is usually higher and will most likely decline with time, which has also been shown in other solid organ transplantation samples (42). The predictive ability of the TERS might be stronger in the longer term after transplantation.

While no association was found between BMI category and TERS tertiles pre-surgery, TERS predicted BMI category at 1-year. It is well-known that obesity is an independent risk factor for mortality and transplant failure after lung transplantation (43). A systematic review and meta-analysis (44) clearly demonstrated that among post-lung transplant recipients underweight and obesity before transplantation were significantly associated with higher mortality and that obesity and overweight were associated with a higher risk of primary graft dysfunction compared to recipients who have normal BMI. A large US registry study including >17,000 patients, confirmed these results and additionally found that BMI increase and decrease from a baseline BMI with the lowest probability of death incrementally increased the odds of mortality at 90 days and 1 year after transplantation (45). The mechanisms are not entirely clear; however, *via* mechanical and probably metabolic effects, lung mechanics are altered in the presence of obesity (46). Weight loss before transplantation was associated with improved short- and long-term clinical outcomes, independent of initial weight (47), and a first case reports describes the successful bariatric surgery in a young women with a BMI of 53.6 kg/m^2 4 years after lung transplantation (48).

Taken together, successfully predicting behavioral outcomes such as adherence to a broad range of medical regimens and unfavorable weight developments might be pivotal for mortality and morbidity in lung transplant patients.

Psychological Outcomes

Even though survival is the key outcome, patients' post-transplant quality of life has become an important component of any evaluation of benefits, specifically as survival times increase (22, 49–52).

As shown in our study, higher levels of pre-transplant "emotional sensitivity" scores might be predictive of lower pulmonary-specific quality of life after transplantation. The PQLS total and subscale values were comparable to the values from the original validation study of the PQLS that provided data at 6 months after transplantation (25). HRQoL was also measured with a generic instrument, the SF-8. TERS tertiles were not predictive of SF-8 subscales; however, in line with other studies in transplant populations, the two subscales—PCS and MCS—have reached values that were comparable to the reference values of the German general population despite differences in life expectancy, treatment-related side effects, and despite the fact that patients after lung transplantation have persistent disabilities (50, 52). Specifically during the first year after transplantation patients experience a substantial benefit from the transplant procedure. Longer-term follow-up will show if we will discover a HRQoL decline after the first post-operative year also in our sample and if this decline is predicted only by co-morbid medical conditions or also by pre-transplant TERS scores.

Comparable to HRQoL measures, depression and anxiety scores were quite low 1 year after transplantation. Again, higher levels of pre-transplant "emotional sensitivity" scores were predictive of higher depression and anxiety scores after transplantation. Overall, predicting post-transplant symptoms of depression might be more important than the presence of pre-transplant mental comorbidity. A meta-analysis on the effect of pre-transplant depression and anxiety on survival following lung transplant (53) did not find that depression or anxiety scores pre-transplant were associated with worse survival. Thus, the presence of affective or anxiety symptoms in a prospective candidate should not be the basis of exclusion from consideration for lung transplantation. However, others found that pre-transplant depression might be a predictor of survival in subgroups of patients (12) and that specifically persistent depression (11) and early post-transplant depressive symptoms might be predictors of long-term outcomes compared with pre-transplant psychosocial assessment alone (7). Smith et al. (7) reported that higher levels of depression and general distress measured 6 months following lung transplantation were associated with increased mortality, independent of baseline characteristics and medical predictors. Also others confirmed that early post-transplant depressive symptoms increase the risk for long-term transplant-related morbidity and mortality (54). Thus, attention should be paid to post-transplant depressive symptoms and putative predictors of the development. Finally, if treatment of comorbid mental disorders would reduce post-transplant mortality requires further study (55).

New psychosocial assessment tools such as the Stanford Integrated Psychosocial Assessment for Transplantation

(SIPAT) (56) have been developed and are increasingly used internationally. The SIPAT comprises 18 psychosocial risk factors, which are divided into four domains. The SIPAT has shown to have good interrater reliability (0.85) and to be predictive of medical and psychosocial post-transplant outcomes in a mixed group of organ transplant recipients (57) including rejection episodes, medical hospitalization, infection rates, psychiatric decompensation, and support system failure. They also reported a trend concerning the relationship with non-adherence. As in our study, effect sizes were small to moderate.

Limitations

Our data are based on a relatively modest sized cohort from a single center with follow-up so far only over 1 year. While our study is the largest to focus on the predictive value of the TERS on multiple post-transplant outcomes, its limitations in size and duration nonetheless are relevant. It has to be kept in mind that patients who get listed and transplanted represent a selected population since transplant centers have to concur with the guidelines of the German Medical Association, which includes the LAS score. In Germany, we also follow the recommendations of International Society for Heart and Lung Transplantation (ISHLT) (58). This must be considered when comparing studies. Psychosocially, our sample was fairly healthy, with overall low TERS values and low mortality rates as well as low levels of depression and anxiety at baseline and at follow-up. Thus, distribution problems caused by considerable ceiling and floor effects, respectively, prevail. Nevertheless, we found significant associations between the pre-transplant TERS and several post-transplant outcomes.

Quantiles are frequently used to facilitate communication of the results to the public and other scientists. Even though the use of quantiles (in our study tertiles) remains highly common in epidemiological research, important problems arise when continuous variables (TERS scores) are categorized, particularly if data dependent cut points are used to form categories (59). Additionally, categorization involves multiple hypothesis testing and assumes homogeneity of risk within groups.

Conclusion

Our results confirm and extend prior evidence suggesting that psychosocial factors as measured with the TERS may predict medical, behavioral and psychological outcomes following lung transplantation, even during the first post-transplant year.

These findings can have several consequences: Higher psychosocial risk might (1) contribute to the determination of transplant candidacy, (2) lead to interventions prior to listing to reduce or minimize psychosocial risk (e.g., achieve smoking cessation, stabilize mood disorder, strengthen support system), and/or (3) might lead to increased clinical attention ("red flags") throughout the transplantation process and guide proactive interventions. Transplant physicians and mental health professionals should discuss the interventions required to be able to safely offer transplantation (10, 16) and the behavioral interventions necessary to avoid or minimize behavioral complications. Longer-term prospective follow-ups are needed since during the first post-transplant year the predictive impact of psychosocial risk factors might differ from that during consecutive years.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Ethics Board of Hannover Medical School. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MZ and JG designed the study. MZ and MB-H were mainly responsible for data acquisition. MZ, CV, and MN analyzed the data. MZ and MN wrote the first draft. All authors contributed significantly to the interpretation of the data and the final version of the manuscript and gave final approval of the version to be published.

FUNDING

The study was supported by a grant from Biomedical Research in End-stage and Obstructive Lung Disease Hannover (BREATH) within the German Center for Lung Research (82DZL002A1). This work was supported in part by the Ellen Schmidt Habilitation Program of Hannover Medical School for Mariel Nöhre.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2021.704319/full#supplementary-material

REFERENCES

1. Yusen RD, Edwards LB, Kucheryavaya AY, Benden C, Dipchand AI, Goldfarb SB, et al. The registry of the international society for heart and lung transplantation: thirty-second official adult lung and heart-lung transplantation report—2015; focus theme: early graft failure. *J Heart Lung Transplant.* (2015) 34:1264-77. doi: 10.1016/j.healun.2015.08.014
2. Gottlieb J, Smits J, Schramm R, Langer F, Buhl R, Witt C, et al. Lung transplantation in Germany since the introduction of the Lung Allocation Score. *Dtsch Arztebl Int.* (2017) 114:179-85. doi: 10.3238/arztebl.2017.0179

3. Parekh PI, Blumenthal JA, Babyak MA, Merrill K, Carney RM, Davis RD, et al. Psychiatric disorder and quality of life in patients awaiting lung transplantation. *Chest.* (2003) 124:1682-8. doi: 10.1378/chest.124.5.1682
4. Dew MA, DiMartini AF, DeVito Dabbs AJ, Fox KR, Myaskovsky L, Posluszny DM, et al. Onset and risk factors for anxiety and depression during the first 2 years after lung transplantation. *Gen Hosp Psychiatry.* (2012) 34:127-38. doi: 10.1016/j.genhosppsych.2011.11.009
5. Søyseth TS, Lund MB, Bjørtuft Ø, Heldal A, Søyseth V, Dew MA, et al. Psychiatric disorders and psychological distress in patients undergoing evaluation for lung transplantation: a national cohort study. *Gen Hosp Psychiatry.* (2016) 42:67-73. doi: 10.1016/j.genhosppsych.2016.07.001
6. Evon DM, Burker EJ, Galanko JA, Dedert E, Egan TM. Depressive symptoms and mortality in lung transplant. *Clin Transplant.* (2010) 24:E201-6. doi: 10.1111/j.1399-0012.2010.01236.x
7. Smith P, Blumenthal J, Trulock E, Freedland K, Carney R, Davis R, et al. Psychosocial predictors of mortality following lung transplantation. *Am J Transplant.* (2016) 16:271-7. doi: 10.1111/ajt.13447
8. Dew MA, DiMartini AF, Dobbels F, Grady KL, Jowsey-Gregoire SG, Kaan A, et al. The 2018 ISHLT/APM/AST/ICCAC/STSW recommendations for the psychosocial evaluation of adult cardiothoracic transplant candidates and candidates for long-term mechanical circulatory support. *J Heart Lung Transplant.* (2018) 37:803-23. doi: 10.1016/j.psym.2018.04.003
9. Woodman CL, Geist LJ, Vance S, Laxson C, Jones K, Kline JN. Psychiatric disorders and survival after lung transplantation. *Psychosomatics.* (1999) 40:293-7. doi: 10.1016/S0033-3182(99)71221-1
10. Hitschfeld MJ, Schneekloth TD, Kennedy CC, Rummans TA, Niazi SK, Vasquez AR, et al. The psychosocial assessment of candidates for transplantation: a cohort study of its association with survival among lung transplant recipients. *Psychosomatics.* (2016) 57:489-97. doi: 10.1016/j.psym.2016.05.003
11. Smith PJ, Rivelli S, Waters A, Reynolds J, Hoyle A, Flowers M, et al. Neurocognitive changes after lung transplantation. *Ann Am Thorac Soc.* (2014) 11:1520-7. doi: 10.1513/AnnalsATS.201406-232OC
12. Smith PJ, Snyder LD, Palmer SM, Hoffman BM, Stonerock GL, Ingle KK, et al. Depression, social support, and clinical outcomes following lung transplantation: a single-center cohort study. *Transpl Int.* (2018) 31:495-502. doi: 10.1111/tri.13094
13. Bundesärztekammer. Richtlinie gemäß § 16 Abs. 1 S. 1 Nrn. 2 u. 5 TPG für die Wartelistenführung und Organvermittlung zur Lungentransplantation. *Dtsch Arztebl.* (2017) 114:A-1948/B-1648/C-1614. doi: 10.3238/arztebl.2017.rili_baek_OrgaWlOvLeberTx20170616
14. Twillman RK, Manetto C, Wellisch DK, Wolcott DL. The transplant evaluation rating scale: a revision of the psychosocial levels system for evaluating organ transplant candidates. *Psychosomatics.* (1993) 34:144-53. doi: 10.1016/S0033-3182(93)71905-2
15. Erim Y, Beckmann M, Marggraf G, Senf W. Psychosomatic evaluation of patients awaiting lung transplantation. *Transplant Proc.* (2009) 41:2595-8. doi: 10.1016/j.transproceed.2009.06.125
16. Solh MM, Speckhart D, Solomon SR, Bashey A, Morris LE, Zhang X, et al. The transplant evaluation rating scale predicts overall survival after allogeneic hematopoietic stem cell transplantation. *Blood Adv.* (2020) 4:4812-21. doi: 10.1182/bloodadvances.2020002204
17. Gazdag G, Horváth GG, Makara M, Ungvari GS, Gerlei Z. Predictive value of psychosocial assessment for the mortality of patients waiting for liver transplantation. *Psychol Health Med.* (2016) 21:525-9. doi: 10.1080/13548506.2015.1109670
18. Yost GL, Bhat G, Ibrahim KN, Karountzos AG, Chandrasekaran M, Mahoney E. Psychosocial evaluation in patients undergoing left ventricular assist device implantation using the Transplant Evaluation Rating Scale. *Psychosomatics.* (2016) 57:41-6. doi: 10.1016/j.psym.2015.07.013
19. Vitinius F, Reklat A, Hellmich M, Klask E, Wahlers T, Rahmanian PB, et al. Prediction of survival on the waiting list for heart transplantation and of posttransplant nonadherence-Results of a prospective longitudinal study. *Clin Transplant.* (2019) 33:e13616. doi: 10.1111/ctr.13616
20. Hoodin F, Kalbfleisch KR. Factor analysis and validity of the Transplant Evaluation Rating Scale in a large bone marrow transplant sample. *J Psychosom Res.* (2003) 54:465-73. doi: 10.1016/S0022-3999(02)00413-0
21. Nöhre M, Paslakis G, Albayrak Ö, Bauer-Hohmann M, Brederecke J, Eser-Valeri D, et al. Factor analyses and validity of the Transplant Evaluation Rating Scale (TERS) in a large sample of lung transplant candidates. *Front Psychiatry.* (2020) 11:373. doi: 10.3389/fpsyt.2020.00373
22. Kolaitis NA, Singer JP. Defining success in lung transplantation: From survival to quality of life. *Semin Respir Crit Care Med.* (2018) 39:255-68. doi: 10.1055/s-0037-1615801
23. Kugler C, Gottlieb J, Warnecke G, Schwarz A, Weissenborn K, Barg-Hock H, et al. Health-related quality of life after solid organ transplantation: a prospective, multiorgan cohort study. *Transplantation.* (2013) 96:316-23. doi: 10.1097/TP.0b013e31829853eb
24. Singer JP, Chen J, Blanc PD, Leard LE, Kukreja J, Chen H. A thematic analysis of quality of life in lung transplant: the existing evidence and implications for future directions. *Am J Transplant.* (2013) 13:839-50. doi: 10.1111/ajt.12174
25. Hoffman BM, Stonerock GL, Smith PJ, O'Hayer CV, Palmer S, Davis RD, et al. Development and psychometric properties of the Pulmonary-specific quality-of-life scale in lung transplant patients. *J Heart Lung Transplant.* (2015) 34:1058-65. doi: 10.1016/j.healun.2015.03.005
26. Bertram A, Fuge J, Suhling H, Tudorache I, Haverich A, Welte T, et al. Adherence is associated with a favorable outcome after lung transplantation. *PLoS ONE.* (2019) 14:e0226167. doi: 10.1371/journal.pone.0226167
27. Drick N, Seeliger B, Fuge J, Tudorache I, Greer M, Welte T, et al. Self-reported non-adherence to immunosuppressive medication in adult lung transplant recipients-A single-center cross-sectional study. *Clin Transplant.* (2018) 32:e13214. doi: 10.1111/ctr.13214
28. Wittchen H-U, Zaudig M, Fydrich T. *Strukturiertes Klinisches Interview für DSM-IV*. Göttingen: Hogrefe (1997).
29. Nghiem DM, Gomez J, Gloston GF, Torres DS, Marek RJ. Psychological assessment instruments for use in liver and kidney transplant evaluations: scarcity of evidence and recommendations. *J Pers Assess.* (2020) 102:183-95. doi: 10.1080/00223891.2019.1694527
30. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* (1987) 40:373-83. doi: 10.1016/0021-9681(87)90171-8
31. Dobbels F, Berben L, De Geest S, Drent G, Lennerling A, Whittaker C, et al. The psychometric properties and practicability of self-report instruments to identify medication nonadherence in adult transplant patients: a systematic review. *Transplantation.* (2010) 90:205-19. doi: 10.1097/TP.0b013e3181e346cd
32. Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affect Disord.* (2010) 122:86-95. doi: 10.1016/j.jad.2009.06.019
33. Löwe B, Gräfe K, Zipfel S, Witte S, Loerch B, Herzog W. Diagnosing ICD-10 depressive episodes: superior criterion validity of the Patient Health Questionnaire. *Psychother Psychosom.* (2004) 73:386-90. doi: 10.1159/000080393
34. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study: primary care evaluation of mental disorders. Patient health questionnaire. *JAMA.* (1999) 282:1737–44. doi: 10.1001/jama.282.18.1737
35. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* (2006) 166:1092–97. doi: 10.1001/archinte.166.10.1092
36. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. *Med Care.* (2008) 46:266–74. doi: 10.1097/MLR.0b013e318160d093
37. Nöhre M, Albayrak Ö, Brederecke J, Claes L, Smits D, Tudorache I, et al. Psychometric properties of the German version of the Pulmonary-Specific Quality-of-Life Scale in lung transplant patients. *Front Psychiatry.* (2019) 10:374. doi: 10.3389/fpsyt.2019.00374
38. Ware J, Kosinski M, Dewey J. *How to Score and Interpret Single-Item Health Status Measures: A Manual for Users of the SF-8^{TM} Health Survey.* Lincoln: Quality Metrics (2001).

39. Beierlein V, Morfeld M, Bergelt C, Bullinger M, Brähler E. Messung der gesundheitsbezogenen Lebensqualität mit dem SF-8. *Diagnostica.* (2012) 58:145-53. doi: 10.1026/0012-1924/a000068
40. Cohen J. *Statistical Power Analysis for Behavioral Sciences.* 2nd ed. New York, NY: Lawrence Erlbaum Associates (1988).
41. Sellarés J, de Freitas DG, Mengel M, Reeve J, Einecke G, Sis B, et al. Understanding the causes of kidney transplant failure: the dominant role of antibody-mediated rejection and nonadherence. *Am J Transplant.* (2012) 12:388-99. doi: 10.1111/j.1600-6143.2011.03840.x
42. Dew MA, Posluszny DM, DiMartini AF, Myaskovsky L, Steel JL, DeVito Dabbs AJ. Posttransplant medical adherence: what have we learned and can we do better? *Curr Transplant Rep.* (2018) 5:174-88. doi: 10.1007/s40472-018-0195-8
43. Lederer DJ, Kawut SM, Wickersham N, Winterbottom C, Bhorade S, Palmer SM, et al. Obesity and primary graft dysfunction after lung transplantation: the Lung Transplant Outcomes Group Obesity Study. *Am J Respir Crit Care Med.* (2011) 184:1055-61. doi: 10.1164/rccm.201104-0728OC
44. Upala S, Panichsillapakit T, Wijarnpreecha K, Jaruvongvanich V, Sanguankeo A. Underweight and obesity increase the risk of mortality after lung transplantation: a systematic review and meta-analysis. *Transpl Int.* (2016) 29:285-96. doi: 10.1111/tri.12721
45. Fernandez R, Safaeinili N, Kurihara C, Odell DD, Jain M, DeCamp MM, et al. Association of body mass index with lung transplantation survival in the United States following implementation of the lung allocation score. *J Thorac Cardiovasc Surg.* (2018) 155:1871-9. doi: 10.1016/j.jtcvs.2017.11.031
46. Brock JM, Billeter A, Müller-Stich BP, Herth F. Obesity and the lung: what we know today. *Respiration.* (2020) 99:856-66. doi: 10.1159/000509735
47. Clausen ES, Frankel C, Palmer SM, Snyder LD, Smith PJ. Pre-transplant weight loss and clinical outcomes after lung transplantation. *J Heart Lung Transplant.* (2018) 37:1443-7. doi: 10.1016/j.healun.2018.07.015
48. El Moussaoui I, De Pauw V, Navez J, Closset J. Roux-En-Y gastric bypass after lung transplantation: case report and literature review. *Surg Obes Relat Dis.* (2021) 17:239-41. doi: 10.1016/j.soard.2020.10.007
49. Singer JP, Katz PP, Soong A, Shrestha P, Huang D, Ho J, et al. Effect of lung transplantation on health-related quality of life in the era of the lung allocation score: A U.S. prospective cohort study. *Am J Transplant.* (2017) 17:1334-45. doi: 10.1111/ajt.14081
50. Seiler A, Klaghofer R, Ture M, Komossa K, Martin-Soelch C, Jenewein J, et al. A systematic review of health-related quality of life and psychological outcomes after lung transplantation. *J Heart Lung Transplant.* (2016) 35:195-202. doi: 10.1016/j.healun.2015.07.003
51. Kugler C, Fischer S, Gottlieb J, Welte T, Simon A, Haverich A, et al. Health-related quality of life in two hundred-eighty lung transplant recipients. *J Heart Lung Transplant.* (2005) 24:2262-8. doi: 10.1016/j.healun.2005.07.005
52. Singer JP, Singer LG. Quality of life in lung transplantation. *Semin Respir Crit Care Med.* (2013) 34:421-30. doi: 10.1055/s-0033-1348470
53. Courtwright AM, Salomon S, Lehmann LS, Wolfe DJ, Goldberg HJ. The effect of pretransplant depression and anxiety on survival following lung transplant: a meta-analysis. *Psychosomatics.* (2016) 57:238-45. doi: 10.1016/j.psym.2015.12.008
54. Rosenberger EM, DiMartini AF, DeVito Dabbs AJ, Bermudez CA, Pilewski JM, Toyoda Y, et al. Psychiatric predictors of long-term transplant-related outcomes in lung transplant recipients. *Transplantation.* (2016) 100:239-47. doi: 10.1097/TP.0000000000000824
55. Dew MA, Rosenberger EM, Myaskovsky L, DiMartini AF, DeVito Dabbs AJ, Posluszny DM, et al. Depression and anxiety as risk factors for morbidity and mortality after organ transplantation: a systematic review and meta-analysis. *Transplantation.* (2015) 100:988-1003. doi: 10.1097/TP.0000000000000901
56. Maldonado JR, Dubois HC, David EE, Sher Y, Lolak S, Dyal J, et al. The Stanford Integrated Psychosocial Assessment for Transplantation (SIPAT): a new tool for the psychosocial evaluation of pre-transplant candidates. *Psychosomatics.* (2012) 53:123-32. doi: 10.1016/j.psym.2011.12.012
57. Maldonado JR, Sher Y, Lolak S, Swendsen H, Skibola D, Neri E, et al. The Stanford Integrated Psychosocial Assessment for Transplantation: a prospective study of medical and psychosocial outcomes. *Psychosom Med.* (2015) 77:1018-30. doi: 10.1097/PSY.0000000000000241
58. Weill D, Benden C, Corris PA, Dark JH, Davis RD, Keshavjee S, et al. A consensus document for the selection of lung transplant candidates: 2014–an update from the Pulmonary Transplantation Council of the International Society for Heart and Lung Transplantation. *J Heart Lung Transplant.* (2015) 34:1-15. doi: 10.1016/j.healun.2014.06.014
59. Bennette C, Vickers A. Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents. *BMC Med Res Methodol.* (2012) 12:21. doi: 10.1186/1471-2288-12-21

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychiatry

ORIGINAL RESEARCH
published: 27 May 2022
doi: 10.3389/fpsyt.2022.873693

A Randomized Clinical Trial Comparing Two Treatment Strategies, Evaluating the Meaningfulness of HAM-D Rating Scale in Patients With Major Depressive Disorder

Junaid Asghar[1], Madiha Tabasam[1], Maha M. Althobaiti[2], Amal Adnan Ashour[3], Mohammed A. Aleid[4], Osamah Ibrahim Khalaf[5] and Theyazn H. H. Aldhyani[6]*

[1] Faculty of Pharmacy, Gomal University, D. I. Khan, Pakistan, [2] Department of Computer Science, Taif University, Taif, Saudi Arabia, [3] Department of Oral & Maxillofacial Surgery, Taif University, Taif, Saudi Arabia, [4] College of Education, King Faisal University, Al-Ahsa, Saudi Arabia, [5] Al-Nahrain Nanorenewable Energy Research Center, Al-Nahrain University, Baghdad, Iraq, [6] Applied College in Abqaiq, King Faisal University, Al-Ahsa, Saudi Arabia

OPEN ACCESS

Edited by:
Jan Ilhan Kizilhan,
University of Duhok, Iraq

Reviewed by:
Francesco Monaco,
Azienda Sanitaria Locale Salerno, Italy
Gaia Sampogna,
University of Campania "L. Vanvitelli", Italy

***Correspondence:**
Junaid Asghar
junaid.asghar@gu.edu.pk

Specialty section:
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

Received: 11 February 2022
Accepted: 02 May 2022
Published: 27 May 2022

Citation:
Asghar J, Tabasam M, Althobaiti MM, Adnan Ashour A, Aleid MA, Ibrahim Khalaf O and Aldhyani THH (2022) A Randomized Clinical Trial Comparing Two Treatment Strategies, Evaluating the Meaningfulness of HAM-D Rating Scale in Patients With Major Depressive Disorder. Front. Psychiatry 13:873693. doi: 10.3389/fpsyt.2022.873693

Introduction: Due to the complexity of symptoms in major depressive disorder (MDD), the majority of depression scales fall short of accurately assessing a patient's progress. When selecting the most appropriate antidepressant treatment in MDD, a multidimensional scale such as the Hamilton Depression Rating scale (HAM-D) may provide clinicians with more information especially when coupled with unidimensional analysis of some key factors such as depressed mood, altered sleep, psychic and somatic anxiety and suicidal ideation etc.

Methods: HAM-D measurements were carried out in patients with MDD when treated with two different therapeutic interventions. The prespecified primary efficacy variables for the study were changes in score from baseline to the end of the 12 weeks on HAM-D scale (i.e., $\leq$8 or $\geq$50% response). The study involved three assessment points (baseline, 6 weeks and 12 weeks).

Results: Evaluation of both the absolute HAM-D scores and four factors derived from the HAM-D (depressed mood, sleep, psychic and somatic anxiety and suicidal ideation) revealed that the latter showed a greater promise in gauging the anti-depressant responses.

Conclusion: The study confirms the assumption that while both drugs may improve several items on the HAM-D scale, the overall protocol may fall short of addressing the symptoms diversity in MDD and thus the analysis of factor (s) in question might be more relevant and meaningful.

Keywords: major depressive disorder, escitalopram, nortriptyline, antidepressant, HAM-D, remission, response

INTRODUCTION

Major depressive disorder (MDD) or unipolar depressive disorder is a syndrome most frequently diagnosed in psychiatric practice. It is characterized by low mood, loss of interest or pleasure and decreased energy, reduced self-esteem and self-confidence in usual activities and is associated with a paralyzed social status (1, 2). Around 280 million people worldwide suffer from depression. MDD is distinct from normal changes in mood and/or short-term emotional responses to everyday challenges. Each year, an estimated 5% of adults globally suffer from depression, yet it continues to be a neglected global health concern, with the majority of cases occurring in young people (3).

There is widespread recognition that this disorder is not a homogenous entity, and that further clinical characterization of the patient is required to customize the treatment plan. A range of pharmacotherapies have been demonstrated to be "equivalent" in the treatment of the syndrome in clinical trials, and these interventions are thus generally considered as interchangeable (4). Pharmacotherapy for depression is generally multifactorial and typically based on the clinician's and/or patient's preference and on tolerability issues and this could be one of the reasons why the majority of people diagnosed with depression do not achieve remission following their initial treatment (5), however, achieving complete remission of depressive symptoms and the return to normal daily functioning are the ultimate goals of antidepressant therapy. It has been demonstrated in studies that achieving remission and maintaining antidepressant therapy for a long period of time after the acute symptoms have subsided can help to prevent relapse or recurrence of the psychiatric episode and restoration of social and occupational functioning (6).

Antidepressants were first introduced to the field of psychopharmacology in the 1960s, and since then the 17-item Hamilton Depression Rating Scale (HAM-D17 or HDR) has become one of the most widely used scales to quantify the severity of symptoms of depression and determine the treatment responses. Response size is a widely used variable in antidepressant clinical trials (7) and it is usually defined as a score reduction of 50% or more on standardized depression scales. HAM-D is still considered the "gold standard" in determining the efficacy of antidepressant treatments (8) however, according to some studies, its score does not appear to be a sufficient measure of the antidepressant outcome during a clinical trial. Because the HAM-D is not a unidimensional scale (9). When developing an antidepressant treatment strategy, a more targeted approach should be used to describe the antidepressant profiles of different therapeutic agents, such as focusing on the individual item scoring, for example, changes in sleep, suicidal behavior, psychosomatic factors, appetite, or weight loss.

The studies show that a depressed patient who responds with a 50% reduction in the HAM-D score may still experience significant symptoms especially if the patient was severely depressed prior to the initiation of therapy. Remission is defined according to post-treatment scores of a depression rating scale and is commonly defined as a low absolute score of ≤ 7 on the HDR (10, 11). However, response does not always imply remission (12).

Many antidepressants, such as SSRIs, have been widely used to treat depressive symptoms, but they have been shown to disrupt sleep and cause other negative effects such as suicidal thoughts and changes in appetite, whereas others with sedative properties (e.g., TCAs) improve sleep, but may cause problems over time due to oversedation. As a result, patients on various antidepressants complain about treatment failure. Due to the activation of 5-HT2 receptors and an increase in noradrenergic and dopaminergic neurotransmission, some antidepressants have been shown to impair sleep quality. Among them, most prominent are selective serotonin reuptake inhibitors (SSRI), serotonin and norepinephrine reuptake inhibitors (SNRI), norepinephrine reuptake inhibitors (NRI), monoamine oxidase inhibitors (MAOI), and tricyclic antidepressants (TCA) (13).

METHODS

Participants

Newly diagnosed community-dwelling outpatients ($n = 500$) with MDD on initial treatment attempt, aged 20–50 years of either gender, living in D.I.Khan city, KPK province of Pakistan were enrolled in the present study. The benefits and potential risks of study participation were fully explained to each patient. All participants met the defined eligibility criteria and gave informed consent for the data collection. Baseline psychiatric and somatic symptoms related to MDD, and the medications' response were evaluated at each visit.

Inclusion Criteria

Patients were included in this trial when they (i) were awaiting to be treated with routine mental health care; (ii) were 20-50 years; (iii) met criteria of a major depressive episode (according to DSM-V); (iv) and who returned the signed informed consent form.

Exclusion Criteria

Patients were excluded in case of (i) presence of any acute or unstable medical condition; (ii) concomitant use of any psychotropic drug; (iii) patients with a history of substance abuse (iv) pregnant and lactating mothers; (iv) patients with multiple disorders e.g., bipolar disorder, obsessive-compulsive disorder (OCD), post-traumatic stress disorder (PTSD) and eating disorders; thyroid dysfunction (v) and terminally ill patients.

Drugs Used

Escitalopram

It is the active enantiomer of citalopram and belongs to the SSRIs (selective serotonin reuptake inhibitors) class. Other members in this therapeutic category include fluoxetine, paroxetine and sertraline and are currently the most widely used antidepressants. Escitalopram has been approved as a first line treatment option for major depressive disorder and generalized anxiety disorder (GAD). It increases the extracellular level of serotonin by inhibiting its reabsorption into the presynaptic cell, thereby increasing the level of serotonin available to bind to the postsynaptic receptor in the synaptic cleft (14, 15).

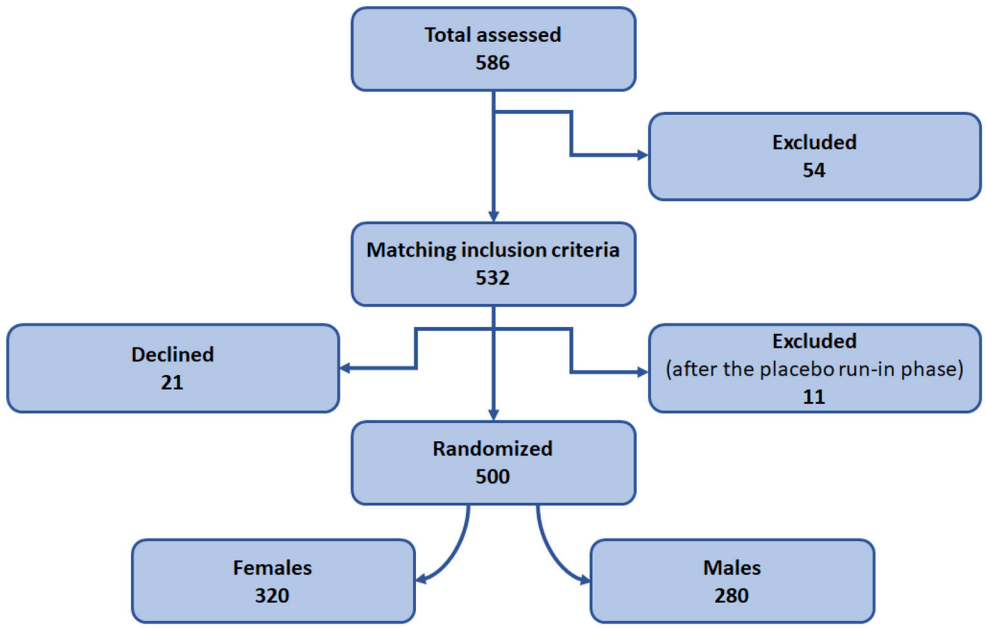


FIGURE 1 | Study design and patient recruitment. Diagram representing design of the randomized control trial.

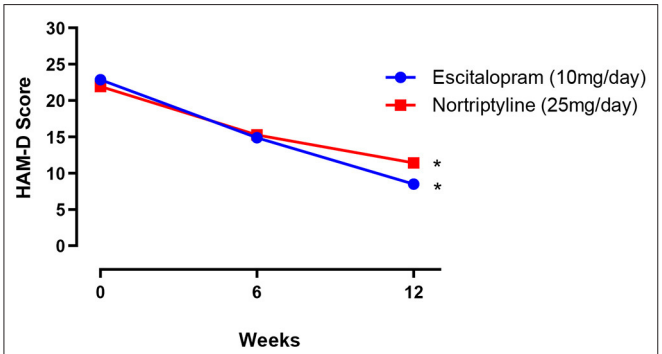


FIGURE 2 | The mean ± SEM on HAM-D depression scale ($n = 250$ in each group), showing baseline and post-treatment scores at 6 and 12 weeks in patients randomly allocated to two drugs: either escitalopram or nortriptyline. *$p < 0.05$ shows the significant improvement in baseline depressive symptoms (one-way ANOVA).

Nortriptyline

It belongs to the tricyclic antidepressants (TCAs) category. These drugs have historically played a significant role in the pharmacotherapy of MDD and are still used as a first line option. Nowadays, other antidepressant agents such as SSRIs and serotonin-noradrenaline reuptake inhibitors (SNRIs) are frequently considered as first line in the treatment of MDD. TCAs are still prescribed in cases of poor tolerability and/or a high rate of non-response to SSRIs and SNRIs (16, 17). The majority of TCAs work as SNRIs by inhibiting the serotonin transporter (SERT) and the norepinephrine transporter (NET), resulting in an increase in synaptic concentrations of these neurotransmitters and hence improved neurotransmission (18). The World Health Organization (WHO) and the World Federation of Societies of Biological Psychiatry (WFSBP) guidelines indicate that TCAs, SSRIs, SNRIs, and the newer antidepressants; mirtazapine and bupropion have no general preference (19, 20).

Study Design and Randomization

It is an open-label, randomized, fixed-dose, parallel-design study conducted at the psychiatric OPD, DHQ/TH, MTI, D.I.Khan. The patients were recruited and randomized into the study as detailed in the **Figures 1**, **3**. A placebo run-in phase was performed in the post-inclusion/ pre-randomization period in which all the patients were given a placebo for a period of 2 weeks and the patients were assessed, and anyone who had improved substantially was excluded from the study. The investigator who conducted the randomization was not engaged in the medication dispensing, patient inclusion, or follow-up. The patients were randomized to receive either escitalopram 10 mg/day or nortriptyline 25 mg/day. The drug dosages were determined by the investigators' clinical judgment, taking into account the newly diagnosed participants' response and tolerability. The baseline data were collected, and the patients were examined at 6 weeks interval for drug responses. Overall, the data were collected at three time points (0, 6, and 12 weeks).

Psychiatric nurses monitored and ensured the drug adherence. DSM-V criteria (HAMD-17) was used to evaluate the total scores and subscore variables pre and post treatment. The answers were scored from 0 to 2 or 0 to 4 and summed up to obtain an overall score, according to the HAM-D protocol. Out of 17 items, nine items were sub scored from 0 to 2 while eight items were sub scored from 0 to 4, in which 0 represents symptoms "not present" while 4 means "severe" symptoms.

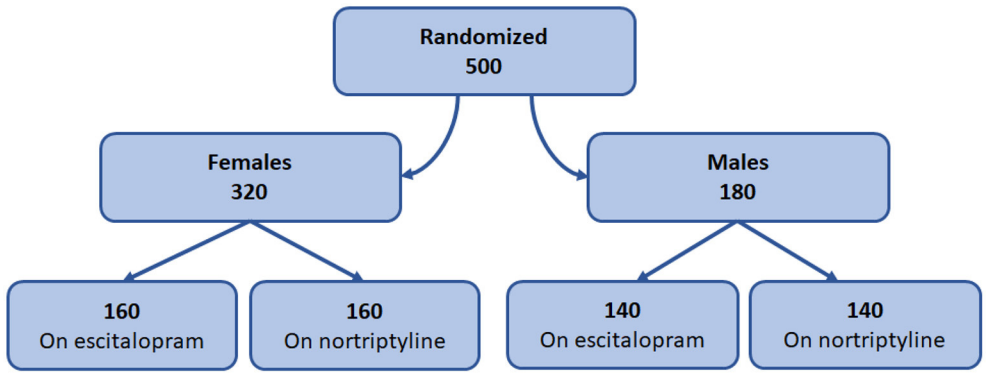


FIGURE 3 | Gender-based treatment protocol. Diagram representing the design of gender-based randomization and treatment plan.

A score of 8 or less was considered equivalent to a remission. Clinical efficacy was defined as 50% or greater reduction in HAM-D rating scores, indicating a positive treatment response. Partial response was defined as an improvement between 25 and 49%. The primary efficacy parameter was measured as the mean change of scores from baseline to end of treatment between escitalopram and nortriptyline treated groups.

Aims

To assess the usefulness and robustness of the HAM-D scale (absolute and individual factors scores) in a two-prong approach, comparing the efficacies of escitalopram (selective serotonin reuptake inhibitor) and nortriptyline (tricyclic antidepressant), particularly targeting the sensitivity of psychiatric and somatic subscales in diagnosing patients with MDD. The prespecified primary efficacy variables for the study were changes in score from baseline to the end of the 12 weeks on HAM-D scale (i.e., $\leq$8 or $\geq$50% reduction in HAMD-17 score from baseline to endpoint).

Data Analyses

The effect size was calculated for the difference in mean change percent for each category. The data is presented as mean $\pm$ standard error mean and the p-value threshold of $\leq$0.05 is considered as significant. Changes in the HAMD-17 absolute scores and subscores from the baseline to endpoints were analyzed using one-way/ or two-way ANOVA. The *post-hoc* analysis included Dunnett's and/or Tukey's tests.

RESULTS

Baseline Characteristics

Clinical characteristics at baseline were assessed for all the patients ($n = 500$) using Clinical Global Impressions (CGI) Scale to ascertain patients' symptoms severity profile, prior to the initiation of pharmacotherapy. Patients were evaluated by a panel of psychiatrists and the CGI-S responses were analyzed on a 7-point scale ranging from 1 = not ill, to 7 = extremely ill as shown in the **Table 1**.

TABLE 1 | Patients' clinical characteristics.

Clinical characteristics	***n* = 500**
CGI-S score	***n* (%)**
1- Normal	0 (0.0 %)
2- borderline ill	20 (4%)
3- Mildly ill	23 (4.6 %)
4- Moderately ill	230 (46 %)
5- Markedly ill	196 (39.2 %)
6- Severely ill	31 (6.2 %)
7- Among the most extremely ill patients	0 (0.0 %)

Following that, five treatment outcomes were evaluated over a 12-week period (i) overall comparative efficacy of the two antidepressants; (ii) gender-based treatment response; (iii) age-based treatment response; (iv) and efficacy in treating psychosomatic disorder.

Overall Comparative Efficacy

Both male and female patients were randomly divided into two treatment groups of equal size (250 patients in each group) either on escitalopram (10 mg/day) or nortriptyline (25 mg/day) monotherapy, administered over a period of 12 weeks. In the first group, patients with depressive symptoms (baseline 22.9 $\pm$ 0.7) were given escitalopram (10 mg/day) over a period of 12 weeks, which resulted in a significant reduction of symptoms (8.50 $\pm$ 0.5) and a clinical response was demonstrated (62.9%) at the end of the treatment plan. Whereas, patients on nortriptyline (25 mg/day), showed a partial improvement (47.9%). Clinical response/ efficacy was achieved only in terms of $\geq$50% reduction in baseline HAM-D) in the escitalopram group (**Figure 2**; **Table 2**).

Gender Based Treatment Response

Of the 500 patients enrolled in the study, 180 (36%) were males and 320 (64%) females. Although the number of male and female patients recruited were different, we avoided the block randomization (21) and the imbalance in the number was kept

TABLE 2 | The data show the mean ± SEM on 17- item HAM-D depression scale ($n = 250$ in each treatment group), compared to baseline at 6 and 12-weeks post-treatment scores.

Weeks	Escitalopram Mean ± SEM	Improvement (%)	Nortriptyline Mean ± SEM	Improvement (%)
0	22.9 ± 0.6	-	21.9 ± 0.6	-
6	14.9 ± 0.8	34.9	15.2* ± 0.6	30.6
12	8.5 ± 0.5	62.9*	11.4 ± 0.8	47.9

*The patients were kept on two drugs: escitalopram and nortriptyline (monotherapy) for up to 12 weeks and the percent improvement of symptoms was recorded for each drug group. *Clinical response/remission was defined as ≤8 or ≥50% reduction in baseline HAM-D.*

the same to prevent the selection bias (22). All the patients were randomly allocated to one of the two treatment groups as shown in the **Figure 3**.

The change in total mean score (from baseline to endpoint) was evaluated for both the groups. At the end of the therapy, improvement in depressive symptoms was associated with a decrease of −6.9 and −8.7 points on escitalopram in male and female patients, respectively, whereas, nortriptyline treatment resulted in an average reduction of −10.1 and −12.9 within male and female patients, respectively (**Figures 4A,B**; **Table 3**). In the male group, a significant clinical response was achieved on escitalopram and nortriptyline-treated patients (69.3 and 51.9%, respectively) at 12 weeks. However, in the female group, only escitalopram was significantly more effective (63.1%) than nortriptyline which demonstrated only partial response (42.9%) as shown in the **Table 3**.

Age-Based Treatment Response

To test whether escitalopram or nortriptyline might differ in efficacy to minimize anxiety/ somatization sub-scores in different age groups, an aged-based comparison was performed. The patients of either sex were divided into 6 age groups: (20–25, 26–30, 31–35, 36–40, 41–45, and 46–50 years) and were randomly allocated to either escitalopram. Both the drugs resulted in significant reduction of symptoms on HAM-D rating scale and produced a statistically significant response in all the age group at the end of the treatment plan ($^*p < 0.05$; One-way ANOVA) (**Figures 5A–F**). However, a varied clinically response was achieved across different age groups as summarized in **Table 4**.

Psychosomatic Disorder and Treatment Response

Some of the HAMD-17 data [depressed mood, psychic anxiety, somatic anxiety symptoms (indigestion, palpitations and headache) and insomnia (initial and middle)] from 500 patients of the 12-week trial comparing the effectiveness of escitalopram and nortriptyline were converted to subscale scores and analyzed during the antidepressant treatment course.

A standard effect size analysis showed improvement in psychosomatic symptoms, following up to 12 weeks of therapy with either escitalopram or nortriptyline monotherapy. Analysis of subscale scores for anxious depression such as depressed mood

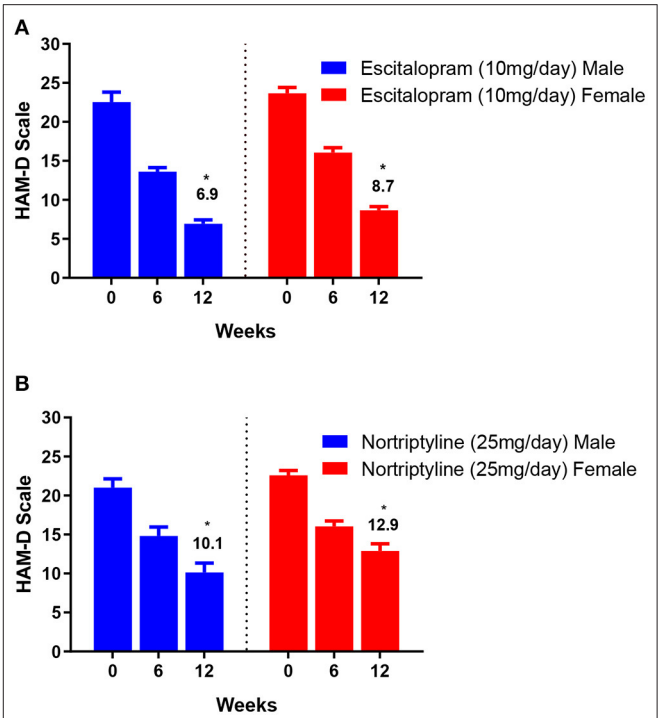


FIGURE 4 | (A,B) Show the gender-based data on 17- item HAM-D depression scale in patients assigned to two different treatment modalities, i.e., escitalopram and nortriptyline for up to 12 weeks. *$p < 0.05$ shows the significant improvement in baseline depressive symptoms (one-way ANOVA).

(sadness, worthlessness, tendency to weep) and psychic anxiety (chronic excessive worry) were assessed. Additionally, analysis of the sub-scores such as insomnia (initial and middle) and somatic anxiety (indigestion, palpitations and headache) were carried out to assess if there was any improvement in the baseline severity. The changes in various psychosomatic parameters/subscale scores are shown in **Figures 6–8** and summarized in **Table 5**.

DISCUSSION

The aim of this research was to evaluate the HAM-D scale's suitability and practicability when comparing two different treatment outcomes in a group of patients who were treated according to the general protocols in a hospital setting. MDD usually goes under-treated as the patients do not respond equally to the available antidepressant choices due to multiple factors such as complexities in psychosocial variables, lack of proper assessment, poor medication response and lack of adherence to the treatment protocols. Consequently, the overall aim of the project was to evaluate the usefulness of HAM-D scale and, followed by a micro-analytic approach derived from HAM-D, in which four specific items were analyzed separately.

We selected and analyzed the data on the basis of a set of primary efficacy variables on HAM-D from baseline to the end of 12 weeks (i.e., ≤8 or ≥50%). According to outcomes of a meta-analysis on MDD and different antidepressants, about one-fourth of the studies showed remission within 12 weeks, one-third of

TABLE 3 | The gender-based data on 17- item HAM-D depression scale.

Drugs	Weeks	Male patients (Mean ± SEM)	Improvement (%)	Female patients (Mean ± SEM)	Improvement (%)
Escitalopram	0	22.5 ± 1.3	-	23.7 ± 0.8	-
	6	13.6 ± 0.5	39.6	16.1 ± 0.6	32.1
	12	6.9* ± 0.5	69.3*	8.7 ± 0.7	63.1*
Nortriptyline	0	21.0 ± 1.1	-	22.6 ± 0.6	-
	6	14.8 ± 1.2	29.5	16.0 ± 0.7	29.2
	12	10.1 ± 1.2	51.9*	12.9 ± 0.9	42.9

*Both male and female patients were randomly assigned to escitalopram and nortriptyline for up to 12 weeks. Mean ± SEM and percent improvement of symptoms were recorded for each drug group. *Clinical response/remission was defined as ≤8 or ≥50% reduction in baseline HAM-D.*

the studies showed remission within 6 months, while one and a half studies showed remission within the period of 12 months (23). A cohort study conducted in primary health care showed the highest remission rate of depressive features in the third and 6 months of the study (24). Antidepressants reach a plateau or stable effect after 6–12 weeks of treatment (25); therefore, a 12-week study was conducted in order to examine the full range of therapeutic efficacies to antidepressants.

The HAMD-17 item scale has been a widely used tool in psychiatric research to assess the severity of depression. Its original version contained 17 items, but it kept updating and its latest revision took place in 1980. The Beck Depression Inventory (BDI) is another widely used depression scoring tool in research that has evolved over time; its most recent version, known as the BDI-II, was introduced in 1996. One of the limitations of these scales is that the side effects of antidepressants could intensify the item scores on these scales and thus mask the true positive effects of the antidepressant agents (26, 27).

Both escitalopram and nortriptyline are the frequently used antidepressant agents in treating MDD. In this study, we used a fixed dose of escitalopram, 10 mg/day and nortriptyline, 25mg/day in our newly diagnosed MDD patients. Numerous placebo-controlled trials have shown that when patients with MDD received escitalopram at a dose of 10 mg/day, it had a significantly greater efficacy than placebo. Furthermore, the escitalopram group had a higher rate of remission than the placebo group. Consequently, 10 mg/day was found to be safe and effective in the initial stages of the MDD. In terms of reduction in depression scores, the efficacy was greater with escitalopram than with placebo at the first or second week and were maintained throughout treatment at these doses (28, 29). Studies reveal that TCAs initial and maintenance dosages are determined empirically and are not substantiated by strong clinical evidence. Lower doses of nortriptyline (25–100 mg/day) were found to be equally efficacious as higher doses with lesser adverse effect events in one review (30).

In our findings, overall comparative efficacy/ target remission (≤8 or ≥50% reduction in baseline HAM-D) of the two drugs revealed that escitalopram is significantly more effective (62.9%) in comparison to nortriptyline (47.9%). None of the drugs could achieve the other efficacy target i.e., ≤8 score (**Table 2**). As a general trend, subjects of all age groups, receiving escitalopram showed highest remission rates than nortriptyline at the end of the therapy. Furthermore, no significant difference was recorded in terms of antidepressant efficacy (absolute HAMD-17 score of ≤8) after 6 weeks of therapy with either drug in all the age groups. Escitalopram offered superior control of depressive symptoms and led to clinical remission at the end of the study (12 weeks) in all age groups with a reduction of 50% or more of the HAMD-17 score, however, in terms of cut-off value on HAM-D scale (≤8), only the age group 26–30 achieved the target score. On the other hand, some interesting data were obtained with nortriptyline which produced a clinical response (≥50% reduction in baseline HAM-D) in the older age groups (41–45 and 46–50), however, it could not produce the same effect in the earlier age groups (20–40 years) (**Table 5**). In order to investigate this differential age-related drug response, a thorough search of the literature led to the extraction of a study where the author recommended TCAs to be more effective antidepressant agents for the acute and/or longer course of antidepressant therapy, particularly in elderly patients (31), however, TCAs are no longer preferred as first-line agents for geriatrics (above 60 years) due to their potential side effects, including postural hypotension, which can contribute to falls and fractures, cardiac conduction disorders and anticholinergic/ antihistaminic effects (32). There is a widely held assumption that individual differences underlie the variability in the association of antidepressants with depressive symptoms (i.e., response). To our knowledge, however, efforts to discover characteristics related with antidepressant response and their impacts on different age groups or gender have been largely ineffective. Nonetheless, depression appears to be a more heterogeneous condition than other psychotic states like schizophrenia, and the unexplained source of this heterogeneity may account for part of the observed variability in antidepressant treatment response in different age groups (33–36).

Depression is more prevalent in women as compared to men (37, 38). Females aged 14 to 25 years have been reported to have twice the prevalence rate of depression as compared to men (39–41); largely due to the hormonal fluctuations, whereas the prevalence rate before puberty remained the same in both genders (23, 39). To see if there were any differences in the rates of improvement based on gender, we tested escitalopram and nortriptyline and observed that the symptoms of males significantly improved by the end of treatment (12 weeks),

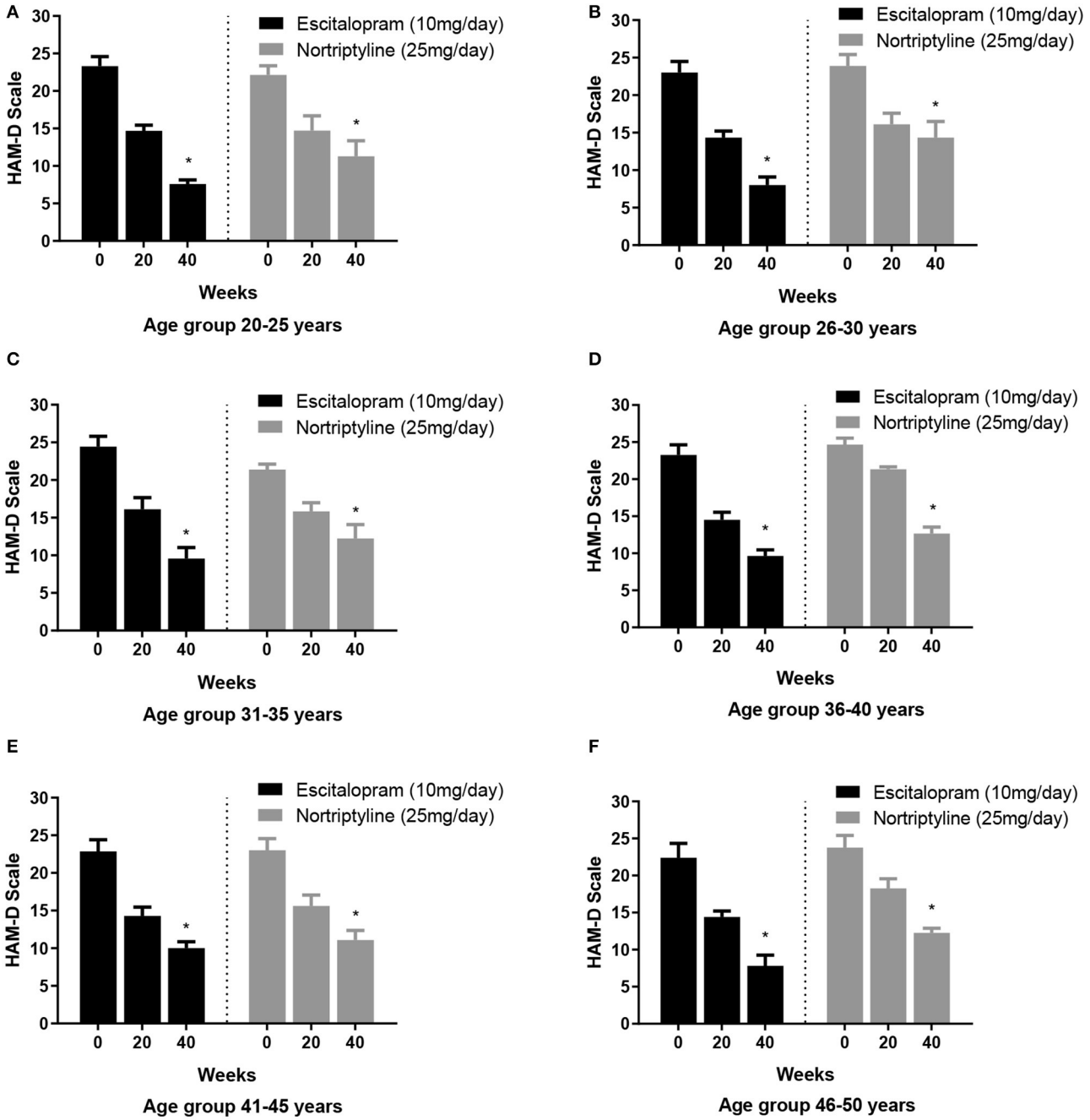


FIGURE 5 | (A–F) Show the age group-based data on 17- item HAM-D depression scale. Participants in different age groups (15–20, 21–25, 26–30, 31–35, 36–40, 41–45, and 46–50 years) were randomly allocated to escitalopram and nortriptyline for up to 6 weeks. *$p < 0.05$ shows the significant improvement in baseline depressive symptoms (one-way ANOVA).

leading to ≥50% reduction of symptoms, while in females, escitalopram was found to be more efficacious than nortriptyline, as the latter showed only 42.9% reduction at the end of the therapy (**Table 3**). Despite decades of research, there is still no clear consensus on whether there are sex-related efficacy differences in antidepressant treatment. For example, males had a considerably better therapeutic response to another tricyclic antidepressant, imipramine, than females. These differences in response could be caused by a multitude of variables, including body fat distribution, liver metabolic rates, hormone physiology, and brain interactions between estrogen and serotonin (40).

To achieve more relevant and robust outcomes, we additionally performed a factor-based evaluation of some key symptoms. Factors/ subscores analyses on HAM-D for

TABLE 4 | The age group-based data on 17- item HAM-D depression scale.

Age group	Weeks	Escitalopram group Mean ± SEM	Improvement (%)	Nortriptyline group Mean ± SEM	Improvement (%)
20–25	0	23.3 ± 1.3	-	22.5 ± 1.2	-
	6	14.7 ± 0.7	36.9	14.7 ± 1.9	33.6
	12	7.6* ± 0.5	67.4*	11.3 ± 2.07	48.9
26–30	0	23.0 ± 1.5	-	23.9 ± 1.5	-
	6	14.3 ± 0.9	37.8	16.1 ± 1.5	32.6
	12	8.0* ± 1.09	65.2*	14.3 ± 2.2	40.2
31–35	0	24.4 ± 1.4	-	21.4 ± 0.7	-
	6	16.1 ± 1.5	34.0	15.8 ± 1.1	26.2
	12	9.6 ± 1.5	60.7*	12.2 ± 1.9	42.9
36–40	0	23.3 ± 1.4	-	24.6 ± 0.8	-
	6	14.5 ± 1.0	37.8	21.3 ± 0.3	13.4
	12	9.6 ± 0.8	58.8*	12.6 ± 0.8	48.8
41–45	0	22.9 ± 1.6	-	23.0 ± 1.6	-
	6	14.3 ± 1.2	37.6	15.6 ± 1.5	32.2
	12	10.0 ± 0.9	56.3*	11.1 ± 1.3	51.7*
46–50	0	22.4 ± 1.9	-	23.8 ± 1.6	-
	6	14.4 ± 0.8	35.7	18.3 ± 1.3	23.1
	12	9.8 ± 1.5	56.3*	11.5 ± 0.6	51.7*

*Participants in different age groups (20–25, 26–30, 31–35, 36–40, 41–45, and 46–50 years) were randomly allocated to escitalopram and nortriptyline for up to 12 weeks. Mean ± SEM and percent improvement of symptoms were recorded for each drug group. *Clinical response/ remission was defined as ≤8 or ≥50% reduction in baseline HAM-D.*

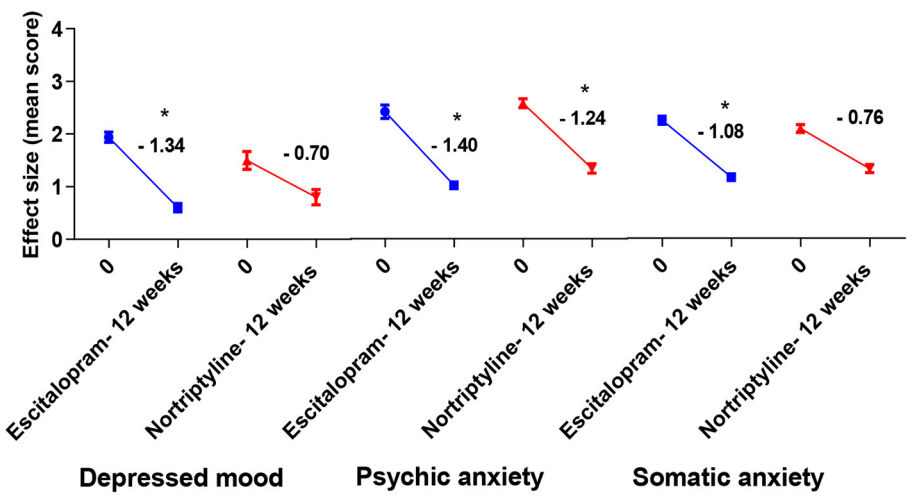


FIGURE 6 | Plots show the mean changes/ reduction in effect size from the baseline for psychiatric symptoms (depressed mood and psychic anxiety) and somatic anxiety symptoms (indigestion, palpitations and headache) on the HAM-D scale. *$p < 0.05$ shows significant improvement (one-way ANOVA, followed by Dunnett's multiple compassion test).

MDD assessment may be more sensitive to antidepressant drug effects (41), so we looked at the sensitivity to depressive change for some key subscales [depressed mood, psychic anxiety, somatic anxiety symptoms (indigestion, palpitations and headache) and insomnia (initial and middle)] which performed better throughout the treatment course, with some subscales having advantage in detecting the treatment effects. Following up to 12 weeks of therapy with either escitalopram or nortriptyline monotherapy, a standard effect size analysis showed improvement in psychosomatic symptoms. Analyses of effect size scores (baseline to week 12) for the different treatment groups showed some interesting results. A *post-hoc* analysis of the effect sizes for each item (**Figure 6**; **Table 5**) showed considerable change in the escitalopram and nortriptyline group (e.g., psychiatric anxiety and somatic anxiety symptoms). The item "somatic anxiety" had the highest impact in the nortriptyline

group. On the other hand, escitalopram significantly improved insomnia-middle, however, its effects on insomnia-initial were very small (**Figure 7**; **Table 5**) which means both the drugs resulted in increased sleep latency, however, the total sleep time was significantly improved in the escitalopram group.

According to Husain et al. (42), both escitalopram and nortriptyline were significantly effective in relieving painful physical symptoms and depression severity. Several studies reveal that the MDD associated somatic symptoms are difficult to treat with the available antidepressant choices (43–47). According to Marangell et al. (48), subjects receiving both nortriptyline and escitalopram for 2 weeks, showed 50% reduction in somatic anxiety, however, clinical response in terms of physical and depressive features was achieved subsequently. For example, on HAM-D 17 item scale, baseline data showed no significant difference in the severity of depression in subjects with or without painful somatic symptoms and regardless of somatic complains, remission rate for MDD remained 84%. Subjects with somatoform disorder reported having severe depressive episodes, which greatly affected the therapeutic outcomes and decreased the clinical response rate in totality. In the current study, we found a significant improvement in depressive and somatoform symptoms with time (12 weeks of therapy). Same has been investigated in some other studies that remission in somatic symptoms is associated with an overall remission of MDD symptoms, with the longer course of antidepressant therapy (43, 46, 49, 50).

Antidepressants' therapeutic efficacy may be hampered by side effects like insomnia, because continual insomnia can exacerbate depressive episodes and mask the true antidepressant effects of these drugs (51, 52). Previous studies show that TCAs produce significant improvement in normalizing sleep pattern when compared to SSRI, because of their anticholinergic and antihistaminic properties. At the same time, the sleep efficiency and depth are substantially reduced in depressed patients and changes in rapid eye movement (REM) are most commonly affected (13). SSRIs might be responsible for a disturbed sleep cycle (particularly difficulty falling asleep) (53) and this has been linked to the activation of 5-HT2 receptor which leads to mental activation and thus insomnia, and therefore, add up to the pre-existing burden of depressive symptoms. TCAs, however, due to their central anticholinergic and H_1 blocking actions could

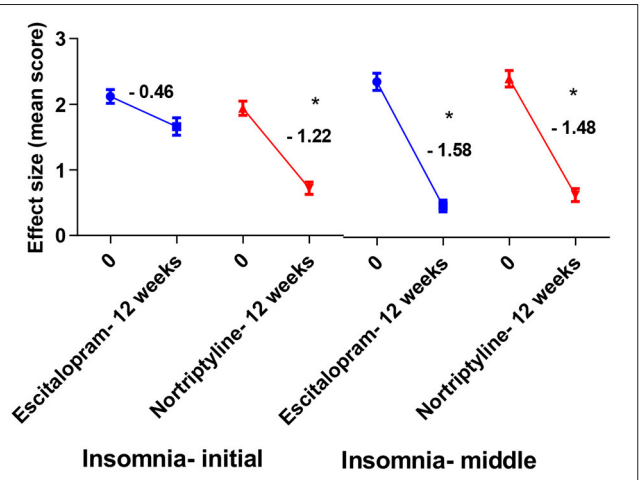


FIGURE 7 | Plots show the mean changes/ reduction in effect size from the baseline for insomnia (initial and middle) on the HAM-D scale. *$p < 0.05$ shows significant improvement (one-way ANOVA, followed by Tukey's multiple compassion test).

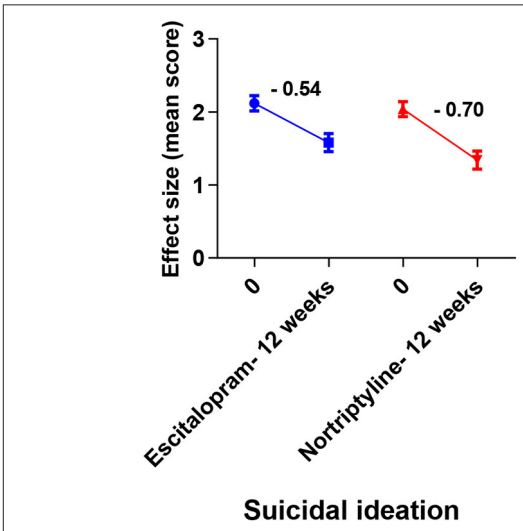


FIGURE 8 | Plots show the mean changes/ reduction in effect size from the baseline for suicidal ideation on the HAM-D scale. No significant change was observed across different data sets (one-way ANOVA, followed by Tukey's multiple compassion test).

TABLE 5 | The comparison of mean changes in effect size compared to baseline for depressed mood, psychic anxiety, somatic anxiety and insomnia and suicidal ideation (subscale scores) on the HAM-D scale at 12 weeks among patients with MDD, treated with escitalopram and nortriptyline.

Drug	Change in effect size (difference of mean scores)					
	Depressed mood	Psychic anxiety	Somatic anxiety	Insomnia		Suicidal ideation
				Initial	Middle	
Escitalopram	−1.34* ± 0.1	−1.40* ± 0.1	−0.94 ± 0.1	−0.46 ± 0.1	−1.58* ± 0.1	0.54 ± 0.1
Nortriptyline	−0.70 ± 0.2	−1.24* ± 0.1	−1.2* ± 0.1	−1.22* ± 0.1	−1.48* ± 0.1	0.70 ± 0.2

*$p < 0.05$ shows significant improvement (one-way ANOVA, followed by Tukey;s multiple compassion test).

improve sleep (54). Accordingly, our findings show that TCAs are significantly better at relieving insomnia than SSRIs, while patients in the latter group reported marked insomnia (**Figure 7**; **Table 5**).

MDD is commonly associated with suicidal thoughts/ideation. More than 60 percent of people who have attempted suicide worldwide have MDD. There is a 20-fold higher risk of suicide among patients with MDD, compared to the general population (55, 56). To treat or prevent suicidal ideation and suicide attempts, antidepressants must be prescribed. According to pharmacoepidemiological studies, the number of suicides decreased as the use of antidepressants increased (57, 58). There have been reports of increased and new-onset suicidal activities since 1988s with TCAs. Also, the SSRIs have been the subject of debate for the past two decades, with a focus on their role in the treatment of depression and anxiety. Controversial results have been found in meta-analysis of randomized trials (59). Since suicidal events are so rare, Gunnell et al. (60) stated in their meta-analysis that SSRIs' effects could not be evaluated. Suicidal thoughts and behavior triggered by antidepressant drugs (primarily with SSRIs) are extremely rare (61). Restlessness and impulsiveness are all possible warning signs in the early stages of psychosis. Based on our study (HAM-D item-analysis protocol), no drug significantly reduced the suicidal thoughts, however, nortriptyline resulted in a larger score reduction as compared with escitalopram (**Figure 8**; **Table 5**). To address the issue, it is recommended that when treating depression for the first time, an appropriate combination therapy may be preferred over monotherapy. However, according to the current study's protocols, switching from nortriptyline to escitalopram resulted in better outcomes than switching from escitalopram to nortriptyline at the end of the study period (data not shown).

In this study, there were no unexpected side effects from the usage of escitalopram or nortriptyline. Escitalopram induced a modest weight increase, as expected, as well as nighttime insomnia. SSRIs have historically been associated with insomnia and poor subjective sleep quality (62). With our participants, we found the same as a general trend. As a result, patients were advised to take escitalopram during the daytime to circumvent nighttime insomnia. Nortriptyline has been a useful antidepressant, though the prevalence and severity of anticholinergic side effects is a downside. We discovered a correlation between efficacy and anticholinergic side effects such as dry mouth and/ or constipation in all the age groups in the current investigation (data not shown). However, no participant dropped out of the trial due to intolerance to these side effects. The delayed onset of antidepressant action has traditionally been an impediment to depression treatment. Antidepressants' complete therapeutic efficacy may take several weeks to manifest, leaving patients to endure prolonged episodes of depressive symptomatology (63) as was the case with this study. One of the most crucial aspects of the relationship of socioeconomic status to psychiatric health, and one of the most consistent associations in the field of psychiatric epidemiology, is the relationship of socioeconomic status to psychiatric disorders (64). With respect to sociocultural context, some of our participants were reluctant to accept that they had depression, and even whether treatment is needed at all. For some, depression was stigmatizing. Furthermore, convincing them to initiate the treatment was challenging in some cases.

Limitations

Several limitations of our study are worth mentioning, including the participants, most of which were females, and all were Asian, thus limiting the study's generalizability to other populations. Similarly, during the administration of the questionnaire, special attention was paid to the evaluation of each element's meaning, without eliciting any significant questions or observations from the participants. The study was only limited to the effects of two drugs; several antidepressants were still very expensive at the time of the study and the participants preferred cost-effective and easily accessible options offered: escitalopram and nortriptyline. Using other anti-depressants such as paroxetine, bupropion, duloxetine and desvenlafaxine, may yield different outcomes.

CONCLUSION

Using data from this clinical trial, we could conclude that the individual effect size analysis has some advantages over the HAM-D absolute scores for depression assessment because of its more focused factor-based approach of evaluating depressive symptoms pre and post treatment. The practicing psychiatrists might follow or want to consider tailoring our methods to their particular needs when comparing different antidepressants' efficacies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

This study was approved by the Ethical Review Board (ERB) of the Gomal University, KP, Pakistan. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

Present research work is supported by Taif University Researchers Supporting Project number (TURSP-2020/130), Taif University, P.O Box 11099, Taif 21944, Saudi Arabia.

REFERENCES

1. Kessler RC, Berglund P, Demler O. The epidemiology of major depressive disorder result from the national comorbidity survey replication (NCS-R). *JAMA*. (2003) 289:3095–105. doi: 10.1001/jama.289.23.3095
2. Angst J, Gamma A, Gastpar M, Lépine JP, Mendlewicz J, Tylee A. Gender differences in depression. epidemiological findings from the European DEPRES I and II studies. *Eur Arch Psychiatry Clin Neurosci.* (2002) 252:201–9. doi: 10.1007/s00406-002-0381-6
3. Herrman H, Patel V, Kieling C, Berk M, Buchweitz C, Cuijpers P, et al. Time for united action on depression: a lancet-world psychiatric association commission. *Lancet*. (2022) 399:957–1022. doi: 10.1016/S0140-6736(21)02141-3
4. Maj M, Stein DJ, Parker G, Zimmerman M, Fava GA, De Hert N, et al. The clinical characterization of the adult patient with depression aimed at personalization of management. *World Psychiatry*. (2020) 19:269–93. doi: 10.1002/wps.20771
5. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW Warden D, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*. (2006)163:1905–17. doi: 10.1176/ajp.2006.163.11.1905
6. Shelton RC. Steps Following attainment of remission: discontinuation of antidepressant therapy. *Prim Care Companion J Clin Psychiatry*. (2001) 3:168–74. doi: 10.4088/PCC.v03n0404
7. Israel JA. Remission in depression: definition and initial treatment approaches. *J Psychopharmacol.* (2006) 20:5–10. doi: 10.1177/1359786806064306
8. Helmreich I, Wagner S, Mergl R, Allgaier AK, Hautzinger M, Henkel V, et al. Sensitivity to changes during antidepressant treatment: a comparison of unidimensional subscales of the inventory of depressive symptomatology (IDS-C) and the hamilton depression rating scale (HAMD) in patients with mild major, minor or subsyndromal depression. *Eur Arch Psychiatry Clin Neurosci.* (2011) 262:291–304. doi: 10.1007/s00406-011-0263-x
9. Bagby RM, Ryder AG, SchuIIer DR, Marshall MB. The hamilton depression rating scale: has the gold standard become a lead weight? *Am J Psychiatry*. (2004) 161:2163–77. doi: 10.1176/appi.ajp.161.12.2163
10. Frank E, Prien RF, Jarrett RB, Keller MB, Kupfer DJ, Lavori PW, et al. Conceptualizationand rationale for consensus definitions of terms in majordepressive disorder remission, recovery, relapse, andrecurrence. *Arch Gen Psychiatry*. (1991) 48:851–5. doi: 10.1001/archpsyc.1991.01810330075011
11. Schlaepfer TE, Bewernick BH. Chapter 18 - Deep brain stimulation for major depression. In: Lozano AM, Hallett M, Editors. *Handbook of Clinical Neurology*. Amsterdam, Netherlands: Elsevier (2013). p. 235–43. doi: 10.1016/B978-0-444-53497-2.00018-8
12. Keller MB. Past, present, and future directions for definingoptimal treatment outcome in depression: remission and beyond. *JAMA*. (2003) 289:3152–60. doi: 10.1001/jama.289.23.3152
13. Wichniak A, Wierzbicka A, Walecka M, Jernajczyk W, Effects of antidepressants on sleep. *Curr Psychiatry Rep*. (2017) 19:63. doi: 10.1007/s11920-017-0816-4
14. Braestrup C, Sanchez C. Escitalopram: a unique mechanism of action. *Int J Psychiatry Clin Pract.* (2004) 8:11–3. doi: 10.1080/13651500410005496
15. Zhong H, Haddjeri N, Sánchez C. Escitalopram, an antidepressant with an allosteric effect at the serotonin transporter- a review of current understanding of its mechanism of action. *Psychopharmacology*. (2012) 219:1–13. doi: 10.1007/s00213-011-2463-5
16. Chockalingam R, Gott BM, Conway CR. Tricyclic antidepressants and monoamine oxidase inhibitors: are they too old for a new look? *Handb Exp Pharmacol.* (2019) 250:37–48. doi: 10.1007/164_2018_133
17. Abbing-Karahagopian V, Huerta C, Souverein PC, Abajo F, de Leufkens HGM, Slattery J, et al. Antidepressant prescribing in five European countries: application of common definitions to assess the prevalence, clinical observations, and methodological implications. *Eur J Clin Pharmacol.* (2014)70:849–57. doi: 10.1007/s00228-014-1676-z
18. Gillman PK. Tricyclic antidepressant pharmacology and therapeutic drug interactions updated. *Br J Pharmacol.* (2007) 151:737–48. doi: 10.1038/sj.bjp.0707253
19. World Health Organization. *Antidepressants (Tricyclic Antidepressants and Selective Serotonine Reuptake Inhibitors) in Treatment of Adults With Depression* (2012).
20. Bauer M, Pfennig A, Severus E, Whybrow PC, Angst J, Moller HJ. World Federation of Societies of biological psychiatry (WFSBP) guidelines for biological treatment of unipolar depressive disorders, part 1: update 2013 on the acute and continuation treatment of unipolar depressive disorders. *World J Biol Psychiatry*. (2013) 14:334–85. doi: 10.3109/15622975.2013.804195
21. Schulz KF, Grimes DA. Unequal group sizes in randomised trials: guarding against guessing. *Lancet*. (2002) 359:966–70. doi: 10.1016/S0140-6736(02)08029-7
22. Rosenberger WF, Lachin JM. *Randomization in Clinical Trials. 2nd ed*. Hoboken, NJ: Wiley (2016). p. 1–14. doi: 10.1002/9781118742112.ch1
23. Whiteford HA, Harris MG, McKeon G, Baxter A, Pennell C, Barendregt JJ, et al. Estimating remission from untreated major depression: a systematic review and meta-analysis. *Psychol Med*. (2013) 43:1569–85. doi: 10.1017/S0033291712001717
24. Vitriol V, Cancino A, Leiva- Bianchi M, Serrano C, Ballesteros S, Potthoff S, et al. The association between adverse childhood experiences with depression in adults consulting in primary care. *Rev Med Chil*. (2017) 145:1147–53. doi: 10.4067/s0034-98872017000901145
25. Machado-Vieira R, Baumann J, Wheeler-Castillo C, Latov D, Henter ID, Salvadore G, et al. The timing of antidepressant effects: a comparison of diverse pharmacological and somatic treatments. *Pharmaceuticals*. (2010) 3:19–41. doi: 10.3390/ph3010019
26. Beck AT, Steer RA, Ball R, Ranieri WF. Dimensions of the beck depression inventory-ii in clinically depressed outpatients. *J OF Clinical Psychol.* (1999). 55:117–28. doi: 10.1002/(sici)1097-4679(199901)55:1<117::aid-jclp12>3.0.co;2-a
27. Snaith P. What do depression rating scales measure? *Br J Psychiatry*. (1993) 163:293–8. doi: 10.1192/bjp.163.3.293
28. Wade A, Michael Lemming O, Bang Hedegaard K. Escitalopram 10 mg/day is effective and well tolerated in a placebo-controlled study in depression in primary care. *Int Clin Psychopharmacol*. (2002) 17:95–102. doi: 10.1097/00004850-200205000-00001
29. Burke WJ, Gergel I, Bose A. Fixed-dose trial of the single isomer SSRI escitalopram in depressed outpatients. *J Clin Psychiatry*. (2002) 63:331–6. doi: 10.4088/JCP.v63n0410
30. Furukawa T, McGuire H, Barbui C. Low dosage tricyclic antidepressants for depression. *Cochrane Database Syst Rev*. (2003) (3):CD003197. doi: 10.1002/14651858.CD003197
31. MacCue RE. Using tricyclic antidepressants in the elderly. *Clin Geriatr Med*. (1992) 8:323–34. doi: 10.1016/S0749-0690(18)30483-X
32. Wiese B. Geriatric depression: the use of antidepressants in the elderly. *BCMJ*. (2011) 53:341–7.
33. Noma H, Furukawa TA, Maruo K, Imai H, Shinohara K, Tanaka S, et al. Exploratory analyses of effect modifiers in the antidepressant treatment of major depression: individual-participant data meta-analysis of 2803 participants in seven placebo-controlled randomized trials. *J Affect Disord*. (2019) 250:419–24. doi: 10.1016/j.jad.2019.03.031
34. Uher R. Genes, environment, and individual differences in responding to treatment for depression. *Harv Rev Psychiatry*. (2011) 19:109–24. doi: 10.3109/10673229.2011.586551
35. Chekroud AM, Gueorguieva R, Krumholz HM, Trivedi MH, Krystal JH, McCarthy G. Reevaluating the efficacy and predictability of antidepressant treatments: a symptom clustering approach. *JAMA Psychiatry*. (2017) 74:370–8. doi: 10.1001/jamapsychiatry.2017.0025
36. Fried EI, Nesse RM. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR*D study. *J Affect Disord*. (2015) 172:96–102. doi: 10.1016/j.jad.2014.10.010
37. Cyranowski JM, Frank E, Young E. Adolescent onset of the gender difference in lifetime rates of major depression: a theoretical model. *Arch Gen Psychiatry*. (2000) 57:21–7. doi: 10.1001/archpsyc.57.1.21
38. Ford DE, Erlinger TP. Depression and C-reactive protein in US adults: data from the third national health and nutrition examination survey. *Arch Intern Med*. (2004) 164:1010–4. doi: 10.1001/archinte.164.9.1010

39. Albert PR, Why is depression more prevalent in women? *J Psychiatry Neurosci.* (2015) 40:219–21. doi: 10.1503/jpn.150205
40. Kornstein SG, Schatzberg AF, Thase ME, Yonkers KA, McCullough JP, Keitner GI, et al. Gender differences in treatment response to sertraline versus imipramine in chronic depression. *Am J Psychiatry.* (2000) 157:1445–2. doi: 10.1176/appi.ajp.157.9.1445
41. Entsuah R, Shaffer M, Zhang J. A critical examination of the sensitivity of unidimensional subscales derived from the hamilton depression rating scale to antidepressant drug effects. *J Psychiatr Res.* (2002) 36:437–48. doi: 10.1016/S0022-3956(02)00024-9
42. Husain MM, Rush AJ, Trivedi MH, McClintock SM, Wisniewski SR, Davis L, et al. Pain in depression: STAR*D study findings. *J Psychosom Res.* (2007) 63:113–22. doi: 10.1016/j.jpsychores.2007.02.009
43. Ang QQ, Wing YK, He Y, Sulaiman AH, Chiu NY, Shen YC, et al. Association between painful physical symptoms and clinical outcomes in East Asian patients with major depressive disorder: a 3-month prospective. *Int J Clin Pract.* (2009) 63:1041–9. doi: 10.1111/j.1742-1241.2009.02107.x
44. Agüera-Ortiz L, Failde I, Mico JA, Cervilla J, López-Ibor JJ. Pain as a symptom of depression: prevalence and clinical correlates in patients attending psychiatric clinics. *J Affect Disord.* (2011) 130:106–12.
45. Shimodera S, Kawamura A, Furukawa TA. Physical pain associated with depression: results of a survey in Japanese patients and physicians. *Compr Psychiatry.* (2012) 53:843–9. doi: 10.1016/j.comppsych.2011.11.004
46. Novick D, Montgomery W, Kadziola Z, Moneta V, Peng X, Brugnoli R, et al. Do concomitant pain symptoms in patients with major depression affect quality of life even when taking into account baseline depression severity? *Patient Prefer Adherence.* (2013) 7:463–70. doi: 10.2147/PPA.S41703
47. Kishi T, Matsuda Y, Mukai T, Matsunaga S, Yasue I, Fujita K, et al. A cross-sectional survey to investigate the prevalence of pain in Japanese patients with major depressive disorder and schizophrenia. *Compr Psychiatry.* (2015) 59:91–7. doi: 10.1016/j.comppsych.2015.02.004
48. Marangell LB, Clauw DJ, Choy E, Wang F, Shoemaker S, Bradley L, et al. Comparative pain and mood effects in patients with comorbid fibromyalgia and major depressive disorder: secondary analyses of four pooled randomized controlled trials of duloxetine. *Pain.* (2011) 152:31–7. doi: 10.1016/j.pain.2010.05.029
49. Lee P, Zhang M, Hong JP, Chua HC, Chen KP, Tang SW, et al. Frequency of painful physical symptoms with major depressive disorder in Asia: relationship with disease severity and quality of life. *J Clin Psychiatry.* (2009) 70:83–91. doi: 10.4088/JCP.08m04114
50. Novick D, Montgomery W, Bertsch J, Peng X, Brugnoli R, Haro JM. Impact of painful physical symptoms on depression outcomes in elderly Asian patients. *Int Psychogeriatr.* (2015) 27:305–12. doi: 10.1017/S1041610214002142
51. Pigeon WR, Hegel M, Unützer J, Fan M, Sateia MJ, Lyness JM, et al. Is insomnia a perpetuating factor for late-life depression in the impact Cohort? *Sleep.* (2008) 31:481–8. doi: 10.1093/sleep/31.4.481
52. Thase MME, Rush AJ, Manber R, Kornstein SG, Klein DN, Markowitz JC, et al. Differential effects of nefazodone and cognitive behavioral analysis system of psychotherapy on insomnia associated with chronic forms of major depression. *J Clin Psychiatry.* (2002) 63:493–500. doi: 10.4088/JCP.v63n0605
53. Moller HJ. Methodological aspects in the assessment of severity of depression by the hamilton depression scale. *Eur Arch Psychiatry Clin Neurosci.* (2001) 251:II13–20. doi: 10.1007/BF03035121
54. Stahl SM, Muntner N, Antidepressants/ classic antidepressants: tricyclic antidepressants. In: Stahl SM, editor. *Essential Psychopharmacology: Neuroscientific Basis and Practical Application*. Cambridge: Cambridge University Press (2003). p. 342–6.
55. Nock MK, Hwang I, Sampson N, Kessler RC, Angermeyer M, Beautrais A, et al. Cross-national analysis of the associations among mental disorders and suicidal behavior: findings from the WHO world mental health surveys. *PLoS Med.* (2009) 6:1000123. doi: 10.1371/journal.pmed.1000123
56. Xin LM, Chen L, Su YA, Yang FD, Wang G, Fang YR, et al. Risk factors for recent suicide attempts in major depressive disorder patients in China: results from a National Study. *Front Psychiatry.* (2018) 9:300. doi: 10.3389/fpsyt.2018.00300
57. Möller HJ. Is there evidence for negative effects of antidepressants on suicidality in depressive patients? a systematic review. *Eur Arch Psychiatry Clin Neurosci.* (2006) 256:476–96. doi: 10.1007/s00406-006-0689-8
58. Brent DA. Antidepressants and suicidality. *Psychiatr Clin North Am.* (2016) 39:503–12. doi: 10.1016/j.psc.2016.04.002
59. Baldessarini RJ, Lau WK, Sim J, Sum MY, Sim K. Suicidal risks in reports of long-term controlled trials of antidepressants for major depressive disorder II. *Int J Neuropsychopharmacol.* (2017) 20:281–4. doi: 10.1093/ijnp/pyw092
60. Gunnell D, Saperia J, Ashby D. Selective serotonin reuptake inhibitors (ssris) and suicide in adults: meta-analysis of drug company data from placebo controlled, randomised controlled trials submitted to the MHRA'S safety review. *BMJ.* (2005) 330:385. doi: 10.1136/bmj.330.7488.385
61. Stübner S, Grohmann R, Greil W, Zhang X, Müller-Oerlinghausen B, Bleich S, et al. Suicidal ideation and suicidal behavior as rare adverse events of antidepressant medication: current report from the amsp multicenter drug safety surveillance project. *Int J Neuropsychopharmacol.* (2018) 2:814–21. doi: 10.1093/ijnp/pyy048
62. Aarts N, Zuurbier LA, Noordam R, Hofman A, Tiemeier H, Stricker BH, et al. Use of selective serotonin reuptake inhibitors and sleep quality: a population-based study. *J Clin Sleep Med.* (2016) 12:989–95. doi: 10.5664/jcsm.5932
63. Dupuy JM, Ostacher MJ, Huffman J, Perlis RH, Nierenberg AA. A critical review of pharmacotherapy for major depressive disorder. *Int J Neuropsychopharmacol.* (2011) 14:1417–31. doi: 10.1017/S1461145711000083
64. Reynolds CF. Optimizing personalized management of depression: the importance of real-world contexts and the need for a new convergence paradigm in mental health. *World Psychiatry.* (2020) 19:266–8. doi: 10.1002/wps.20770

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Psychiatry

TYPE Mini Review
PUBLISHED 03 August 2022
DOI 10.3389/fpsyt.2022.984878

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University, Germany

REVIEWED BY
Axel Steiger,
Ludwig Maximilian University of Munich, Germany

*CORRESPONDENCE
Jorge Lopez-Castroman
jorge.lopezcastroman@chu-nimes.fr

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 02 July 2022
ACCEPTED 15 July 2022
PUBLISHED 03 August 2022

CITATION
Pastre M and Lopez-Castroman J (2022) Actigraphy monitoring in anxiety disorders: A mini-review of the literature. *Front. Psychiatry* 13:984878. doi: 10.3389/fpsyt.2022.984878

Actigraphy monitoring in anxiety disorders: A mini-review of the literature

Martin Pastre[1] and Jorge Lopez-Castroman[1,2,3]*

[1]Department of Psychiatry, CHU Nimes, Nimes, France, [2]Institut de Génomique Fonctionnelle (IGF), Université de Montpellier, CNRS, INSERM, Montpellier, France, [3]Centro de Investigacion Biomedical en Salud Mental (CIBERSAM), Instituto de Salud Carlos III (ISCIII), Madrid, Spain

Sleep disturbances and changes of activity patterns are not uncommon in anxiety disorders, but they are rarely the object of attention. Actigraphic monitoring of day and night activity patterns could provide useful data to detect symptom worsening, prevent risk periods, and evaluate treatment efficacy in those disorders. Thus, we have conducted a systematic search of the scientific literature to find any original study using actigraphic monitoring to investigate activity and sleep patterns in patients affected by any type of anxiety disorder according to the definition of the DSM-5. We found only six studies fulfilling these criteria. Three studies report significant findings in patients suffering from anxiety disorders. Overall, the samples and methods are heterogeneous. Although the authors support the interest of actigraphic monitoring in anxiety disorders, the evidence to date is very limited.

KEYWORDS

physical activity, circadian rhythm, sleep disturbances, wearable sensor, polysomnography, phobic and anxiety disorders

Introduction

Anxiety disorders are prevalent and disabling conditions characterized by excessive fear or anxiety, as well as a range of other cognitive and somatic symptoms. Comorbidity with other anxiety disorders and other mental disorders is very frequent, as well as with non-psychiatric medical conditions (1). Large epidemiological samples have estimated their lifetime prevalence at 14.5% in Europe (2) and 33.7% in the US (3), but these numbers comprise obsessive-compulsive and related disorders and trauma and stressor-related disorders, which no longer belong to the category of anxiety disorders in the DSM-5 (4). The 12-month prevalence reported in DSM-5 for adult anxiety disorders ranges from 1 to 3% in the case of panic disorder (PD), 2–7% for social phobia, 0.4–9% for generalized anxiety disorder (GAD), 1–2% for agoraphobia and 6–9% for specific phobias, with the highest prevalence rates being generally reported in the US (4). The wide ranges in these figures are due, among other reasons, to differences in the diagnostic assessment methods and the target populations [for a detailed review see Bandelow and Michaelis (5)]. The 2019 Global Burden of Disease study provides an estimate of the disability associated with these disorders, which are the 24th leading cause in disability-adjusted life-years (the 6th in young people aged 10–24 years) (6).

Among the symptoms of anxiety disorders, sleep disturbances and changes in physical activity (PA) patterns are rarely the object of attention. Contrary to

depression, these symptoms are not part of the diagnostic criteria except for GAD, which includes an item about sleeping difficulties, but all anxiety disorders seem to be associated with some degree of sleep disturbances and changes in PA. With regards to sleep, a recent meta-analysis (7) based on polysomnography or self-reported sleep data in controlled studies has shown that patients suffering from anxiety disorders have less sleep continuity (Hedge's $g = -0.49$), an average of 21 min less in total sleep time ($g = -0.40$) and more subjective sleep disturbances ($g = 2.16$) compared to healthy controls with no mental disorder. It should be noted that GAD patients reported the highest scores of subjective sleep disturbances ($g = 5.55$).

Concerning PA patterns, anxiety disorders are characterized by excessive daytime arousal or restlessness according to heart rate and activity monitoring (8). Excessive arousal however does not imply more PA. Two recent meta-analyses found that anxiety disorders and anxiety symptoms are associated with sedentary behavior (9, 10). Also, symptomatic forms of anxiety were prospectively associated with less PA two years later according to a large epidemiological survey with almost 3,000 persons in the Netherlands (11). The association seems to be bidirectional since low sports participation at baseline was associated with symptomatic anxiety two years later. This has implications for treatment. A meta-analysis of randomized controlled trials proved that both aerobic and anaerobic activity reduces the intensity of anxiety symptoms (12), and PA has been proposed as an effective adjunctive treatment for anxiety disorders (13).

Actigraphy can be used to monitor activity rhythms and sleep in mental disorders (14). Actigraphic devices can be routinely used in daily life, have a limited cost compared to polysomnography, and they provide quantifiable objective data that substantially improve the utility of self-reported measures (15). A recent retrospective study investigated the phenomenon of "misperception of sleep" (discrepancies between objective and subjective measures of sleep), and showed that it is a common feature in anxiety disorders (16). Considering all the above and the absence of any review on the topic, we decided to conduct a systematic review of scientific papers using actigraphic monitoring to measure activity patterns and sleep in anxiety disorders.

Methods

Selection of studies

We selected all studies according to the following eligibility criteria: (i) original studies published until June 2022 in English, French or Spanish language, (ii) actigraphy measures were used for activity monitoring, including activity patterns during the day and/or sleep parameters at night, (iii) the study samples comprised adult patients with any anxiety disorder diagnosis included in the corresponding DSM-5 category. Therefore, studies investigating trauma or stress-related disorders and obsessive-compulsive and related disorders were excluded. We also excluded studies in which the diagnostic criteria of anxiety disorders were not clearly respected (for example, when only anxiety symptoms were reported) or the actigraphic monitoring was limited to an experimental procedure in a clinical setting (not reflecting daily activity). We followed PRISMA 2020 checklist for systematic reviews and published the protocol on the PROSPERO registry for systematic reviews (CRD42022323708).

Data sources and search strategy

To identify potential papers, we searched three databases: PubMed, WebOfScience, and PsycINFO until June 2022 with the following equation terms: [("Anxiety Disorders" OR "Social Anxiety" OR "Generalized Anxiety Disorder" OR "Panic disorder" OR "Social, Phobia" OR "phobic disorder" OR "Phobia, Specific" OR agoraphobia) NOT ("Obsessive-Compulsive Disorder" OR "Anxiety, Separation" OR "Neurocirculatory Asthenia" OR "Neurotic Disorders")] AND (actigrap* OR actimet* OR actograp* OR actomet* OR accelerometer).

The title and abstract of each potential paper were screened by two reviewers working independently (MP and JLC). Zotero software was used for the management of records. The full text of eligible studies was then reviewed independently by the same two reviewers to assess all inclusion and exclusion criteria.

Data extraction and quality assessment

We extracted all relevant data from selected papers using a data chart. The quality of each study was assessed using a modified version of the Effective Public Health Practice Project Tool (17) that we built for the purpose of this review. We used the EPHPP tool because the designs of selected studies were highly heterogeneous. Sections D, G, and H of the original scale were not relevant because our review did not include any interventional study and were therefore suppressed from the global rating. Likewise, section C was also revised because the first part of the section (Q1, "Were there important differences between groups prior to the intervention?") was not applicable. Thus, we considered only the second part of the section (Q2, "Indicate the percentage of relevant confounders that were controlled") for the rating of section C.

Abbreviations: GAD, Generalized Anxiety Disorder; PA, Physical Activity; PD, Panic Disorder; TST, Total Sleep Time; WASO, Wake After Sleep Onset.

TABLE 1 Description of selected studies.

	Sample	Comparison group	Type of anxiety disorder	Outcomes	Main results
Helgadóttir et al. (25)	22	Patients with major depression or comorbid anxiety and depression	GAD, PD, SAD, (PTSD)	Average counts per min, % of sedentary bouts, % of activity bouts, total time in sedentary bouts, number of sedentary bouts	No significative differences in physical activity patterns between depressive and anxious participants
Koolhaas et al. (24)	147 + 59*	Populational cohort	GAD, PD, AgPh, SAD, SPh	Hours per day of sedentary behavior (defined as <199 count per min during waking hours)	Cross-sectionally: no significative association between anxiety disorders and sedentary behavior after adjustment on cofounders. Longitudinally: no significative association between sedentary time and subsequent development of anxiety disorder
Luik et al. (20)	144	Populational cohort	GAD, PD, AgPh, SAD, SPh	Fragmentation of the rhythm, stability of the rhythm over days, timing of the rhythm. TST, sleep onset latency and WASO	Anxiety disorders associated with more fragmented rhythm (intradaily variability), independent of covariates (OR: 1.39 per 1 SD, 95% CI: [1.13; 1.70], $p = 0.002$). GAD (n = 39) associated with more fragmented rhythms (OR: 1.75 per 1 SD, 95% CI: [1.20; 2.55], $p = 0.004$), but also shorter TST (OR: 0.66 per 1 h, 95% CI: [0.45; 0.97], $P = 0.033$)
Sakamoto et al. (27)	16	None	PD	Mesor, circadian amplitude, acrophase	Association between frequency of panic attacks and mesor ($r = 0.55$, $p = 0.03$), and between HAM-A score and mesor ($r = 0.62$, $p = 0.01$)
Todder and Baune (21)	15	Healthy controls	PD	Sleep time (%), sleep efficiency (%), index of fragmentation of sleep	No significative difference between patients and controls, or before and after treatment by escitalopram
Wainberg et al. (22)	4847	Psychiatric outpatients	GAD, PD, AgPh, SAD, SPh	Sleep efficiency, longest sleep bout, wake-up/bed-time, WASO, number of awekenings, number of naps, bedtime variability, sleep duration variability	Anxiety disorders associated with sleep disturbances: WASO (beta coefficient for linear regression = 0.04), later bed-time (0.03), later wake-up time (0.04), sleep efficiency (−0.05), number of awakenings (0.04), longest sleep bout (−0.04), number of naps (0.04), bedtime variability (0.03) and sleep duration variability (0.04)

*They were 147 prevalent cases of anxiety disorders used in the cross-sectional analysis, and 59 incident cases used in the longitudinal analysis. PD, Panic Disorder; GAD, Generalized Anxiety Disorder; SAD, Social Anxiety Disorder; SPh, Specific Phobia; AgPh, Agoraphobia; PTSD, Post Traumatic Stress Disorder.

Results

The search retrieved 201 potential papers (80 from PubMed, 85 from WebOfScience, and 36 from Psychinfo). After the removal of duplicates (46 papers), 129 articles were excluded based on title or abstract (not relevant to the topic). Among the 26 papers that were read in full to assess eligibility, 20 were excluded because they did not fulfill the criteria. Most of the excluded papers focused on anxiety symptoms only and did not consider anxiety disorder diagnoses. Three papers (15, 18, 19) used pooled diagnostic data of anxiety disorders and mood disorders. We contacted the authors to obtain specific data on anxiety disorders, but we did not receive an answer. Six articles were included in the mini-review (see Table 1 for a summary of principal results). A flow diagram based on PRISMA 2020 guidelines is shown in Figure 1. Overall, the studies presented a substantial risk of bias (see Supplementary Table 1 for quality assessment).

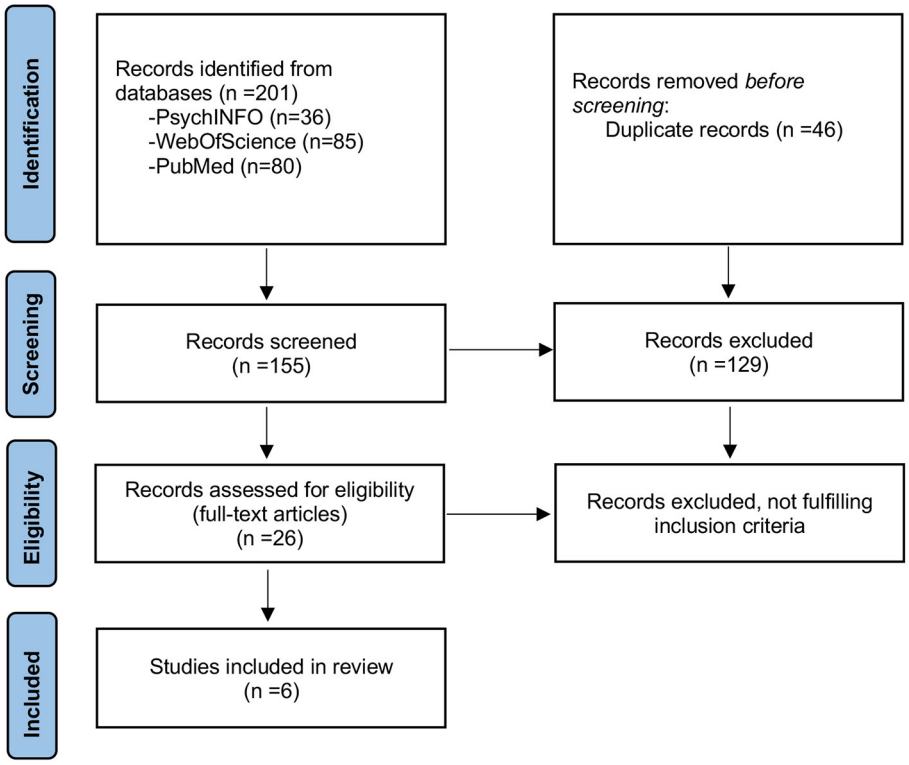


FIGURE 1
Flowchart of selected studies.

Hereafter we describe the six included studies according to their quality score (highest quality studies are presented first).

Todder and Baune (21) followed prospectively a cohort of 15 women with PD before and after the instauration of the antidepressant escitalopram (up to 10 mg/day), seeking changes in actigraphic parameters of sleep. There was a "wash-out" period followed by 4 weeks of treatment with continuous actigraphic monitoring. Assessment scales such as the Panic and Agoraphobic scale, the Hamilton anxiety scale (HAM-A) and the Pittsburgh sleep quality index were completed once per week. Patients under benzodiazepine treatment were excluded. The Mini-International Neuropsychiatric Interview (MINI) inventory was used to confirm the diagnoses. Night-time activity was characterized by sleep time (percentage of time asleep between onset and end of sleep), and sleep efficiency (ratio between actual sleep time and total time in bed). These outcomes did not change after treatment and did not differ with those of a control group of female healthy administration workers. At 4 weeks, there was no significant difference in sleep patterns between patients that showed a clinical improvement (>50% HAM-A score) and those that did not.

The study by Luik et al. (20) analyzed cross-sectionally the circadian activity and sleep patterns of patients with anxiety disorders in a populational cohort (>45 years old) from the Rotterdam Study (23). 96 h of actigraphic data was collected from 1,714 people. Anxiety disorders ($n = 141$) were diagnosed using the Composite International Diagnostic Interview (CIDI). Nonparametric measures were used to assess activity rhythms, namely: interday stability, intraday variability (indicative of rhythm fragmentation, i.e., transitions from an active to an inactive state), and dominant rest phase onset (start time of lowest activity period). Concerning sleep, total sleep time (TST), sleep onset latency and wake after sleep onset (i.e., time periods of wakefulness after sleep onset, WASO) were recorded. The reference category for logistic regression analysis comprised participants with no clinical symptoms of depression or anxiety ($n = 1,441$). There was a significant association between fragmented rhythms and the prevalence of anxiety disorders, independently of covariates (OR: 1.39 per 1 SD of intradaily variability, [1.13; 1.70], $p = 0.002$). The significant difference persisted after the exclusion of 47 patients with anxiety disorders and substantial depressive symptoms (Center for Epidemiologic Studies-Depression scale >15). The authors also found that GAD ($n = 39$) was associated with more fragmented rhythms (OR: 1.75 per 1 SD, [1.20; 2.55], $p = 0.004$) and a shorter TST (OR: 0.66 per 1 h, [0.45; 0.97], $p = 0.033$) than the reference group after adjusting on covariates.

Koolhaas et al. (24) studied the relationship of anxiety disorders and sedentary behavior with the data of the Rotterdam Study. A subsample of participants was monitored with an

actigraph for a period of 7 days (n = 1,841). Activity level during waking hours was measured by the number of counts per minute (a count corresponding to a single movement in any direction captured by the actigraphic sensors). For a given subject, the time of sedentary behavior corresponds to the time during which the activity is <199 counts per min. Diagnoses of anxiety disorders were obtained at baseline using the CIDI. Participants with anxiety disorders (n = 147 prevalent cases) reported significantly more sedentary time than the rest of the sample in unadjusted analyses, but after controlling for lifestyle factors (namely disability, smoking, and occupational status) the association did no longer exist. Of note, sedentary behavior at baseline was not associated with the emergence of anxiety disorders during the average 5.7 years of follow-up time (n = 59 incident cases).

Wainberg et al. (22) conducted a *post-hoc* cross-sectional analysis of 89 000 individual actigraphic data from the UK-Biobank (a community-based prospective cohort study) to study sleep parameters in anxiety disorders. Anxiety disorders were identified through registered codes of the International Classification of Diseases-10th edition (F40, F41). Several sleep features were measured: bed and wake-up times, sleep duration (defined here as the total duration of night sleep bouts), WASO, sleep efficiency, number of awakenings, duration of longest sleep bout, number of naps, and variability in bedtime/in sleep duration. The presence of any anxiety disorder was associated with sleep disturbances, but effect sizes were small. Compared to healthy participants, patients with anxiety disorders presented a longer WASO (with a beta coefficient for linear regression of 0.04), as well as longer bedtime and wake-up time (0.03 and 0.04 respectively). The same pattern was observed for bedtime variability and sleep duration variability. Sleep efficiency (−0.05) and the duration of the longest sleep bout (−0.04) was decreased, and they experienced more awakenings (0.04).

Helgadóttir et al. (25) used actigraphic data of 165 anxious and/or depressed Swedish adults from the Regassa randomized controlled study (26) to investigate their level of sedentary behavior. All participants, who had a minimum score of 10 on the Patient Health Questionnaire, wore an actigraph for seven days. Diagnoses were obtained with the MINI. They measured activity level using the same proxy described above (number of counts per minute), considering <100 counts per min during twenty consecutive minutes as a sedentary activity bout, 100–1,951 counts per minute as light PA, and more than 1,951 counts per min as moderate to vigorous activity. They then calculated the total time spent in sedentary bouts, as well as the number of sedentary bouts. Twenty-two participants were diagnosed with an anxiety disorder, while 121 had depressive and anxiety disorders at the same time. All the participants were rather sedentary, but there was no statistical difference in activity measures between diagnostic groups.

Finally, Sakamoto et al. (27) investigated the effect of PD severity on 24h activity patterns in 16 outpatients. The participants were recruited through advertisements and assessed with the HAM-A and the Panic Disorder Severity Scale on the first day of the study. They all received a diagnosis of PD with agoraphobia (DSM-IV). Only two patients were male. Most of them were treated with antidepressants and/or benzodiazepines. The patients used electronic diaries (watch-type computers) for 14 days to note information on any panic attack. Also, the intensity of the symptomatology was actively assessed with daily ecological momentary assessment questions. Investigators ran a "cosinor" analysis to describe the timing and amplitude of PA, considered as a circadian process with a particular rhythm. This model provides the estimation of the "mesor" (or corrected amplitude mean) of the circadian rhythm as well as the "acrophase" (peak time in the model). Pearson's correlation analysis showed a significant association between the mesor (from double cosinor analysis) and the frequency of panic attacks ($r = 0.55$, $p = 0.03$) as well as the mesor and the HAM-A score ($r = 0.62$, $p = 0.01$).

Discussion

A large share of the recent literature about actigraphic measures in psychiatry is focused on mood disorders. In contrast, we decided to review systematically the objective alterations of sleep and activity patterns associated with anxiety disorders that so far have been the object of only a handful of studies. Although the results of our review show that these symptoms can be objectively detected in anxiety disorders, for the moment there is very limited evidence supporting the use of actigraphic measures to monitor their evolution or severity.

Four papers studied anxiety disorders as a general category. All of them were secondary studies, based on large datasets. Luik et al. (20) found fragmented 24 h circadian rhythm measures in anxiety disorders and specifically in GAD, which was also associated with shorter TST. Participants diagnosed with anxiety disorders in the UK-Biobank were more likely to have a disturbed sleep (i.e., notably higher WASO, more awakenings and less sleep efficiency) than healthy controls, but this pattern was shared across psychiatric conditions. These results are consistent with previous self-reported or polysomnographic data regarding altered sleep continuity and lower TST in anxiety disorders (7). In contrast, Koolhaas et al. (24) did not find any association, either cross-sectionally or longitudinally, between diurnal sedentary behavior and anxiety disorders, contradicting studies based on self-reports (9, 10). According to the authors, this discrepancy can be explained by the insufficient precision of actigraphic measures and the fact that previous studies did not control for important confounders, such as disability or occupational status. Helgadóttir et al. (25) also failed to find any differences in sedentary behavior when comparing anxiety-disordered participants and those suffering from depression or comorbid anxiety and depression.

The physiopathological relationship between sleep disturbances and anxiety can be better understood with the results of a recent study. The anxiety symptoms that emerged in patients submitted to sleep deprivation were associated in functional MRI with an hypoactivity of the medial prefrontal cortex, involved in emotional control, and an hyperactivity in the amygdala and dorsal anterior cingulate cortex, responsible of the reactivity to negative emotions (28). Also, the amount of slow-wave sleep predicted the reactivation of medial prefrontal cortex the next day, suggesting an anxiolytic effect of this particular phase of sleep that is known to be shortened in patients suffering from GAD (29) and PD (30).

Some actigraphic studies focused on specific anxiety disorders. Todder and Baune (21) expected to find an actigraphic marker of the clinical response to antidepressants in PD but they did not find any association with sleep disturbances. Sakamoto et al. (27) showed, by using a proxy to measure circadian amplitude in a clinical sample of panic disorders with phobic avoidance, that patients with a more severe form of panic disorder showed greater motor activity. In the same way, a study using a motion sensor found that PD patients with a higher level of phobic avoidance had greater motor activity than controls and PD patients with a lower level of avoidance (8). In the case of GAD, a recent paper (31) investigated restlessness in patients with this diagnosis and healthy controls using actigraphy through a threat-exposure task. Restlessness, which is a subjective feeling close to hyperarousal, is a core feature of GAD and one of its diagnostic criteria. In this study the GAD group did not show greater actigraphic movement magnitude than controls, despite having a significantly higher self-reported restlessness level at baseline and during the threat exposure. Participants with restlessness had a significantly heightened movement level at baseline and through the stages of the increasingly threatening task compared to those without. Moreover, objective measures failed to confirm the subjective restlessness reported by people with GAD, a contradiction that was described by the authors as the "reactivity paradox": self-reported restlessness does not match objective measures of threat reactivity (31). Overall, these findings suggest that restlessness in GAD could constitute a chronic state of arousal rather than a tendency to overreact while anticipating or being exposed to a threat.

In this review, we excluded papers based on patients presenting anxiety symptoms only because of the transdiagnostic and unspecific nature of these symptoms. However, anxiety symptoms can also impair sleep and activity features. Spira et al. (32) in a sample of older adults with primary insomnia, showed that trait anxiety was associated with greater actigraphy-measured WASO. Studies with pooled samples of depressive and anxiety disorders were also excluded, although we retrieve in clinical samples with this comorbidity the same types of activity and sleep alterations patterns as for anxiety disorders alone. By monitoring sleep, circadian rhythm and PA in a sample of 359 participants with anxiety and/or depressive disorders, Difransesco et al. (15) found that currently anxious and/or depressed patients were less active (with a lower circadian relative amplitude between day-time and night-time activity levels) than controls. Interestingly, the more severe the symptoms of anxiety and depression, the lower the level of PA and the relative amplitude of circadian rhythms. In the same study, participants diagnosed with anxiety and/or depressive disorders reported more insomnia and longer sleep duration, but this difference was not present with objective measures.

Actigraphy has also been used as a prognostic biomarker in the field of anxiety disorders. Jacobson et al. (33) investigated actigraphic measures of movement patterns in GAD and PD. Participants were followed up to 18 years, and a deep learning model based on various activity and sleep features could predict symptom worsening over time with an AUC = 0.696 (84.6% sensitivity, 52.7% specificity). The same authors found with passive data from a wearable accelerometer that patients with higher social anxiety symptoms had lower movement amplitudes (34).

There are several limits to the scope of this review. First, each anxiety disorder might have distinct sleep and activity patterns, despite their common physiopathology and very frequent comorbidity, but most of the included studies considered them in a single category. Second, the restricted number and quality of the studies precludes any strong interpretation of the existing evidence. We were unable to obtain detailed data on the studies that pooled depressive and anxiety disorders. Furthermore, only two of the selected studies (20, 21) took into account the potential interaction of benzodiazepine and/or antidepressant medication in the relationship between sleep and anxiety, despite their wide prescription in anxiety disorders (and particularly among anxious patients presenting sleep disorders) and the well-documented effects of these drugs on sleep architecture (35, 36).

In summary, few studies have yet examined objectively sleep and daily activity changes associated with anxiety disorders. The extant studies are heterogeneous, with an overall high risk of bias. Only half of those included report statistically significant results linking anxiety disorders with disturbances in sleep and activity patterns, and the results are sometimes conflicting. Overall, we want to point out the need for new and specific research in the field, given the burden caused by these disorders and the potential interest of ecological interventions, i.e., based on daily life activities, to improve their prognosis. Characterizing activity and sleep change patterns in people suffering from anxiety disorders might provide useful knowledge to monitor the effects of pharmacological and behavioral interventions.

Author contributions

MP and JL-C reviewed the articles and drafted the manuscript. All authors contributed to the conception and design of the study and approved the final version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2022.984878/full#supplementary-material

References

1. Meuret AE, Tunnell N, Roque A. Anxiety disorders and medical comorbidity: treatment implications. *Adv Exp Med Biol.* (2020) 1191:237–61. doi: 10.1007/978-981-32-9705-0_15

2. Alonso J, Lépine JP. ESEMeD/MHEDEA 2000 Scientific Committee. Overview of key data from the European Study of the Epidemiology of Mental Disorders (ESEMeD). *J Clin Psychiatry.* (2007) 68:3–9.

3. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen HU. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res.* (2012) 21:169–84. doi: 10.1002/mpr.1359

4. Association AP. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). US: American Psychiatric Pub (2013). 1505. p.

5. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. *Dialogues Clin Neurosci.* (2015) 17:327–35. doi: 10.31887/DCNS.2015.17.3/bbandelow

6. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet.* (2020) 396:1204–22. doi: 10.1016/S0140-6736(20)30925-9

7. Cox RC, Olatunji BO. Sleep in the anxiety-related disorders: a meta-analysis of subjective and objective research. *Sleep Med Rev.* (2020) 51:101282. doi: 10.1016/j.smrv.2020.101282

8. Clark DB, Taylor CB, Hayward C, King R, Margraf J, Ehlers A, et al. Motor activity and tonic heart rate in panic disorder. *Psychiatry Res.* (1990) 32:45–53. doi: 10.1016/0165-1781(90)90134-Q

9. Allen MS, Walter EE, Swann C. Sedentary behaviour and risk of anxiety: a systematic review and meta-analysis. *J Affect Disord.* (2019) 242:5–13. doi: 10.1016/j.jad.2018.08.081

10. Stanczykiewicz B, Banik A, Knoll N, Keller J, Hohl DH, Rosińczuk J, et al. Sedentary behaviors and anxiety among children, adolescents and adults: a systematic review and meta-analysis. *BMC Public Health.* (2019) 19:459. doi: 10.1186/s12889-019-6715-3

11. Hiles SA, Lamers F, Milaneschi Y, Penninx BWJH. Sit, step, sweat: longitudinal associations between physical activity patterns, anxiety and depression. *Psychol Med.* (2017) 47:1466–77. doi: 10.1017/S0033291716003548

12. Jayakody K, Gunadasa S, Hosker C. Exercise for anxiety disorders: systematic review. *Br J Sports Med.* (2014) 48:187–96. doi: 10.1136/bjsports-2012-091287

13. Kandola A, Vancampfort D, Herring M, Rebar A, Hallgren M, Firth J, et al. Moving to beat anxiety: epidemiology and therapeutic issues with physical activity for anxiety. *Curr Psychiatry Rep.* (2018) 20:63. doi: 10.1007/s11920-018-0923-x

14. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP, et al. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep.* (2003) 26:342–92. doi: 10.1093/sleep/26.3.342

15. Difrancesco S, Lamers F, Riese H, Merikangas KR, Beekman ATF, van Hemert AM, et al. Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: a 2-week ambulatory assessment study. *Depress Anxiety.* (2019) 36:975–86. doi: 10.1002/da.22949

16. Liang Y, Zhao X, Zhang C, Liu G, Lu B, Han L, et al. Sleep misperception and associated factors in patients with anxiety-related disorders and complaint of insomnia: a retrospective study. *Front Neurol.* (2022) 13:836949. doi: 10.3389/fneur.2022.836949

17. Thomas BH, Ciliska D, Dobbins M, Micucci SA. process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid Based Nurs.* (2004) 1:176–84. doi: 10.1111/j.1524-475X.2004.04006.x

18. Difrancesco S, Riese H, Merikangas K, Shou H, Zipunnikov V, Antypa N, et al. Sociodemographic, health and lifestyle, sampling, and mental health determinants of 24-hour motor activity patterns: observational study. *J Med Internet Res.* (2021) 23:e20700. doi: 10.2196/20700

19. Klumpp H, Roberts J, Kapella MC, Kennedy AE, Kumar A, Phan KL, et al. Subjective and objective sleep quality modulate emotion regulatory brain function in anxiety and depression. *Depress Anxiety.* (2017) 34:651–60. doi: 10.1002/da.22622

20. Luik AI, Zuurbier LA, Direk N, Hofman A, Van Someren EJW, Tiemeier H, et al. 24-hour activity rhythm and sleep disturbances in depression and anxiety: a population-based study of middle-aged and older persons. *Depress Anxiety.* (2015) 32:684–92. doi: 10.1002/da.22355

21. Todder D, Baune BT. Quality of sleep in escitalopram-treated female patients with panic disorder. *Hum Psychopharmacol.* (2010) 25:167–73. doi: 10.1002/hup.1088

22. Wainberg M, Jones SE, Beaupre LM, Hill SL, Felsky D, Rivas MA, et al. Association of accelerometer-derived sleep measures with lifetime psychiatric diagnoses: A cross-sectional study of 89,205 participants from the UK Biobank. *PLoS Med.* (2021) 18:e1003782. doi: 10.1371/journal.pmed.1003782

23. Hofman, A, Darwish Murad, S, van Duijn CM, Franco OH, et al. (2013). The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol.* 28, 889–926. doi: 10.1007/s10654-013-9866-z

24. Koolhaas, C, van Rooij F, Kocevska D, Luik A, Ikram M, et al. (2019). Objectively measured sedentary time and mental and cognitive health: cross-sectional and longitudinal associations in The Rotterdam Study. *Ment Health Phys Act*. 17. doi: 10.1016/j.mhpa.2019.100296

25. Helgadóttir B, Forsell Y, Ekblom Ö. Physical activity patterns of people affected by depressive and anxiety disorders as measured by accelerometers: A cross-sectional study. *PLoS ONE.* (2015) 10:e0115894. doi: 10.1371/journal.pone.0115894

26. Hallgren M, Kraepelien M, Öjehagen A, Lindefors N, Zeebari Z, Kaldo V, et al. Physical exercise and internet-based cognitive-behavioural therapy in the treatment of depression: randomised controlled trial. *Br J Psychiatry J Ment Sci.* (2015) 207:227–34. doi: 10.1192/bjp.bp.114.160101

27. Sakamoto N, Yoshiuchi K, Kikuchi H, Takimoto Y, Kaiya H, Kumano H, et al. Panic disorder and locomotor activity. *Biopsychosoc Med*. (2008) 2:23. doi: 10.1186/1751-0759-2-23

28. Ben Simon E, Rossi A, Harvey AG, Walker MP. Overanxious and underslept. *Nat Hum Behav*. (2020) 4:100–10. doi: 10.1038/s41562-019-0754-8

29. Fuller KH, Waters WF, Binks PG, Anderson T. Generalized anxiety and sleep architecture: a polysomnographic investigation. *Sleep*. (1997) 20:370–6. doi: 10.1093/sleep/20.5.370

30. Arriaga F, Paiva T, Matos-Pires A, Cavaglia F, Lara E, Bastos L, et al. The sleep of non-depressed patients with panic disorder: a comparison with normal controls. *Acta Psychiatr Scand*. (1996) 93:191–4. doi: 10.1111/j.1600-0447.1996.tb10630.x

31. Franklin AR, Mathersul DC, Raine A, Ruscio AM. Restlessness in generalized anxiety disorder: using actigraphy to measure physiological reactions to threat. *Behav Ther*. (2021) 52:734–44. doi: 10.1016/j.beth.2020.09.004

32. Spira AP, Friedman L, Aulakh JS, Lee T, Sheikh JI, Yesavage JA, et al. Subclinical anxiety symptoms, sleep, and daytime dysfunction in older adults with primary insomnia. *J Geriatr Psychiatry Neurol*. (2008) 21:149–53. doi: 10.1177/0891988707317120

33. Jacobson NC, Lekkas D, Huang R, Thomas N. Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17-18 years. *J Affect Disord*. (2021) 282:104–11. doi: 10.1016/j.jad.2020.12.086

34. Jacobson NC, Summers B, Wilhelm S. Digital biomarkers of social anxiety severity: digital phenotyping using passive smartphone sensors. *J Med Internet Res*. (2020) 22:e16875. doi: 10.2196/16875

35. Hutka P, Krivosova M, Muchova Z, Tonhajzerova I, Hamrakova A, Mlyncekova Z, et al. Association of sleep architecture and physiology with depressive disorder and antidepressants treatment. *Int J Mol Sci*. (2021) 22:1333. doi: 10.3390/ijms22031333

36. de Mendonça FMR, de Mendonça GPRR, Souza LC, Galvão LP, Paiva HS, de Azevedo Marques Périco C, et al. Benzodiazepines and sleep architecture: A systematic review. *CNS Neurol Disord Drug Targets*. (2021). doi: 10.2174/1871527320666210618103344. [Epub ahead of print].

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 18 August 2022
DOI 10.3389/fpsyt.2022.956566

OPEN ACCESS

EDITED BY
Luodan Yang,
Louisiana State University in Shreveport, United States

REVIEWED BY
Jing Shen,
Tianjin Stomatological Hospital, China
Shunming Zhang,
Lund University, Sweden
Yeqing Gu,
Chinese Academy of Medical Sciences and Peking Union Medical College, China
Jie Zhang,
Peking University Hospital of Stomatology, China

*CORRESPONDENCE
Dongwang Zhu
zhudongwang126@126.com

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 30 May 2022
ACCEPTED 28 July 2022
PUBLISHED 18 August 2022

CITATION
Qiao F, Zhang M, Zhang T and Zhu D (2022) Dental anxiety is related to postoperative symptoms in third molar surgery. *Front. Psychiatry* 13:956566. doi: 10.3389/fpsyt.2022.956566

Dental anxiety is related to postoperative symptoms in third molar surgery

Feng Qiao[1], Menghua Zhang[2], Tingting Zhang[1] and Dongwang Zhu[1]*

[1]Department of Oral and Maxillofacial Surgery, School and Hospital of Stomatology, Tianjin Medical University, Tianjin, China, [2]Graduate School and Hospital of Stomatology, Tianjin Medical University, Tianjin, China

Purpose: To examine the association of preoperative dental anxiety with the severity of postoperative symptoms among patients undergoing lower third molar (LM3) extraction surgery.

Materials and methods: We conducted a hospital-based prospective study with a sample size of 213 patients. All the patients underwent LM3 extraction surgery at the Stomatology Hospital of Tianjin Medical University. Preoperative dental anxiety was measured using the Dental Anxiety Scale for Third Molar Surgery (DAS-TMS) and classified into four categories: No anxiety, Some unease, Anxious, and Very anxious. The primary outcome was defined using the postoperative symptom severity scale on the seventh day after surgery. The patients' clinical characteristics, radiologic features, and surgery-related variables were used as control variables. Bivariate analysis involved Fisher's exact test and Kruskal–Wallis test. Multivariable logistic analysis was used to assess preoperative dental anxiety in relation to the severity of postoperative symptoms. We applied a two-piecewise regression model to examine the potential non–linear associations.

Results: The mean (SD) dental anxiety score was 10.56 (3.84). The proportion of dental anxiety was as follows: No anxiety, 7.5%; Some unease, 46.9%; Anxious, 31.0%; Very anxious, 14.6%. The multivariable-adjusted ORs with 95% CIs of postoperative symptoms were 1.00 for No anxiety, 3.63 (0.90–14.68) for Some unease, 5.29 (1.25–22.33) for Anxious, and 4.75 (1.02–22.18) for Very anxious (P for trend = 0.047). The risk of serious postoperative symptoms increased with the dental anxiety level up to 7 points (adjusted OR 1.94, 95% CI 1.12–3.74; $P = 0.012$). When the dental anxiety level exceeded 7 points, the level of DAS-TMS was not associated with the risk of serious postoperative symptoms (OR 0.98, 95% CI 0.88–1.08; $P = 0.756$).

Conclusions: Findings suggest that dental anxiety is associated with a risk of serious postoperative symptoms following LM3 removal. The degree of dental anxiety in patients before LM3 extraction surgery should be of concern to clinicians.

KEYWORDS

dental anxiety, postoperative symptom, logistic regression, non–linear relation, lower third molar extraction

Introduction

Removal of third molars is a common surgical procedure performed in maxillofacial and oral surgery. Patients with wisdom tooth extractions have the highest dental anxiety levels (1, 2). Preoperative dental anxiety results in a delay or avoidance of dental treatment and, consequently, poorer oral health and oral health-related quality of life (3, 4). Recent studies have shown that psychological anxiety leads to the activation of the body's stress reaction and slower postoperative recovery (5). Dental anxiety has also been shown to be associated with postoperative pain (5–7).

Recently, patient recovery has attracted considerable attention in the field of third molar surgery (8–12). Questionnaires on evaluating the severity of postoperative symptoms have become more useful and are widely used (13–15). Additionally, a previous study found that patient anxiety affects the difficulty of impacted lower third molar extraction (16). The surgical difficulty is frequently associated with considerable postoperative adverse effects such as pain, edema, and trismus (9, 17, 18). Previous studies have also indicated that the complexity of the surgical operation has been associated with postoperative symptoms (14). These results suggest that preoperative dental anxiety may be associated with postoperative symptoms. However, the specific relationship between preoperative dental anxiety and the severity of postoperative symptoms remains unclear.

Several studies have evaluated the association between dental anxiety and LM3 surgery (19–21). Patients with high dental anxiety experience greater trismus and more pain (19). Onwuka et al. reported that preoperative dental anxiety is more common in women (21). However, confounding factors have not been fully incorporated into multivariable regression models for control (19). A direct independent association between preoperative dental anxiety and postoperative symptoms in patients undergoing lower third molar removal is still unestablished. Moreover, the significance of the non–linear relationship still requires further clarification. Therefore, exploring the non–linear relationships between anxiety and postoperative symptoms using non–linear methods is important.

The present study assessed the relationship between preoperative dental anxiety and postoperative symptoms after lower third molar (LM3) extraction surgery.

Methods

Study design and patients

Between May 2019 and June 2020, we performed a hospital-based prospective cohort study of patients who underwent LM3 extraction surgery at the Department of Oral and Maxillofacial Surgery, Stomatology Hospital of Tianjin Medical University. The inclusion criteria for this study were adult patients who had complete mandibular permanent dentition between 18 and 60 years of age and underwent LM3 extraction surgery under local anesthesia. The exclusion criteria were as follows: individuals aged 17 years or younger, no need for mucosal incision or high-speed turbine for extraction, inability to tolerate the procedure, presence of current pain, edema, trismus, and infection, poor compliance to postoperative care instructions and those who had previously undergone this surgery. Ethical approval was obtained from the Ethics Committee of the Stomatology Hospital of Tianjin Medical University (Tianjin, China) (No: TMUhMEC20210508). This study adhered to the ethical guidelines of the Declaration of Helsinki. Written informed consent was obtained from all patients.

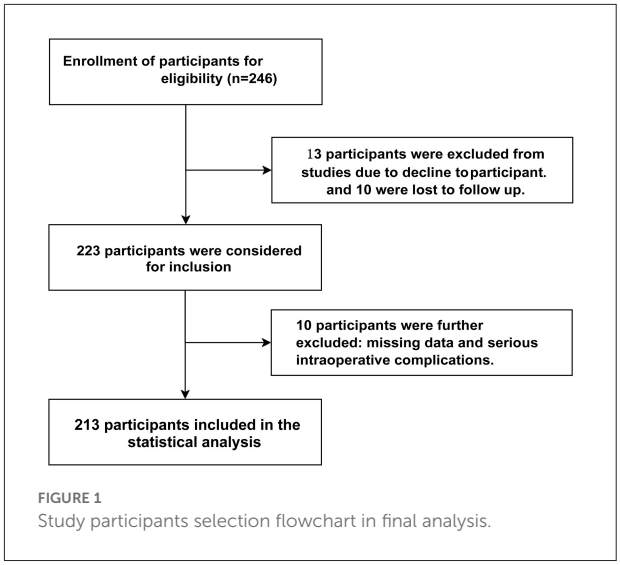


FIGURE 1
Study participants selection flowchart in final analysis.

Evaluation of dental anxiety

Preoperative dental anxiety (exposure) was measured using a self-report questionnaire (Dental Anxiety Scale for Third Molar Surgery, DAS-TMS). DAS-TMS was developed specifically for mandibular third molar extraction and is based on the Chinese version of the Modified Dental Anxiety Scale (22, 23). There were four questions on this scale. Each question was answered by a single choice among five options, representing a score of 1 to 5. The scale had a total score of 4–20. Dental anxiety levels were classified by grouping linear variables on a scale of 4–5 as No anxiety, 6–10 as Some unease, 11–15 as Anxious, and 16–20 as Very anxious. The No anxiety (4–5 points) group was defined as the reference group. We evaluated dental anxiety while sitting in a dental chair, ready for local anesthesia.

The following reliability ($n = 213$) and validity ($n = 30$) of the DAS-TMS were assessed in a randomly selected sample

TABLE 1 Baseline characteristics of the study population for total samples and subgroups according to categories of DAS-TMS*.

Variables	*N*	DAS-TMS				*P*
		No anxiety (4–5 points)	Some unease (6–10 points)	Anxious (11–15 points)	Very anxious (16–20 points)	
Sample size (%)	213	16 (7.5)	100 (46.9)	66 (31.0%)	31(14.6%)	
Postoperative symptoms score > median# No (%)	106	3 (2.8)	48 (45.3)	37 (34.9)	18(17.0)	
Gender, No. (%)						0.019
Male	76	10 (62.5)	36 (36.0)	16 (24.2)	14 (45.2)	
Female	137	6 (37.5)	64 (64.0)	50 (75.8)	17 (54.8)	
Age, median (IQR)		27.5 (22, 31.5)	29 (22, 36)	27 (23,33)	30 (23,45)	0.15
Winter, No. (%)						0.99
Vertical	54	5 (31.3)	24 (24.0)	15 (22.7)	10 (32.3)	
Mesioangular	95	8 (50.0)	45 (45.0)	31 (47.0)	11 (35.5)	
Horizontal	55	3 (18.8)	26 (26.0)	17 (25.8)	9 (29.0)	
Distoangular	3	0 (0.0)	2 (2.0)	1 (1.5)	0 (0.0)	
Inverted	6	0 (0.0)	3 (3.0)	2 (3.0)	1 (3.2)	
PG-ramus, No. (%)						0.3
I	109	8 (50.0)	55 (55.0)	34 (51.5)	12 (38.7)	
II	71	6 (37.5)	33 (33.0)	23 (34.8)	9 (29.0)	
III	33	2 (12.5)	12 (12.0)	9 (13.6)	10 (32.3)	
PG-class, No. (%)						0.76
A	109	7 (43.8)	53 (53.0)	30 (45.5)	19 (61.3)	
B	86	8 (50.0)	38 (38.0)	29 (43.9)	11 (35.5)	
C	18	1 (6.3)	9 (9.0)	7 (10.6)	1 (3.2)	
Operation time, median (IQR)		17.5 (11, 27.5)	20 (14, 29)	20 (15, 30)	25 (20, 30)	0.22

*Continuous variables were verbalized as the median (interquartile range), while categorical variables were verbalized as absolute frequencies, n (%). Continuous and categorical data were compared by using the Kruskal–Wallis test and exact fisher Chi-squared test, respectively.
#Total postoperative symptoms score was classified patients into two groups (0 = low-risk group and 1 = high-risk group) by median (20.89). A higher PoSS score reflected more severe symptoms.

from the study population: (a) internal consistency, (b) temporal stability, and (c) criterion-related validity (i.e., association with the Index of Dental Anxiety and Fear, IDAF-4C) (24). (d) discrimination validity, and (e) the construct validity from a confirmatory factor analysis (CFA).

Outcome measurement

The total severity of postoperative symptoms was used to gauge the study outcome. Postoperative symptoms were measured with the Postoperative Symptom Severity Scale (13, 15, 25). The full postoperative symptom scale was first proposed by Ruta in 2000, which contains a 7-item subscale: eating, speech, sensation, appearance, pain, sickness, and interference with daily activities (15). The total number of postoperative symptoms was the sum of the subscales. The postoperative symptom score was calculated from the self-reported questionnaire items as follows: Full Postoperative symptom score = eating scores + speech scores + sensation scores + appearance scores + pain scores + sickness scores + interference scores (15). Patients were asked to record postoperative symptoms observed during the first seven days immediately following surgery. Suture removal was done on the seventh day after extraction surgery. For the analysis, patients were classified into two groups (0 = low-risk group and 1 = high-risk group) based on the median total postoperative symptom score (20.89). A higher postoperative symptom score reflects more severe symptoms. The same investigator conducted the follow-up for all patients. Follow-up for the first postoperative week *via* interview was done during suture removal.

TABLE 2 Adjusted association for the categories of the dental anxiety with the severity of postoperative symptom[a].

DAS-TMS		Model 1[d]	Model 2[e]	Model 3[f]
	N	OR (95%CI)	OR (95%CI)	OR (95%CI)
4 ~ 5(no anxiety)	16	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)
6 ~ 10(some unease)	100	4.00 (1.07, 14.90)[c]	4.20 (1.06, 16.65)	3.63 (0.90, 14.68)
11 ~ 15(anxious)	66	5.53 (1.44, 21.25)	6.00 (1.46, 24.61)	5.29 (1.25, 22.33)
16 ~ 20(very anxious)	31	6.00 (1.42, 25.42)	5.76 (1.26, 26.33)	4.75 (1.02, 22.18)
P for trend[b]	0.019	0.031	0.058	

[a] Multivariate logistic regression analysis was used to sequentially adjusted for covariates.
OR, odds ratio; CI, confidence interval.
[b]P for trend: P for linear trend was calculated by modeling the median of the dental anxiety for each quintile as a continuous variable.
[c]Continues variables.
[d]Crude model.
[e]Adjusted for impaction status (Pell-Gregory's classification, Pell-Gregory's occlusion, Winter classification) and operation time. Adjusted odds ratios (95% confidence interval) (all such values).
[f] Additionally adjusted for gender, age.

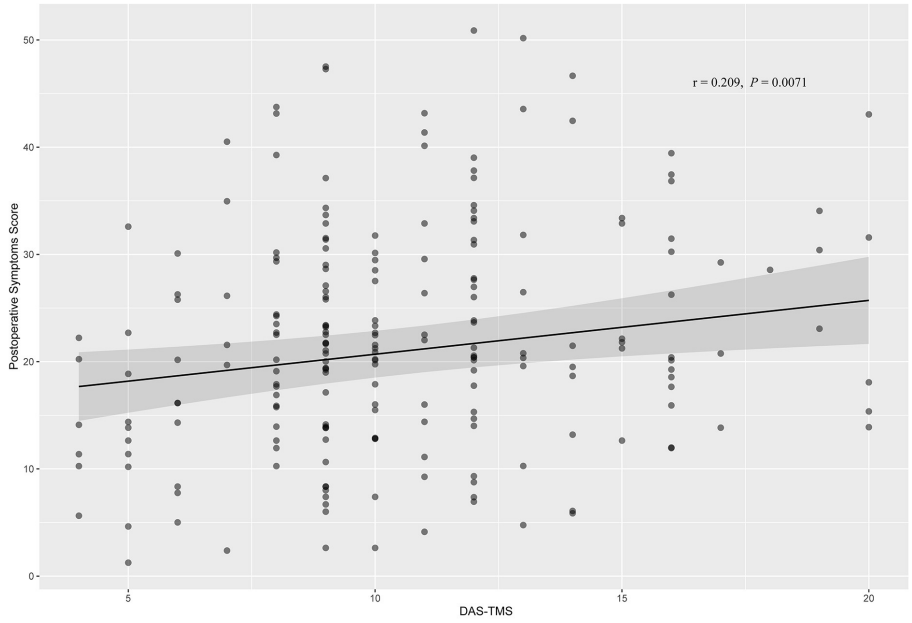


FIGURE 2
Correlation between DAS-TMS and postoperative symptoms by Pearson's test. DAS-TMS showed slightly positive correlation with postoperative symptoms ($r = 0.209$, $P = 0.0071$).

Control of confounders

Standardized tables were used to collect confounding variables for each operation. Demographic variables included sex (male/female) and age. Radiographic variables were specified using Winter's classification (26), the Pell-Gregory ramus classification, and the Pell-Gregory occlusal position (27). The operation time was also included, which was defined as the interval between the first incision and placement of the last suture.

Panoramic films were taken before surgery to evaluate and classify LM3 radiologic variables (Winter classification, Pell-Gregory ramus classification, and Pell-Gregory occlusal position). This classification method is based on the area covered by the leading edge of the mandibular ascending ramus to the teeth (Class I-III) and the depth of impaction relative to the adjacent teeth (Positions A, B, or C) (26–28).

All surgical procedures were performed in the same surgical unit. Local anesthesia was administered with 2% lidocaine and

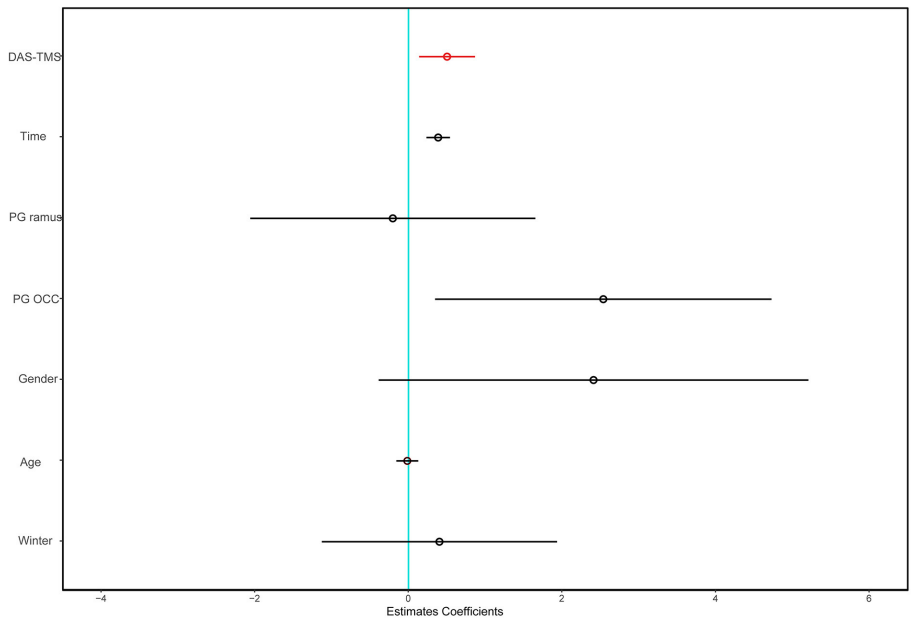


FIGURE 3
The forest plot for DAS-TMS with postoperative symptoms as a continuous variable.

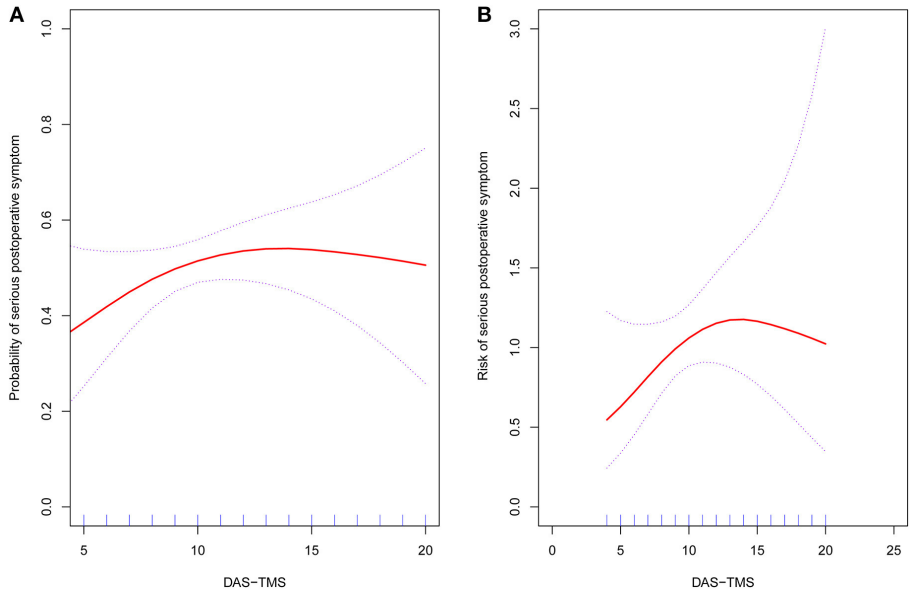


FIGURE 4
The relationship between DAS-TMS and the risk of severity of postoperative symptoms following LM3 surgery. A non−linear relationship between the DAS-TMS and risk of severity of postoperative symptoms was observed after adjusting for impaction status (Pell-Gregory's classification, Pell-Gregory's occlusion, Winter classification), operation time, gender, and age. **(A)** Probability of serious postoperative symptom; **(B)** Risk of serious postoperative symptom.

4% articaine hydrochloride under the same conditions. A full-layer mucoperiosteal flap was elevated, and either of the two techniques was employed: cases that used a triangular flap and cases that did not need a triangular flap. After determining the necessity and extent of bone removal, bone was removed from the occlusal surface of the teeth using a high-speed turbine with sufficient speed and torque. A tungsten steel crack needle drill (NSK Ltd.) was used to section the tooth (29). Machines

TABLE 3 Threshold effect analysis of DAS-TMS on the severity of postoperative symptoms using piecewise linear regression[a].

Inflection point of DAS-TMS	Odds ratio (95%CI)	*P* value
≤ 7 point	1.94 (1.12-3.74)	0.012
>7point	0.98 (0.88, 1.08)	0.756

[a]Adjusted for impaction status (Pell-Gregory's classification, Pell-Gregory's occlusion, Winter classification), operation time, gender, and age.

used during the procedure were obtained from Japan (NSK Ltd., Tokyo, Japan). Patients were given routine medication and wound dressing guidance immediately after surgery. Antibiotics were administered for 3 days after surgery.

Data management and statistical analysis

All statistical analyses were performed using R version 4.1.0, an open-source language for statistical calculations (R Foundation for Statistical Computing, Vienna, Austria). Continuous variables are expressed as median (interquartile range), while categorical variables are expressed as absolute frequencies [*n* (%)]. Continuous and categorical data were compared using the Kruskal–Wallis test and Fisher's exact test, respectively. Postoperative symptoms (binomial) were used as the dependent variable, and DAS-TMS was used as the independent variable for logistic regression. We built three multivariable logistic models (model 1, model 2, and model 3) to determine the association between preoperative dental anxiety and the severity of postoperative symptoms. Adjustments were not made in the crude model. In model 2, adjustments were made for impaction status (Pell-Gregory ramus classification, Pell-Gregory occlusal position, Winter classification) (categorical variables) and operation time (continuous variables, minutes). In model 3, additional adjustments were made for sex (categorical variables) and age (continuous variables).

We also performed secondary analyses with postoperative symptoms as continuous variables in the multivariable linear regression model. We then explored the relationship between DAS-TMS and postoperative symptoms following LM3 extraction surgery using a smoothing plot with an adjustment for potential confounders. We further applied a two-piecewise regression model to examine the threshold effect of DAS-TMS. A trial method was used to determine the threshold level of DAS-TMS at which the relationship between DAS-TMS and postoperative symptoms began to change and became notable. The trial inflection point was moved along a predefined interval, and the inflection point that gave the maximum model likelihood was detected. Differences were considered statistically significant at a two-sided *P* value of 0.05.

Results

A total of 213 patients were enrolled in this study (Figure 1). The baseline patient characteristics are shown in Table 1. The proportion of dental anxiety was as follows: No anxiety, 7.5%; Some unease, 46.9%; Anxious, 31.0%; Very anxious, 14.6%; Overall, the mean (SD) dental anxiety score was 10.56 (3.57), the median (IQR) age was 27.5 (22.0–31.5) years, and 35.7% (76 of 213) were males. The mean (SD) dental anxiety score was 10.22 (3.94) in males and 10.75 (3.34) in females.

The DAS-TMS revealed good internal consistency (Cronbach's $\alpha = 0.905$) and temporal stability ($\rho = 0.67$; $p < 0.001$). The score was significantly correlated with the IDAF-4C score ($\rho = 0.63$, $p < 0.001$), supporting good criterion-related and discrimination validity. For construct validity, the CFA revealed that the data from the DAS-TMS fit well with the two-factor model ($\chi^2 = 0.654$, $P = 0.419$, with a root mean square error of approximation = 0, comparative fit index = 1.001, goodness of fit index = 0.998, normed fit index = 0.999).

The prevalence of serious postoperative symptoms across the categories of DAS-TMS scores was 2.8% for 4–5 points, 45.3% for 6–10 points, 34.9% for 11–15 points, and 17.0% for 16–20 points. The distribution of DAS-TMS based on sex was statistically significant ($P = 0.02$). Age, Winter classification, Pell-Gregory occlusal position, Pell-Gregory ramus classification, and operation time were not significantly different across the categories of DAS-TMS (all *P* values > 0.05) (Table 1).

As shown in Table 2, a higher DAS-TMS level was associated with a higher incidence of postoperative symptoms before multivariate adjustment (Table 2). The multivariate-adjusted ORs (95% CIs) for postoperative symptom severity across categories of DAS-TMS were 1, 3.63 (0.90, 14.68), 5.29 (1.25, 22.33), and 4.75 (1.02, 22.18). The results of the secondary analysis did not significantly change the estimated associations (Figure 2). The forest plot for DAS-TMS with postoperative symptoms as a continuous variable is shown in Figure 3.

After adjusting for these possible factors related to postoperative symptoms, a non–linear relationship between DAS-TMS and postoperative symptoms was observed (Figure 4). The risk of serious postoperative symptoms increased with the dental anxiety level up to 7 points (adjusted OR 1.94, 95% CI 1.12–3.74; $P = 0.012$). When the dental anxiety level exceeded 7 points, the level of DAS-TMS was not associated with the risk of serious

postoperative symptoms (OR 0.98, 95% CI 0.88–1.08; $P = 0.756$) (Table 3).

Discussion

Dental anxiety is a common problem in patients undergoing third molar extraction (30). Dental anxiety is a significant problem for both patients and dental professionals. Information provided to the patient and dentist concerning dental anxiety is important and based on scientific evidence. Our findings indicate that dental anxiety is associated with postoperative symptoms during third molar extraction surgery. Previous studies indicated that dental anxiety is associated with surgical difficulties and postoperative pain (5, 6, 19). Management of postoperative complications is important for faster recovery. Identifying the relationship between preoperative dental anxiety and associated postoperative symptoms can help minimize and prevent postoperative complications.

Patients with high dental anxiety often require longer operation times (19). Studies have also reported that operation time is related to postoperative analgesia and the severity of postoperative complications (13). The "Very anxious" category of dental anxiety was associated with a high incidence of postoperative symptoms (OR 4.75, 95% CI 1.02–22.18) after multivariate adjustments. This also supports the conclusion that preoperative anxiety is related to postoperative symptoms. The effect of dental anxiety on the risk of serious postoperative symptoms weakened after additional adjustment, indicating that these confounders may be associated with serious postoperative symptoms.

Our results may be due to inflammatory reactions. Preoperative dental anxiety can lead to the change of prostaglandin E2 (PGE2) (31, 32). PGE2 is thought to enhance inflammation by causing vasodilation and increasing the local blood flow (33). An increase in blood PGE2 concentration caused by preoperative anxiety may affect the severity of postoperative complications. Additionally, preoperative anxiety can significantly change the release of Serotonin (5-HT) (34), which is manifested by the increased secretion of 5-HT by platelets, mast cells, and chromaffin cells. Among them, 5-HT3 can directly excite nociceptors or sensitize them through the internal messenger system. 5-HT2A acts on 5-HT2A receptors at the terminals of primary afferent fibers, resulting in aggravation of pain and edema.

In our research center, 54.4% of the patients had a DAS-TMS score of 10 or less. Our research showed that preoperative anxiety is related to the severity of postoperative symptoms. More importantly, the data further show that the risk of severe postoperative symptoms increases with an increase in DAS-TMS levels of up to 7 points. It also shows that early intervention for preoperative dental anxiety is significant in preventing the occurrence of severe postoperative symptoms. Whether the use of preoperative anxiolytic drugs affects postoperative symptoms merits further exploration.

Limitations

This study has several limitations. First, the sample was selected from only one local hospital. The patients included in this study were relatively young, which may have caused a selection bias. The representativeness of the sample might be limited, and our results may have poor generalizability. Second, the exposure and outcome variables were collected through self–completed questionnaires, which may not reflect the real situation. Third, despite controlling and adjusting many confounders, the existence of residual confounding factors may have affected the results.

Conclusion

Our findings suggest that preoperative dental anxiety is associated with a risk of serious postoperative symptoms following LM3 extraction surgery. The degree of dental anxiety in patients before LM3 extraction surgery should be of concern to clinicians.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of Stomatology Hospital of Tianjin Medical University (Tianjin, China) (No: TMUhMEC20210508). The patients/participants provided their written informed consent to participate in this study.

Author contributions

FQ and DZ contributed to the study concept and design. Material preparation, data collection and analysis was performed by FQ, MZ, and TZ. The first draft of the manuscript was written by FQ. All authors contributed to the article, commented on

previous versions of the manuscript and approved the submitted version.

Funding

This work was supported by the Science and Technology Project of Tianjin Health Committee (Grant Number: ZC20134) and the Joint Research Fund of Tianjin Medical University-DeepWise Health care (Grant Number: 1KQ026).

Acknowledgments

We would like to thank Editage (www.editage.cn) for English language editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Sirin Y, Humphris G, Sencan S, Firat D. What is the most fearful intervention in ambulatory oral surgery? Analysis of an outpatient clinic. *Int J Oral Maxillofac Surg.* (2012) 41:1284–90. doi: 10.1016/j.ijom.2012.06.013

2. Oosterink FM, De Jongh A, Aartman IH. What are people afraid of during dental treatment? Anxiety-provoking capacity of 67 stimuli characteristic of the dental setting. *Eur J Oral Sci.* (2008) 116:44–51. doi: 10.1111/j.1600-0722.2007.00500.x

3. Zinke A, Hannig C, Berth H. Comparing oral health in patients with different levels of dental anxiety. *Head Face Med.* (2018) 14:1–5. doi: 10.1186/s13005-018-0182-4

4. Coxon JD, Hosey MT, Newton JT. How does dental anxiety affect the oral health of adolescents? A regression analysis of the child dental health survey 2013. *Br Dent J.* (2019) 227:823–8. doi: 10.1038/s41415-019-0895-1

5. Potter CM, Kinner DG, Tellez M, Ismail AI, Heimberg RG. Clinical implications of panic symptoms in dental phobia. *J Anxiety Disord.* (2014) 28:724–30. doi: 10.1016/j.janxdis.2014.07.013

6. Heaton LJ. Self-reported dental anxiety is associated with both state anxiety and dental procedure-related pain. *J Evid Based Dent Pract.* (2017) 17:45–7. doi: 10.1016/j.jebdp.2017.01.007

7. Lin CS, Wu SY Yi CA. Association between anxiety and pain in dental treatment: a systematic review and meta-analysis. *J Dent Res.* (2017) 96:153–62. doi: 10.1177/0022034516678168

8. White RP, Shugars DA, Shafer DM, Laskin DM, Buckley MJ, Phillips C. Recovery after third molar surgery: clinical and health-related quality of life outcomes. *Int J Oral Maxillofac Surg.* (2003) 61:535–44. doi: 10.1053/joms.2003.50106

9. Duarte-Rodrigues L, Miranda EFP, Souza TO, de Paiva HN, Falci SGM, Galvao EL. Third molar removal and its impact on quality of life: systematic review and meta-analysis. *Qual Life Res.* (201) 27:2477–89. doi: 10.1007/s11136-018-1889-1

10. Bradshaw S, Faulk J, Blakey GH, Phillips C, Phero JA, White RP. Quality of life outcomes after third molar removal in subjects with minor symptoms of pericoronitis. *J Oral Maxillofac Surg.* (2012) 70:2494–500. doi: 10.1016/j.joms.2012.05.013

11. Conrad SM, Blakey GH, Shugars DA, Marciani RD, Phillips C, White RP. Patients' perception of recovery after third molar surgery. *J Oral Maxillofac Surg.* (1999) 57:1288–94. doi: 10.1016/S0278-2391(99)90861-3

12. Zheng X, Zhao J, Wang Z, Jia B, Zhang Z, Guo J, et al. Postoperative online follow-up improves the quality of life of patients who undergo extraction of impacted madibular third molars: a randomized controlled trial. *Clin Oral Investig.* (2021) 25:993–9. doi: 10.1007/s00784-020-03388-0

13. Qiao F, Li L, Zhang J, Dong R, Sun J. Operation time is independent associated with serious postoperative symptom in patients with mandibular third molar removal. *Ann Palliat Med.* (20121) 10:4080–9. doi: 10.21037/apm-20-2340

14. Qiao F, Huang X, Li B, Dong R, Huang X, Sun J, et al. validated model to predict postoperative symptom severity after mandibular third molar removal. *J Oral Maxillofac Surg.* (2020) 78:893–901. doi: 10.1016/j.joms.2020.02.007

15. Ruta DA, Bissias E, Ogston S, Ogden GR. Assessing health outcomes after extraction of third molars: the postoperative symptom severity (PoSSe) scale. *Br J Oral Maxillofac Surg.* (2000) 38:480–7. doi: 10.1054/bjom.2000.0339

16. Aznar-Arasa L, Figueiredo R, Valmaseda-Castellon E, Gay-Escoda C. Patient anxiety and surgical difficulty in impacted lower third molar extractions: a prospective cohort study. *Int J Oral Maxillofac Surg.* (2014) 43:1131–6. doi: 10.1016/j.ijom.2014.04.005

17. Rafetto LK. Managing impacted third molars. *Oral Maxillofac Surg Clin North Am.* (2015) 27:363–71. doi: 10.1016/j.coms.2015.04.004

18. Miloro M, Ghali G, Larsen PE, Waite PD (2004). *Peterson's Principles of Oral and Maxillofacial Surgery*. USA: Springer.

19. Lago-Mendez L, Diniz-Freitas M, Senra-Rivera C, Seoane-Pesqueira G, Gandara-Rey JM, Garcia-Garcia A. Postoperative recovery after removal of a lower third molar: role of trait and dental anxiety. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* (2009) 108:855–60. doi: 10.1016/j.tripleo.2009.07.021

20. de Jongh A, van Wijk AJ, Lindeboom JA. Psychological impact of third molar surgery: a 1-month prospective study. *J Oral Maxillofac Surg.* (2011) 69:59–65. doi: 10.1016/j.joms.2010.05.073

21. Onwuka CI, Udeabor SE, Al-Hunaif AM, Al-Shehri WAK, Al-Sahman LA. Does preoperative dental anxiety play a role in postoperative pain perception after third molar surgery? *Ann Afr Med.* (2020) 19:269. doi: 10.4103/aam.aam_68_19

22. Lin CS, Lee CY, Wu SY, Chen LL, Lee KT, Wang MC, et al. Translation and validation of modified dental anxiety scale based on adult Taiwan population. *BMC Oral Health.* (2021) 21:647. doi: 10.1186/s12903-021-02017-w

23. Yuan S, Freeman R, Lahti S, Lloyd-Williams F, Humphris G. Some psychometric properties of the Chinese version of the modified dental anxiety scale with cross validation. *Health Qual Life Outcomes.* (2008) 6:22. doi: 10.1186/1477-7525-6-22

24. Armfield JM. Development and psychometric evaluation of the index of dental anxiety and fear (IDAF-4C+). *Psychol Assess.* (2010) 22:279–87. doi: 10.1037/a0018678

25. Grossi GB, Maiorana C, Garramone RA, Borgonovo A, Creminelli L, Santoro F. Assessing postoperative discomfort after third molar surgery: a prospective study. *J Oral Maxillofac Surg.* (2007) 65:901–17. doi: 10.1016/j.joms.2005.12.046

26. Winter GB. *Principles of Exodontia as Applied to the Impacted Mandibular Third Molar*. St Louis, MO: American Medical Book Company. (1926).

27. Pell GJ. Report on a ten year study of a tooth division technique for removal of impacted teeth. *Am J Orthod Oral Surg.* (1942) 28:660. doi: 10.1016/S0096-6347(42)90021-8

28. Pederson GW. *Oral Surgery*. (1988). Philadelphia, PA: WB Saunders.

29. Farish SE, Bouloux GF. General technique of third molar removal. *Oral Maxillofac Surg Clin North Am.* (2007) 19:23–43. doi: 10.1016/j.coms.2006.11.012

30. Earl P. Patients' anxieties with third molar surgery. *Br J Oral Maxillofac Surg.* (1994) 32:293–7. doi: 10.1016/0266-4356(94)90049-3

31. Winston JH, Sarna SK. Enhanced sympathetic nerve activity induced by neonatal colon inflammation induces gastric hypersensitivity and anxiety-like behavior in adult rats. *Am J Physiol-Gastrointest Liver Physiol.* (2016) 311:G32–9. doi: 10.1152/ajpgi.00067.2016

32. Moons WG, Shields GS. Anxiety, not anger, induces inflammatory activity: an avoidance/approach model of immune system activation. *Emotion.* (2015) 15:463. doi: 10.1037/emo0000055

33. Kawahara K, Hohjoh H, Inazumi T, Tsuchiya S, Sugimoto Y. Prostaglandin E2-induced inflammation: rRelevance of prostaglandin E receptors. *Biochim Biophys Acta.* (2015) 1851:414–21. doi: 10.1016/j.bbalip.2014.07.008

34. Akimova E, Lanzenberger R, Kasper S. The serotonin-1A receptor in anxiety disorders. *Biol Psychiatry.* (2009) 66:627–35. doi: 10.1016/j.biopsych.2009.03.012

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 08 September 2022
DOI 10.3389/fpsyt.2022.888234

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University, Germany

REVIEWED BY
Gaia Sampogna,
University of Campania L. Vanvitelli, Italy
Delfina Janiri,
Sapienza University of Rome, Italy

*CORRESPONDENCE
Ozélia Sousa Santos
ozeliasousa@ufpa.br
Fernanda Nogueira Valentin
fer_valentin@yahoo.com.br

†These authors have contributed equally to this work and share last authorship

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 02 March 2022
ACCEPTED 05 August 2022
PUBLISHED 08 September 2022

CITATION
Sousa ILMd, Silveira R, Takito MY, Pereira AL, Lucianelli-Júnior D, Carmona GS, Viegas APdV, Teixeira FB, Santos OS and Valentin FN (2022) The impact of the social isolation in elderly Brazilian mental health (anxiety and depression) during the COVID-19 pandemic.
Front. Psychiatry 13:888234.
doi: 10.3389/fpsyt.2022.888234

The impact of the social isolation in elderly Brazilian mental health (anxiety and depression) during the COVID-19 pandemic

Isabella Louise Morais de Sousa[1], Rodrigo Silveira[2], Mônica Yuri Takito[2], Adenilson Leão Pereira[1,3], Dalberto Lucianelli-Júnior[1], Giselle Sousa Carmona[4], Ana Paula do Vale Viegas[1], Francisco Bruno Teixeira[5], Ozélia Sousa Santos[1*†] and Fernanda Nogueira Valentin[1,3*†]

[1]Faculty of Medicine, Federal University of Pará, Altamira, Brazil, [2]Campus 'University City Armando de Salles Oliveira (CUASO)', University of São Paulo, São Paulo, Brazil, [3]Graduate Program in Biodiversity and Conservation (PPGBC), Altamira, Brazil, [4]Faculty of Medicine, Federal University of Piauí, Teresina, Brazil, [5]Faculty of Medicine, State University of Pará, Belém, Brazil

The impact of social isolation in the pandemic context on elderly Brazilian mental health is little known, especially about the occurrence of depressive symptoms. In this study, we evaluated elderly people undergoing social isolation in order to identify factors associated with depression and which of these are more important to characterize elderly Brazilians with depression. In a cross-sectional, exploratory, and analytical study of a quantitative nature, the mental profile of elderly individuals subjected to social isolation during the COVID-19 pandemic period was used. A total of 450 participants was divided into normal and depressive groups, and a form covering sociodemographic data, opinions/perceptions about the pandemic, and a Reduced Geriatric Depression Scale was used to assess participants' mental health. To assess the statistical significance between the variables, chi-square test was applied, considering the p-value <0.05. The effect size was analyzed to identify the magnitude of the difference between groups. To identify the most important characteristics to define the groups Multilayer Perceptron algorithm were applied. We found that elderly people with a depressive profile are (in Multilayer Perceptron rank order) (1) showing signs of anxiety during the COVID-19 pandemic, (2) of low education, (3) being divorced, (4) having more than one mental disorder, (5) reading, watching, or listening to information about COVID-19, and (6) being previously diagnosed with depression. In conclusion, elderly Brazilians in social isolation tend to develop depressive disorders during quarantine. Thus, we can consider that the pandemic requires effective and safe gerontological care and monitoring, especially with regard to mental health.

KEYWORDS

COVID-19, social isolation, elderly, mental health, anxiety, depression

Introduction

The emergence and rapid increase in the number of cases of COVID-19, an infectious disease caused by the new coronavirus, which in most cases can lead the patient to the severe acute respiratory syndrome, presents complex challenges for health, economy, and society. COVID-19 is currently a public health emergency of international concern, as declared on 30 January 2020 by the World Health Organization. In early July of 2022, there were more than 552 million confirmed cases of COVID-19 worldwide and more than 6.34 million deaths (1).

The first confirmed case of COVID-19 in Brazil was announced on 26 February 2020. Currently, the number of cases exceeds 28 million and more than 670,000 victims, making Brazil the third country with more cases and is the second deaths by COVID-19 in the world (2).

The COVID-19 pandemic has been compared to catastrophic events such as earthquakes, tsunamis, conflicts, and wars. However, unlike these cases, the pandemic was and is still something unusual and obscure for world society, because until some time ago, it was not known what was ahead, and the possibility of contagion by the virus was everywhere and is still a threat (3). In addition, the excess of information transmitted and still may generate panic, favoring situations of stress and fear. Studies show that these factors can trigger traumatic stress, which may manifest itself in the main models of post-traumatic stress disorder (4, 5), which may have even more catastrophic impacts on vulnerable groups such as the elderly people (6, 7).

In the beginning of the confrontation of the COVID-19 pandemic, in 2020, Brazil adopted many public health measures, such as quarantine and mandatory social isolation, suspensions from school and non-essential services, in order to mitigate the risks and impact of the disease on the population. A study carried out in Hong Kong showed that sudden changes in daily life are risk factors that can substantially affect mental health, and this fact can be brought to the Brazilian context (8).

The elderly are more vulnerable to COVID-19 because they have a higher risk of developing the most severe form of the disease, especially those with preexisting comorbidities, such as heart, hypertension, diabetes, kidney, lung, cancer, and immunosuppression diseases (9). In Brazil, the mortality rate in 2020 among people with aged $\geq$80 years was higher (14.8% died), when compared to the elderly aged 70–79 years (8% died) and 60–69 years (8.8% died), in other words, a rate of 3.82 times higher than the general average, reinforcing the concerns regarding the elderly population (10). However, after the start of vaccination for the elderly in January 2021, these numbers have been reduced (11). Orellana and collaborators in 2022 (12) observed changes in the pattern of hospitalizations and deaths from COVID-19 after substantial vaccination of the elderly in Manaus, Amazonas, and Brazil. According to him, there was an overall reduction of approximately 62% in hospitalization and death rates, especially in the elderly aged 60–69 years.

Social distancing and isolation are among the recommended guidelines for the safety of the elderly during the pandemic. However, social isolation is a major danger to the health and wellbeing of the elderly as it is associated with an increased mortality risk and is linked to worsening mental health (13). The incidence and prevalence of the depressive disorder in the elderly population is high globally, and although it affects both sexes, the incidence is higher in women (14). Recently, Santini, Jose (15), observed that social disconnection exposes the elderly to a high risk of depression. In addition, it is believed that the health risks associated with the social isolation and loneliness are equivalent to the prejudicial effects caused by smoking and obesity (16).

The causes of depression can be genetic, brain biochemistry, or vital events. Events that cause stress and anxiety, also called vital events, are mostly triggering factors for depressive episodes, especially in those who already have a genetic predisposition to the development of the disorder. The imbalance of neurotransmitters such as serotonin, dopamine, and noradrenaline responsible for controlling appetite, mood, and motor activity are also closely associated with depression (17).

The situations of daily life trigger different reactions in individuals, among which depressive symptoms are present. In these situations, individuals demonstrate general or non-specific responses of a physiological and psychological nature of the body to a stressor or external and internal threats (18, 19).

The causes and symptoms that trigger the depressive disorder are well characterized; however, in elderly individuals, these symptoms are more difficult to diagnose and, consequently, to treat. Therefore, the main difficulty in the treatment of this clinical condition is the correct diagnosis, which is partly associated with the fact that many elderly people do not accept their depressive clinical condition and do not seek adequate psychiatric treatment (20). In this scenario, the context of the COVID-19 global pandemic can make this situation more aggravating, since the fear of the unknown can lead to depression, and social isolation measures limit people's daily activities, especially of the elderly (21). Therefore, due to the pandemic conditions to which the elderly are being subjected, the development or worsening of depressive conditions is expected, since these disorders are closely related to social isolation, affecting physical and mental health and aggravating underlying diseases (22).

The Brazilian population has a cultural and religious plurality that is very subjective (23, 24), and it is possible that it does not behave in the same way in relation to other population groups. In this context, the use of machine learning can be useful to create robust models that can provide more accurate data for this population.

Although there are previous works based on bibliographic reviews in Brazil (25–28), and some cross-sectional studies on mental health of the elderly in the pandemic in other countries such as China, Spain, and Italy (8, 29–33), in Brazil, cross-sectional studies have not yet been found, nor combined with k-means cluster analysis (an unsupervised machine learning algorithm) that explored the association of COVID-19 impacts and physical isolation on the mental health of elderly Brazilians, especially in terms of depression levels.

In this study, we aimed to identify whether there are distinct groups in the elderly population (with and without depression). We also analyzed the main characteristics of elderly Brazilian people with and without depression in the period of social isolation and we identified which of these characteristics are more important to characterize Brazilian elderly people with depression. Thus, based on the literature cited, it is believed that elderly Brazilians may develop or worsen depressive symptoms during the COVID-19 pandemic due to social isolation.

Materials and methods

Participants

The study included 450 male and female subjects, over 60 years of age (67.2 ± 6.7 years), representing all Brazilian states. The form was in Portuguese and was available online from 26 June to 8 September 2020, through social networks and e-mail.

Data collection was performed after approval of the research project by the Ethics Committee of the Institute of Health Sciences of the Federal University of Pará (CAAE number: 32893620.8.0000.0018). All participants who agreed to participate in the research signed the Informed Consent Form.

In this study, only people who lived in Brazil at the time of data collection were included. The questionnaires were distributed mostly by e-mail to universities, institutes, and personal e-mails of project participants. In addition, another part of the participants, the application of the form, was carried out through the whatsapp application and/or telephone call. The elderly who could not answer the form alone were helped by someone close (family member, friends, or project participants) to whom the questions were dictated and the respective alternatives were answered verbally. Participants unable to answer verbally and/or provide decisions regarding the alternatives to the questions by cognitive or psychiatric disability were excluded. In addition, for all participants who filled it more than once, only the first participation was maintained, excluding the remaining.

To ensure better quality of the data obtained, a pilot study was conducted before starting the official form dissemination with a dataset of 100 participants (not counted in the sample) to evaluate the dissemination strategy, responses obtained, and the quality of the anchoring questions.

For the sample calculation, the G* Power 3.0.10 software was used to simulate all the analyses performed. The sample size was determined by the analysis that estimated the largest number of participants, being a Chi-square test with up to 6 degrees of freedom, assuming an intermediate effect size, a significance of $p < 0.05$ and a statistical power of 95%, estimating a minimum sample of $n = 232$. However, to ensure better representativeness of the Brazilian population, this minimum sample size was estimated to be increased by 90%. Thus, based on cultural plurality rooted in the great social and regional diversity in the set of 27 Brazilian states (34), the estimated minimum sample size increased by 186 (∼80%) with an additional 22 (∼10%) for possible sample loss, totaling a minimum sample size of $n = 440$.

The online form was structured with multiple choice questions and covering general demographic data such as age, gender, race, marital status, religion, having children, education, city and previous diagnosis of mental disorder. The questions on the opinions and perceptions of the elderly regarding the COVID-19 pandemic were as follows: (a) If the participants claim to know what the pandemic and COVID-19 are?; (b) What are the main ways to obtain information about the pandemic?; (c) How much time do you spend getting this information?; (d) Do you know what social isolation is?; (e) Do you agree with the imposed social isolation?; (f) How do you feel about the whole pandemic scenario?; and (g) Who are they with passing the period of social isolation?

Mental health measurements

To assess anxiety, the Brazilian version of the Geriatric Anxiety Inventory (GAI) with 20 objective questions was applied (34). The GAI is characterized by being a self-applicable instrument with dichotomous responses (agree/disagree) (35). The instrument has a cutoff score between 10/11 (non-case/case), where a score of 0–10 indicates no anxiety, 11–15 indicates mild or moderate anxiety, and 16–20 indicates severe anxiety. In this study, only the absence (score 0–10) or presence (score 11–20) of anxiety was considered.

To assess depression, the Brazilian version of the reduced Geriatric Depression Scale with 15 objective questions was applied (36). Its score ranges from 0 to 15 points, being divided into three categories. A score of 0–5 is considered normal, 6–10 mild depressive symptoms, and 11–15 severe depressive symptoms. We only considered the absence (score 0–5) or presence (score 6–15) of depression.

Bias

To avoid possible interpretation errors and potential sources of bias, a pilot study was conducted, which served to improve the form questions.

Data analysis

Continuous data were presented as the median and interquartile range, while categorical data as percentages. To analyze the significance between the proportions of the sample with and without depressive disorder, 95% confidence intervals were observed. To analyze the associations between the groups with and without depression and the different categorical variables, Pearson's chi-square test was applied. Correction by Fisher's exact test was applied when in any contingency table there was $n < 6$ in any cell. For all tests, the statistical significance adopted was p-value <0.05. In contingency tables >2 x 2 with statistical significance, adjusted residuals >2 were analyzed to identify which categories influenced the p-value <0.05.

To analyze the magnitude of differences between groups, effect sizes were observed using Φ (ϕ) in 2 x 2 tables, assuming "no effect" for $\phi < 0.10$, "small effect" for $\phi < 0.30$, "moderate effect" for $\phi < 0.50$ and "large effect" for higher values. In >2 x 2 tables, the sizes were observed by Cramer's V, whose interpretations of null, small, moderate, and large effects were performed considering the variations according to an increase in degrees of freedom (37, 38).

To assess the characteristics that most influence the classification of the participants as depressive or non-depressive, the Multilayer Perceptron algorithm was used (p-value <0.05). This supervised machine learning algorithm, through an artificial neural network, identifies non-linear patterns among different variables in a dataset and, in response, provides a prediction of some predetermined variable of interest. When executed, this learning algorithm perform through the following steps: (1) the weights are initialized; (2) the flow and analysis of information flows through the input, hidden, and output layers; (3) error rate in output layer predictions is calculated and weights are adjusted; and (4) all previous steps are repeated until the error rate becomes as low as possible (39).

Quantitative variables were rescheduled at intervals between 0 and 1. The samples were randomly divided into two datasets, where 70% of the samples were used for training the algorithm and 30% for the test. For training and optimization, Minibatch and Descending Gradient methods were selected, respectively. Because Multilayer Perceptron can give different results each time it is run due to randomization of dataset partitions for cross validation and initialization of weights, the algorithm was run three times. The trial chosen was the one with the lowest mean value of cross-entropy error ([training error + test error]/2). Therefore, the chosen attempt was the second.

The ability of the predictors to determine the artificial neural network was tested by using sensitivity analysis, combining the training and test samples. In addition, a table that shows the degree of importance of each predictor was created. Data analyses were processed using the SPSS v.23.0 software.

Results

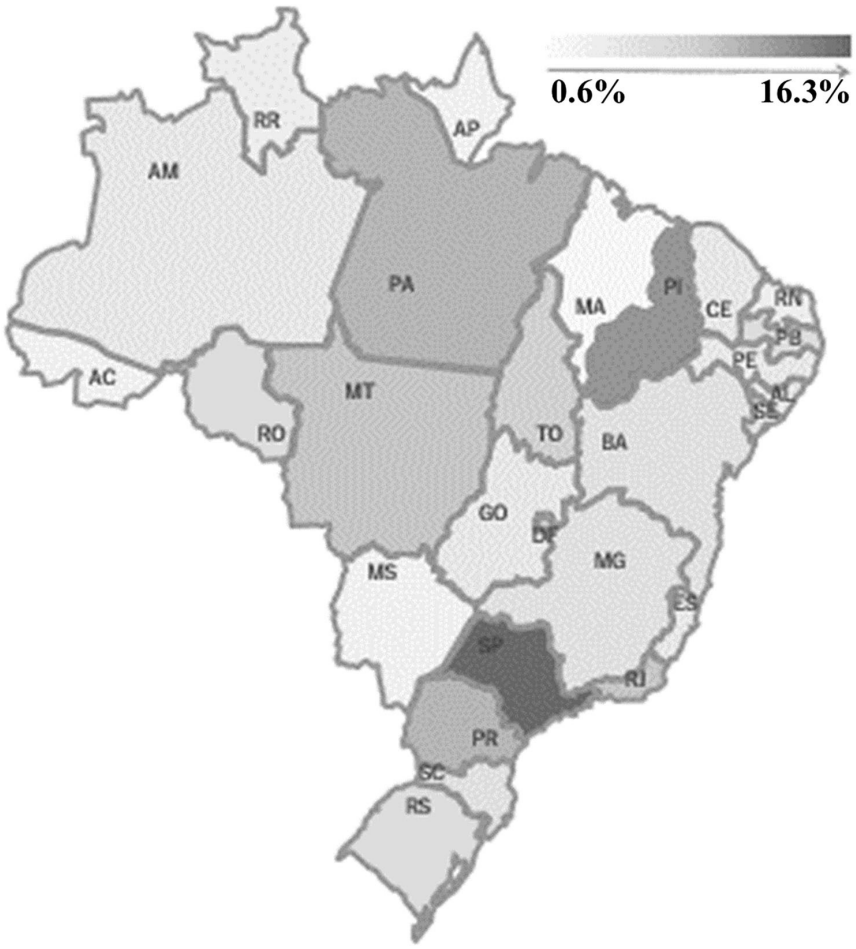
The sample distribution ($n = 450$) across Brazilian states ranged from $n = 3$ in Acre to $n = 69$ in São Paulo (Figure 1). Of the 450 subjects, 31.1% showed depressive symptoms (IC: Normal = 64.6–73.2; IC: Depressive = 26.8–35.4).

The sociodemographic characteristics between the groups are detailed in Supplementary Table S1. Individuals with depressive symptoms are characterized by having a higher proportion of women (80.7%), divorced (23.6%), and with low education (32.9%) ($p < 0.01$). Regarding the religion of the elderly without depressive symptoms, there was a higher proportion of individuals without religion (14.2%), while among the elderly with depressive symptoms, there was a higher proportion of subjects who adhere to Afro-Brazilian religions (2.1%) ($p < 0.05$). There was no statistical difference between the depressive and normal groups regarding ethnicity and whether they had children.

Table 1 shows the participants' psychological responses and perceptions about the COVID-19 pandemic. The elderly people were divided into groups with and without depression. On declaring themselves to have a mental disorder, it was observed that the elderly people who claimed to be anxious and those who already had a diagnosis of depressive disorder were present in the group with depression ($p < 0.001$). On declaring themselves to have a mental disorder, it was observed that the elderly who claimed to be anxious and those who already had a diagnosis of depressive disorder were present in the depressive group ($p < 0.001$). There is a direct relationship between the number of mental diseases and the group that has depression, while people without any mental disorder are mostly present in the non-depressive group.

Elderly with signs of severe (30.7%) and mild-to-moderate (25.0%) anxiety were predominant in depressive group ($p < 0.001$), having a large effect size ($p < 0.001$). Those who declared that they do not understand the situation that the world is going through, and who do not understand what a pandemic and COVID-19 is, most of them are present in the depressive group. Those who usually obtain information through reading, viewing, or listening to news about COVID-19 are present in the non-depressive group. There was no relevance among the sources of information used by the elderly to find out about the pandemic. As for the reason why the elderly person maintains social isolation, the elderly who declared not knowing or not understanding the reason for physical isolation predominated in the non-depressive group; this variable was the most influential ($p < 0.05$).

Of the characteristics with statistical value ($p < 0.05$), the most important to identify the groups is the presence or absence of anxiety symptoms, followed by education and civil status (Figure 2). In addition, the ranking showed that the importance



State	n
São Paulo (SP)	69
Piauí (PI)	48
Pará (PA)	34
Paraná (PR)	34
Mato Grosso (MT)	26
Rio de Janeiro (RJ)	25
Distrito Federal (DF)	19
Tocantins (TO)	19
Paraíba (PB)	15
Rio Grande do Sul (RS)	14
Rondônia (RO)	14
Bahia (BA)	12
Sergipe (SE)	12
Santa Catarina (SC)	11
Minas Gerais (MG)	11
Goiás (GO)	10
Amazonas (AM)	10
Pernambuco (PE)	9
Maranhão (MA)	9
Espírito Santo (ES)	8
Ceará (CE)	8
Alagoas (AL)	8
Rio Grande do Norte (RN)	7
Roraima (RR)	6
Amapá (AP)	5
Mato Grosso do Sul (MS)	4
Acre (AC)	3
TOTAL	**450**

FIGURE 1
Sample divided by Brazilian states.

of the other variables varies in a complex way among biological, psychological, and social factors.

Discussion

The COVID-19 pandemic was and is still considered an acute stressor for the general population (40). In addition, studies have shown that this event generated emotional deregulation that culminates in high psychological distress, triggering anxious and depressive symptoms, especially for older age groups (41, 42). Such an event contributed to a large number of people developing and exacerbating neurological disorders, which are determined by individual factors that affect the way each patient deals with a traumatic event, such as the pandemic (43).

In this study, we evaluated and identified the characteristics of Brazilian elderly people with and without depression in social isolation during the COVID-19 pandemic period, and which of them are more appropriate to characterize the elderly in depressive conditions. In general, the elderly in depressive conditions are mostly diagnosed with anxiety, have low education, and are widowed or unmarried.

We observed a predominance of women in the depressive group. This fact may be linked to the fact that women tend to be more vulnerable when subjected to stress and when developing post-traumatic symptoms, as a consequence of the intense routine required by the demands of work, child care, and daily routines (44). Our results corroborate previous studies (45, 46) that found an association between the female sex and psychological distress increasing.

Studies have observed that during social isolation there has been an increase in the number of cases of domestic violence against women in Brazil, in part, as a result of the longer time spent with couples or spouses (47, 48). The rise in this type of violence was an important

TABLE 1 Psychological responses and participants' perceptions about the COVID-19 pandemic.

	Psychological responses and participants' perceptions of the COVID-19 pandemic				
		Depression ($n = 140$)	Normal ($n = 310$)	Effect size	*P*
Diagnosed with some mental disorder	Anxiety	24.30%	8.10%	$\phi = 0.22^{+}$	<0.001***
	Depression	29.30%	9.00%	$\phi = 0.26^{+}$	<0.001***
	Bipolar affective disorder	1.40%	1.00%	$\phi = 0.20^{+}$	0.65
	ADHD	0.70%	1.30%	$\phi = 0.25^{+}$	1
	Panic syndrome	2.10%	0.60%	$\phi = 0.07^{+}$	0.18
Number of mental disorders	Any mental disorder	53.6%[a]	82.3%[a]	$v = 0.31^{++}$	<0.001***
	Has one mental disorder diagnosed	35.0%[b]	15.5%[b]		
	Has two mental disorders diagnosed	11.4%[c]	2.3%[c]		
Classification of anxiety	Shows signs of severe anxiety during the COVID-19 pandemic	30.7%	1.6%	$v = 0.54^{+++}$	<0.001***
	Shows signs of mild (leve) to moderate anxiety during the COVID-19 pandemic	25.0%	6.8%		
Understanding about COVID-19 pandemic	Declares not understand the world situation due to the COVID-19 pandemic	19.30%	9.00%	$\phi = 0.14^{+}$	<0.01**
	Declares not understand what is a pandemic and COVID-19	23.60%	11.90%	$\phi = 0.15^{+}$	<0.01**
Source of information about COVID-19	Declares usually reads. watches or listens to news related to COVID-19	87.90%	94.80%	$\phi = 0.12^{+}$	<0.01**
Information source most consulted for news about the COVID-19 pandemic	World Health Organization Guidelines	7.90%	6.80%	$v = 0.11^{++}$	0.13
	Radio and Televison	62.90%	54.20%		
	Internet and magazines	20.70%	31.60%		
	Family	8.60%	7.40%		
Reasons for social isolation reported by participants	Prevent the spread of the virus	37.90%	42.90%	$v = 0.14^{++}$	<0.05*
	He / she is in the risk group	10.70%	6.10%		
	For him / her not to be contaminated	49.30%	43.20%		
	Does not know the reason for the social isolation or did not know how to explain	2.1%[a]	7.7%[a]		

Different letters represent the categories that influenced the statistical significance ($p < 0.05$) between the groups, with the letter "a" corresponding to the highest adjusted-value residual (>2) and the subsequent letters characterizing smaller values, respectively.
$^{+}$small effect, $^{++}$moderate effect, $^{+++}$large effect.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

factor for the development of depressive symptoms in women (47, 48).

We found that marital status is also associated with depression levels. In fact, widowed or divorced elderly people have a higher risk of feeling lonely and depressed (49). The loss of the spouse can cause an increase in depressive symptoms, and the absence of a partner is among the factors that lead the elderly to a state of social and emotional loneliness, favoring the onset of depressive symptoms (50).

Regarding religious conviction, the elderly of Afro-descendant religions belonged to the depressive group, while the elderly belonging to the non-depressive group and more informed about the pandemic declared not to have a religion. Therefore, we emphasize that new studies considering religious conviction among depressed elderly people need to be conducted to better investigate, characterize and understand the impact of this variable on the mental health of elderly people.

The fact of having or not having children was not statistically significant in determining the groups with and without depression. Nóbrega et al. (51) observed that the presence of depression in elderly Brazilians was independent of the fact of having children. Oliveira et al. (52) observed that elderly people who do not live with their children have a higher risk of feeling depressed, probably due to the feeling of

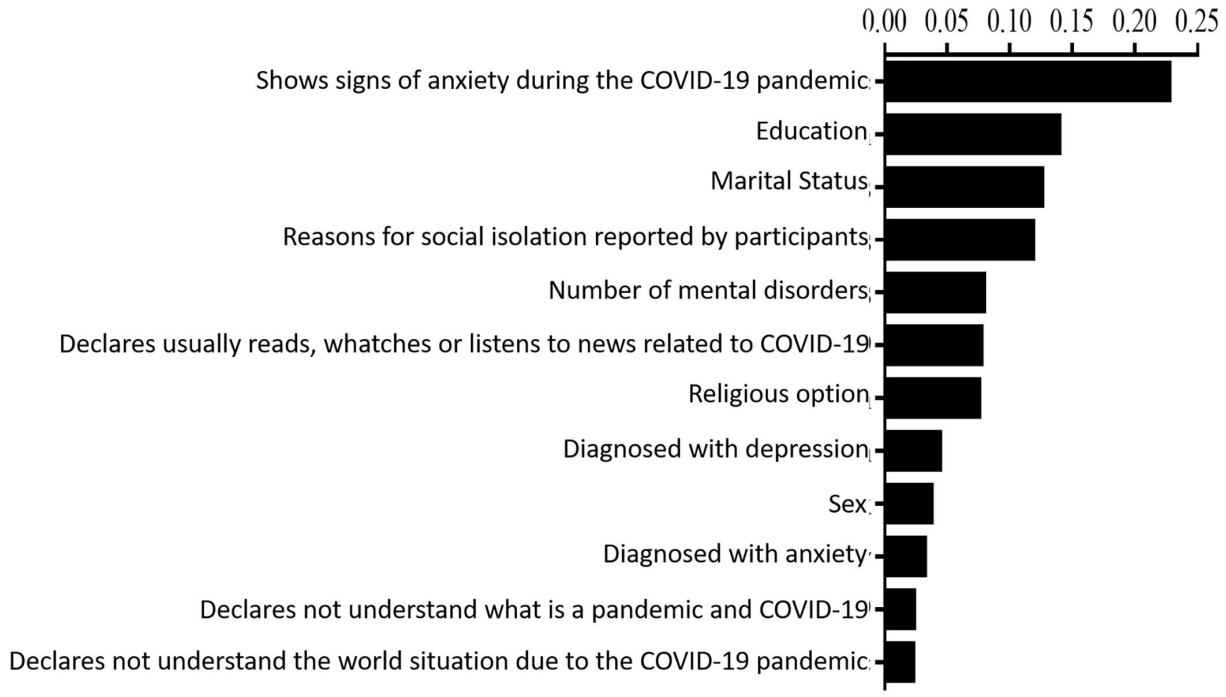


FIGURE 2
Importance of the variables for the characterization of the groups with and without depression. The figure shows the percentage importance of each variable that has statistical significance.

loneliness. We emphasize that there is no consensus whether this variable is a factor directly related to the presence of depression in elderly.

The second most important variable to characterize depressed elderly people was their low educational level. These results corroborate previous studies that report that this condition influences the onset of anxiety and depression symptoms during old age (53, 54). The educational level is directly related to the economic level and quality of life, factors that are determinant for the index of depressive symptoms (55). It is recognized that the educational level is directly related to the economic level and quality of life, factors that are determinant for the index of depressive symptoms (44). These combined characteristics provide a state of pessimism that may result in the inability to confront these situations (56). In addition, the inability to read and interpret texts combined with limited access to information can be an obstacle for the elderly to obtain a minimum level of knowledge about protective measures against the coronavirus and to update themselves on their reality. Thus, this group may develop more concerns and, consequently, become more prone to the development of depressive symptoms (56).

Regarding the fact of having depression and previous diagnosis of other mental illnesses, the most elderly people with depressive disorders claimed to have another type of psychiatric disorder, mainly anxiety. We also identified that the most influential variable in determining elderly people with depressive disorder is the previous diagnosis of anxiety, since 55.7% of the elderly reported having symptoms of anxiety during the COVID-19 pandemic. These results corroborate the results of studies carried out in other countries during the pandemic (29). Anxiety is considered a possible risk factor for the onset of depression, and the simultaneous occurrence of these two psychopathologies among the elderly is frequent (57).

Elderly people in the non-depressive group stood out in terms of obtaining information about the pandemic and COVID-19 when compared to the depressive group. We emphasize that the individual in depression may develop feelings and thoughts of pessimism, helplessness, deep sadness, apathy, lack of initiative, physical discontent, difficulty in organizing and fluidity of ideas, impaired cognitive judgment, among other symptoms (58). Thus, such factors can compromise the ability of an individual affected by depression to obtain information, especially when related to COVID-19.

Participants who declared not knowing or not understanding the reason for physical isolation were predominant in the non-depressive elderly group. This result may be a consequence of data collection since the data were collected at the beginning of the pandemic, when the rigor of preventive measures imposed on the elderly

population was lower and this group had no discernment of the COVID-19 complications. Thus, they probably became more prone to social isolation and, consequently, did not develop depressive symptoms.

We consider that the use of the electronic form could be a limitation for this study, since it could induce subjectivity in the interpretation of questions by the participants. To minimize this bias and before starting the study, we applied a pilot form with the aim of evaluating and improving the quality of the questions, alternative answers and avoiding possible misinterpretations. As a result of the adjustments, the final form is easier and clearer for elderly understanding.

Another limitation of this study was the impossibility of selecting, through "selection criteria," only elderly people with the ability to handle electronic devices. This fact may have restricted the number of people who could have participated in the study, and consequently, may have been a bias. However, many of the elderly participants had the help of family members with such skill during the completion of the form, which may have reduced this bias. Although the study included participants from all Brazilian states, the predominance of women among the participants may have interfered with gender representation and may be a bias in terms of Brazilian population representation.

This study is important because it evaluated elderly people from all Brazilian states, which allowed the identification of the main mental characteristics of Brazilian elderly people affected by the pandemic period, considering the ethnic, social, and cultural plurality of this population (59). In additon, in this study, it was possible to recruit a large number of the participants and it was the only one to characterize the profile of mental health and the prevalence of depression associated with the pandemic period in the Brazilian elderly population.

With the results obtained in the study, which made it possible to know the characteristics of the elderly who developed or worsened symptoms of anxiety and depression, therapeutic strategies aimed at groups that are more likely to be anxious and depressive can be devised. People with mental illness or who share the characteristics found in the research may be unable or unwilling to protect themselves against COVID-19 due to apathy, depression, paranoia, or other psychiatric symptoms. Therefore, early identification of these symptoms is of fundamental importance for the resolution of the condition of these patients (60).

Conclusion

Overall, this study identified that for the sample of elderly people studied, the most important characteristics to identify the group with depression during the COVID-19 pandemic were (1) showing signs of anxiety during the COVID-19 pandemic; (2) of low education; (3) being divorced; (4) having more than one mental disorder; (5) reading, watching, or listening to information about COVID-19, and (6) being previously diagnosed with depression.

In conclusion, elderly Brazilians in social isolation tend to develop depressive disorders during quarantine. Having anxiety, low education, and marital status were the most important variables to characterize the depressive group. Thus, we can consider that the pandemic requires effective and safe gerontological care and monitoring, especially with regard to mental health.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

Data collection was performed after approval of the Research Project by the Ethics Committee of the Institute of Health Sciences of the Federal University of Pará (CAAE number: 32893620.8.0000.0018). The patients/participants provided their written informed consent to participate in this study.

Author contributions

IS, OS, and FV: study conception and design. IS, RS, DL-J, AP, GC, AV, FT, OS, and FV: methodology and data collection. RS and MT: modeling and statistical analysis. IS, GC, AV, and FT: descriptive analysis. AP and DL-J: article editors. RS, MT, AP, DL-J, OS, and FV: scientific consultants and correction supervision. All authors contributed to the article and approved the submitted version.

Funding

We are grateful to PROPESP/UFPA (PAPQ) for funding the article publication fee. This study was partially funded by the Institutional Program for Scientific Initiation Scholarships (PIBIC) of the Federal University of Pará (UFPA), contemplating a scholarship (PRO4538-2020).

Acknowledgments

We thank the Study and Research Group on Health and Education in the Amazon – GEPSEA and the PIBIC of the Federal University of Pará (UFPA), which awarded a scientific initiation scholarship for the IS (PRO4538-2020).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2022.888234/full#supplementary-material

References

1. WHO. World Health Organization. *Coronavirus (COVID-19) Dashboard*. (2021). Available online at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance (accessed Sept 2, 2021),

2. Brasil MS. *Coronavirus (covid-19) Panel in Brazil by the Ministry of Health*. (2020). Available at: https://covid.saude.gov.br (accessed Aug 10, 2020).

3. Fiorillo A, Gorwood P. The consequences of the COVID-19 pandemic on mental health and implications for clinical practice. *J Eur Psychiatry*.(2020) 63. doi: 10.1192/j.eurpsy.2020.35

4. Bridgland VM, Moeck EK, Green DM, Swain TL, Nayda DM, Matson LA, et al. Why the COVID-19 pandemic is a traumatic stressor. *PloS ONE*. (2021) 16:e0240146. doi: 10.1371/journal.pone.0240146

5. Unützer J, Kimmel RJ, Snowden M. Psychiatry in the age of COVID-19. *World J Psychiatry*. (2020) 19:130. doi: 10.1002/wps.20766

6. Perracini MR, De Amorim JSC, Lima CA, Da Silva A, Trombini-Souza F, Pereira DS, et al. Impact of COVID-19 pandemic on life-space mobility of older adults living in Brazil: REMOBILIZE study. *J Front Public Health*. (2021) 9:643640. doi: 10.3389/fpubh.2021.643640

7. Romero DE, Muzy J, Damacena GN, Souza NAd, Almeida WdSd, Szwarcwald CL, et al. Older adults in the context of the COVID-19 pandemic in Brazil: effects on health, income and work. *J Cadernos de Saúde Pública*. (2021) 37:e00216620. doi: 10.1590/0102-311x00216620

8. Choi EPH, Hui BPH, Wan EYF. Depression and anxiety in Hong Kong during COVID-19. *Int J Environ Res Public Health*. (2020) 17:3740. doi: 10.3390/ijerph17103740

9. Jordan RE, Adab P, Cheng K. Covid-19: risk factors for severe disease and death. *BMJ*. (2020) 368:m1198. doi: 10.1136/bmj.m1198

10. Barbosa IR, Galvão MHR, Souza TA, Gomes SM, Medeiros AdA, Lima KC. Incidence of and mortality from COVID-19 in the older Brazilian population and its relationship with contextual indicators: an ecological study. *Rev Bras Geriatr Gerontol*. (2020) 23:e200171. doi: 10.1590/1981-22562020023.200171

11. da Silva Marçal DF, Gaspar PAA. de Lima LT, Coelho DT, Reinaldo JZ, Dechen VM, et al. Mortality from COVID-19 and vaccination in the elderly: an ecological study in the city of Curitiba, Paraná, Brazil. *J Infects Dis*. (2022) 26:102039. doi: 10.1016/j.bjid.2021.102039

12. Orellana JDY. Cunha GMd, Marrero L, Leite IdC, Domingues CMAS, Horta BL. Changes in the pattern of hospitalizations and deaths from COVID-19 after substantial vaccination of the elderly in Manaus, Amazonas, Brazil. *Cad Saude Publica*. (2022) 38:PT192321. doi: 10.1590/0102-311xpt192321

13. Klinenberg E. Social isolation, loneliness, and living alone: identifying the risks for public health. *Am J Public Health*. (2016) 106:786–7. doi: 10.2105/AJPH.2016.303166

14. Armitage R, Nellums LB. COVID-19 and the consequences of isolating the elderly. *J Lancet Public Health*. (2020) 5:e256. doi: 10.1016/S2468-2667(20)30061-X

15. Santini ZI, Jose PE, Cornwell EY, Koyanagi A, Nielsen L, Hinrichsen C, et al. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. *J Lancet Public Health*. (2020) 5:e62–70. doi: 10.1016/S2468-2667(19)30230-0

16. Lábadi B, Arató N, Budai T, Inhóf O, Stecina DT, Sík A, et al. Psychological wellbeing and coping strategies of elderly people during the COVID-19 pandemic in Hungary. *J Aging Mental Health*. (2022) 26:570–7. doi: 10.1080/13607863.2021.1902469

17. Aurélio SS. *Physical Activity in Combating the Incidence of Depression and Anxiety in the COVID-19 Pandemic: A Literature Review*. (2020). Available online at: https://repositorio.animaeducacao.com.br/handle/ANIMA/12635 (accessed May 20, 2021).

18. Kurebayashi LFS, Gnatta JR, Borges TP, da Silva MJP. Traditional Chinese Medicine diagnostic evaluation of stress symptoms treated by auriculotherapy: clinical trial. *J Rev Eletr Enf*. (2014) 16:68–76. doi: 10.5216/ree.v16i1.20167

19. Telles-Correia D, Barbosa A. Anxiety and depression in medicine: theoretical models and evaluation. *J Acta Med Port*. (2009) 22:89–98.

20. Drago SMMS, Martins RML. Depression in the elderly. *Asia Pac Psychiatry*. (2011) 3:46–53 doi: 10.1111/j.1758-5872.2011.00119.x

21. World Health Organization. *The Injury Chart Book: A Graphical Overview of the Global Burden of Injuries*. World Health Organization 2002 (2002).

22. Webb L. Covid-19 lockdown: a perfect storm for older people's mental health. *J Psychiatr Ment Health Nurs*. (2020) 28:300. doi: 10.1111/jpm.12644

23. Queiroz MIP. Cultural identity, national identity in Brazil. *J Tempo soc*. (1989) 1:29–46. doi: 10.1590/ts.v1i1.83318

24. Mendonça AG. Republic and religious plurality in Brazil. *J Revista USP*. (2003) 59:144–63. doi: 10.11606/issn.2316-9036.v0i59p144-163

25. Viana SAA, Lima Silva M, Lima PT. Impact on the mental health of the elderly during the period of social isolation due to the spread of the COVID-19 disease: a literary review. *Diálogos Saúde*. (2020) 3:1–16.

26. Lima Monteiro IV, Figueiredo JFC, Cayana EG. Elderly and mental health: impacts of the COVID-19 Pandemic. *Braz J Med Biol Res*. (2021) 4:6050–61. doi: 10.34119/bjhrv4n2-162

27. Almeida Costa F, Santos Silva AO. Caio Bismarck Silva, Costa LCS, Silva Paixão MÉ, Celestino MNS, et al. COVID-19: its clinical and psychological impacts on the elderly population. *Braz Dev*. (2020) 6:49811–24. doi: 10.34117/bjdv6n7-580

28. Oliveira VV, Oliveira LVV, Rocha MR, Leite IA, Lisboa RSA, Kelly CL. Impacts of social isolation on the mental health of the elderly during the COVID-19 pandemic. *Braz J Health Rev*. (2021) 4:3718–27. doi: 10.34119/bjhrv4n1-294

29. Wang Z, Qi S, Zhang H, Mao P, He Y, Li J, et al. Impact of the COVID-19 epidemic on anxiety among the elderly in community. *J Zhonghua yi xue za zhi*. (2020) 100:3179–85. doi: 10.3760/cma.j.cn112137-20200720-02167

30. Ozamiz NE, Dosil MS, Picaza MG, Idoiaga NM. Niveles de estrés, ansiedad y depresión en la primera fase del brote del COVID-19 en una muestra recogida en el norte de España. *Cad Saude Publica*. (2020) 36:e00054020. doi: 10.1590/0102-311x00054020

31. Gorrochategi MP, Munitis AE, Santamaria MD, Etxebarria NO. Stress, anxiety, and depression in people aged over 60 in the COVID-19 outbreak in a sample collected in Northern Spain. *Am J Geriatr Psychiatry*. (2020) 28:993–8. doi: 10.1016/j.jagp.2020.05.022

32. García-Fernández L, Romero-Ferreiro V, López-Roldán P, Padilla S, Rodriguez-Jimenez R. Mental health in elderly Spanish people in times of COVID-19 outbreak. *Am J Geriatr Psychiatry*. (2020) 28:1040–5. doi: 10.1016/j.jagp.2020.06.027

33. Iasevoli F, Fornaro M, D'Urso G, Galletta D, Casella C, Paternoster M, et al. Psychological distress in patients with serious mental illness during the COVID-19 outbreak and one-month mass quarantine in Italy. *Psychol Med.* (2021) 51:1054–6. doi: 10.1017/S0033291720001841

34. Martiny C, de Oliveira Silva AC, Nardi AE, Pachana NA. Translation and cross-cultural adaptation of the Brazilian version of the geriatric anxiety inventory (GAI)/Traducao e adaptacao transcultural da versao Brasileira do inventario de ansiedade geriatrica (GAI). *Rev Psiquiatr Clin.* (2011) 38:8–13. doi: 10.1590/S0101-60832011000100003

35. Pachana NA, Byrne GJ, Siddle H, Koloski N, Harley E, Arnold E. Development and validation of the geriatric anxiety inventory. *Int Psychogeriatr.* (2007) 19:103–14. doi: 10.1017/S1041610206003504

36. Almeida OP, Almeida SA. Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. *Int J Geriatr Psychiatry.* (1999) 14:858–65. doi: 10.1002/(SICI)1099-1166(199910)14:10<858::AID-GPS35>3.0.CO;2-8

37. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen.* (2012) 141:2–18. doi: 10.1037/a0024338

38. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Cambridge, MA: Academic press. (2013). doi: 10.4324/9780203771587

39. Russell Stuart J, Norvig P. *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Prentice Hall Press. (2009).

40. Moccia L, Janiri D, Giuseppin G, Agrifoglio B, Monti L, Mazza M, et al. Reduced hedonic tone and emotion dysregulation predict depressive symptoms severity during the COVID-19 outbreak: an observational study on the Italian general population. *Int J Environ Res Public Health.* (2021) 18:255. doi: 10.3390/ijerph18010255

41. Janiri D, Kotzalidis G, Giuseppin G, Molinaro M, Modica M, Montanari S. Psychological distress after COVID-19 recovery: reciprocal effects with temperament and emotional dysregulation. An exploratory study of patients over 60 years of age assessed in a post-acute care service. *Front Psychiatry.* (2020) 11:590135. doi: 10.3389/fpsyt.2020.590135

42. Zaninotto P, Iob E, Demakakos P, Steptoe A. Immediate and longer-term changes in the mental health and wellbeing of older adults in England during the COVID-19 pandemic. *J JAMA Psychiatry.* (2022) 79:151–9. doi: 10.1001/jamapsychiatry.2021.3749

43. Piano C, Di Stasio E, Primiano G, Janiri D, Luigetti M, Frisullo G, et al. An Italian neurology outpatient clinic facing SARS-CoV-2 pandemic: data from 2,167 patients. *Front Neurol.* (2020) 11:564. doi: 10.3389/fneur.2020.00564

44. Sareen J, Erickson J, Medved MI, Asmundson GJ, Enns MW, Stein M, et al. Risk factors for post-injury mental health problems. *Depress Anxiety.* (2013) 30:321–7. doi: 10.1002/da.22077

45. Barros MBA, Lima MG. Malta DC, Szwarcwald CL, Azevedo RCSd, Romero D, et al. Report of sadness/depression, nervousness/anxiety and sleep problems in the Brazilian adult population during the COVID-19 pandemic. *Epidemiol serv* saude. (2020) 29:e2020427.

46. Meng H, Xu Y, Dai J, Zhang Y, Liu B, Yang H. Analyze the psychological impact of COVID-19 among the elderly population in China and make corresponding suggestions. *J Psychiatry Res.* (2020) 289:112983. doi: 10.1016/j.psychres.2020.112983

47. Barbosa TR. *Confronting Domestic Violence Against Women in the Context of a Pandemic*. (2020). Available online at: http://ri.ucsal.br:8080/jspui/handle/prefix/2729 (accessed Jun 2, 2021).

48. Morais AC. *Depression in Women Victims of Domestic Violence*. (2009). [Masters dissertation]. Available online at: https://repositorio.ufba.br/ri/bitstream/ri/11425/1/Disserta%C3%A7%C3%A3o_Enf_Ariane%20Cedraz%20Morais.pdf (Accessed May 12, 2021).

49. Oliveira LM, Abrantes GG, Ribeiro GS, Cunha NM, Pontes MLF, Vasconcelos SC. Loneliness in senescence and its relationship with depressive symptoms: an integrative review. *Rev Bras Geriatr Gerontol.* (2020) 22:e190241. doi: 10.1590/1981-22562019022.190241

50. Ramos M, Wilmoth J. Social relationships and depressive symptoms among older adults in southern Brazil. *J Gerontol B Psychol Sci Soc Sci.* (2003) 58:S253–S61. doi: 10.1093/geronb/58.4.S253

51. Nóbrega IRAP, Leal MCC. Marques APO, Vieira JCM. Factors associated with depression in institutionalized elderly: an integrative review. *J Saúde em Debate.* (2015) 39:536–50. doi: 10.1590/0103-110420151050002020

52. Oliveira MCGM, Salmazo-Silva H, Gomes L, Moraes CF, Alves VP. Elderly individuals in multigenerational households: Family composition, satisfaction with life and social involvement. *Estud Psicol.* (2020) 37:e180081. doi: 10.1590/1982-0275202037e180081

53. Almeida MASO, Lemes AG, do Nascimento VF, da Fonseca PIMN, da Rocha EM, Volpato RJ, et al. Risk factors associated with depression in the elderly in the interior of Mato Grosso. *Rev Saude Publica.* (2015) 39:627. doi: 10.22278/2318-2660.2015.v39.n3.a1895

54. Pardal A. *Social Support, Symptoms of Anxiety and Depression and Life Satisfaction in the Elderly Under Social Response*. (2011). Available online at http://repositorio.ismt.pt/jspui/handle/123456789/129 (accessed May 10, 2021).

55. Minghelli B, Tomé B, Nunes C, Neves A, Simões C. Comparison of anxiety and depression levels between active and sedentary elderly. *Arch Clin Psychiatry.* (2013) 40:71–6. doi: 10.1590/S0101-60832013000200004

56. Almeida-Filho N, Lessa I, Magalhães L, Araújo MJ, Aquino E, James SA, et al. Social inequality and depressive disorders in Bahia, Brazil: interactions of gender, ethnicity, and social class. *Soc Sci Med.* (2004) 59:1339–53. doi: 10.1016/j.socscimed.2003.11.037

57. Maximiano-Barreto MA, Oliveira Fermoseli AF. Prevalence of anxiety and depression in elderly people with low education in Maceió/Al. *J Psicologia, Saúde e Doenças*. (2017) 18:801–13. doi: 10.15309/17psd180314

58. Sougey EB, Del Porto JA. Brasil MA, Juruena MF. Guidelines of the Brazilian medical association for the treatment of depression. *Braz J Psychiatry.* (2003) 25:114–22. doi: 10.1590/S1516-44462003000200013

59. Canen A. Multicultural education, national identity and cultural plurality: tensions and curricular implications. *Cad Pesqui.* (2000) 111:135–49. doi: 10.1590/S0100-15742000000300007

60. Stewart DE, Appelbaum PS. COVID-19 and psychiatrists' responsibilities: a WPA position paper. *World J Psychiatry.*(2020) 19:406. doi: 10.1002/wps.20803

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 06 October 2022
DOI 10.3389/fpsyt.2022.972628

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University, Germany

REVIEWED BY
Xenia Gonda,
Semmelweis University, Hungary
Anne Toussaint,
University Medical Center
Hamburg-Eppendorf, Germany

*CORRESPONDENCE
Jelena Vrublevska
vrublevskaja@inbox.lv

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 18 June 2022
ACCEPTED 20 September 2022
PUBLISHED 06 October 2022

CITATION
Vrublevska J, Renemane L, Kivite-Urtane A and Rancans E (2022) Validation of the generalized anxiety disorder scales (GAD-7 and GAD-2) in primary care settings in Latvia. *Front. Psychiatry* 13:972628. doi: 10.3389/fpsyt.2022.972628

Validation of the generalized anxiety disorder scales (GAD-7 and GAD-2) in primary care settings in Latvia

Jelena Vrublevska[1,2]*, Lubova Renemane[1], Anda Kivite-Urtane[2] and Elmars Rancans[1]

[1]Department of Psychiatry and Narcology, Riga Stradins University, Riga, Latvia, [2]Institute of Public Health, Riga Stradins University, Riga, Latvia

Background: Anxiety disorders are the most prevalent mental disorders in the world and have an important impact on the global burden of disease. Generalized anxiety disorder (GAD) is the most prevalent anxiety disorder encountered in primary care. There are no available validated anxiety screening tools in primary care in Latvia. We aimed to validate both a seven-item and a two-item generalized anxiety disorder scale (GAD-7 and GAD-2) in the Latvian and Russian languages, to detect generalized anxiety disorder (GAD) in primary care settings in Latvia.

Methods: During a 1-week period, all patients aged 18 years or older visiting their GP (general practitioners) with any health concern at 24 primary care settings throughout Latvia were invited to complete the GAD-7 in their native language (Latvian or Russian). Criterion validity was assessed against the Mini International Neuropsychiatric Interview (MINI).

Results: The study sample included 1,459 participants who completed the GAD-7 and the MINI. The GAD-7 items showed good internal reliability [Cronbach's alpha 0.87 for Latvian version and 0.85 for Russian version (for Latvia) of the GAD-7]. A cut-off score for detecting GAD of 5 or above was estimated for Latvian version of the GAD-7 (sensitivity 75.4%, specificity 68.9%, respectively) and 7 or above for Russian version of the GAD-7 (sensitivity 73.3%, specificity 84.1%, respectively). The internal reliability of the GAD-2 was lower for both languages (Cronbach's alpha 0.75 for Latvian version and 0.68 for Russian version of the GAD-2). A cut-off score of 2 or above was established for both the Latvian, and Russian versions of the GAD-2 (sensitivity 78.9 and 83.3%; specificity 63.7 and 69.1% for the Latvian and Russian versions of the GAD-2, accordingly) for detecting GAD.

Conclusions: This is the first study to report criterion validity of the Latvian and Russian (for Latvia) versions of the GAD-7 and GAD-2, assessed in a nationwide study conducted at the primary care level.

KEYWORDS

generalized anxiety disorder (GAD), mental disorder, primary care, validated anxiety screening, Mini International Neuropsychiatric Interview (MINI), Latvia, GAD-2, GAD-7

Introduction

Anxiety disorders are the most prevalent mental disorders in the general population in the world and have a significant impact on the global burden of disease (1). They are receiving increasing attention because of their early onset as well as their tendency to recur and cause disability (2, 3). Estimates of the prevalence of anxiety disorders vary widely across studies and population groups. Different studies demonstrate lifetime prevalence rates of anxiety disorders ranging from 5.1 to 16.6% in general population, and from 7.2 to 19.5% in primary care (4–9). Moreover, anxiety disorders are often undetected, undertreated, and associated with the global health-related, personal and societal burden. In addition, they can cause substantial impairment of quality of life (10).

According to the latest evidence, anxiety disorders are becoming more prevalent. A recent systematic review estimated an additional 76.2 million cases of anxiety disorders globally (an increase of 25.6%). Additionally, the data suggest that anxiety disorders caused 44.5 million disability-adjusted life-years globally in 2020 (11). Another systematic review indicates that the rates of anxiety disorders in the general population could be more than 3 times higher in recent years (12).

Generalized anxiety disorder (GAD) is the most prevalent anxiety disorder encountered in primary care, with an estimated point prevalence of 8%. The disorder is present in 22% of primary care patients who complain of anxiety symptoms (9, 13). The high prevalence rates underline the necessity of identification and assessment of GAD in primary care settings, but many people who might benefit from treatment are not recognized. Moreover, of those patients who are diagnosed as suffering from GAD, 41% do not receive the adequate treatment (5). The data from previous studies suggest that GAD could be the most frequent anxiety disorder causing 'completed' suicides; also sub-threshold GAD is clearly linked to suicide ideation (14). Anxiety disorders rank as the second leading diagnostic category (15.8%) in primary care in Latvia, based on the assessment with the MINI, with the prevalence of GAD of 6.1% (95% CI 4.9–7.3) (7).

It is estimated that the prevalence of diagnosis and treatment of anxiety disorders in primary care is much lower than expected, given their prevalence (15). The major problems in primary care are time constraints and the existence of comorbid depressive disorders and chronic physical health problems (16). Therefore, self-reported rating scales are often preferred in primary care level. The underdiagnosis of anxiety disorders appears to be a worrying issue for Latvia as well, since the data from the National Health Service Register show that the most prevalent diagnosed mental disorders in Latvia are organic mental disorders, schizophrenia spectrum disorders, but not neurotic and affective mental disorders, which are the most prevalent worldwide. Moreover, among neurotic spectrum disorders, Latvian GPs most frequently diagnose somatoform autonomic dysfunction (17, 18).

The NICE (National Institute for Health and Care Excellence) provides the evidence-based clinical guidelines for identification and assessment of common mental health problem, and recommends the use of the 2-item generalized anxiety disorder (GAD-2) tool for identification, and the 7-item generalized anxiety disorder scale (GAD-7) for assessment of anxiety disorder severity (19). A recent systematic review of validated screening tools for anxiety disorders that included 58 articles and 77 screening tools, demonstrated that the GAD-7 was one of the most commonly validated tools for anxiety disorders (20).

The GAD-7 was developed as a brief self-reported screening tool to detect probable cases of GAD among primary care patients, and assess its severity in clinical practice and research (21). The GAD-2, consists of the first two questions of the GAD-7, is a shorter version of the tool, and is used as a screening test for detection of GAD (5). The GAD-7 and the GAD-2 were validated in primary care patients and have been widely used by general practitioners (16). Earlier studies suggested that the GAD-7 and GAD-2 perform well for screening not only GAD, but can also be used for detecting other anxiety disorders such as panic disorder, social anxiety disorder and post-traumatic stress disorder (5, 16).

Till now, there are no published studies examining the psychometric properties of anxiety screening tools among the Latvian- and Russian-speaking population of Latvia. As the ethnic distribution of the Latvian population is more than 61% Latvian and the remaining are mostly Russian-speaking, it is critical to perform validation in both Latvian and Russian languages (22). Therefore, we aimed to investigate psychometric properties of the GAD-7 and GAD-2 to provide the reliability and validity of these tools, and recommended screening cut-off scores for GAD, using the Mini International Neuropsychiatric Interview (MINI) as the reference standard in a large sample among the Latvian primary care population.

Materials and methods

Procedure and participants

The study was conducted within the framework of the National Research Program, BIOMEDICINE 2014–2017, which aimed to estimate the prevalence of mental disorders in primary care settings in Latvia. The program was funded by the Latvian Ministry of Education and Science. The main aim of this program was to develop new methods and practices for the prevention, treatment and diagnosis of mental disorders, as also biomedical technologies to improve public health in Latvia. It comprised certain areas: cardiovascular and metabolic

diseases, oncological diseases, and childhood and infectious diseases. Mental health was included in the program for the first time. Study participants did not receive any financial compensation for their participation. Within the project the validity of the PHQ-9 and the PHQ-2 was assessed and a cut-off score to identify depression was established (23). Patients visiting their general practitioners (GPs) for any medical reason were recruited from 24 primary care settings (16 in urban and 8 in rural regions) that covered all regions of Latvia. The survey was conducted in Latvian or in Russian, as per patient preference.

All patients, aged 18 years or older, visiting a primary care physician with any health concern, during a 1-week period, were invited to participate in the study. Those who visited their GPs for administrative reasons were not included. The others who were excluded were the patients who refused to participate in the study, patients younger than 18 years of age, and those who were not able to participate due to acute medical conditions requiring hospitalization or other general medical conditions (one patient was deaf-mute). All consecutive patients were invited to complete the paper-and-pencil form of the GAD-7 in their preferred language (Latvian or Russian) before seeing the GP, and were requested to complete a structured socio-demographic questionnaire. All ambiguities and questions that arose were clarified by the researcher.

The Mini International Neuropsychiatric Interview (MINI) Version 6.0.0 was conducted over the phone by four trained psychiatrists (who were unaware of the GAD-7 scores), no more than 2 weeks after the first contact with the patient. The MINI was used as the standard to determine the presence of GAD and other anxiety disorders. Participants with high scores of the GAD-7, the PHQ-9 and those who were diagnosed with GAD, or any other diagnostic category according to the MINI, were referred for appropriate care.

Riga Stradins University Ethics Committee approved this study (No. 8/18.06.2015.), and written informed consent was obtained from all participants. The study was carried out in accordance of the Declaration of Helsinki and its subsequent amendments.

Measures

The GAD-7 consists of 7 self-reported items, measuring symptoms of anxiety, allowing the rapid screening for GAD. Each item has a Likert-response format on a 4-point scale (0–3 points). Respondents were asked to consider the previous 2 weeks and to rate symptom frequency as 'not at all' (0), 'several days' (1), 'more than half of all days' (2) or 'nearly all days' (3). The total score response ranged from 0 to 21. In the initial validation study of the GAD-7, estimated sensitivity and specificity were identified at 89 and 82%, respectively, at a cut-off score of 9 (21).

The GAD-2 is a shorter version of the tool that is composed of the first two questions of the GAD-7. The GAD-2 in its initial validation study had a sensitivity of 86% and a specificity of 83% at a cut-off score of 2 (5).

A forward/backward translation of the GAD-7 into the Latvian and Russian languages was performed by professional translators and was reviewed by Latvian and Russian language speaking psychiatrists. Additionally, the evaluation of potential problems in comprehension or cultural differences of scale was discussed in a professional focus group. The final agreement of both language versions of the GAD-7 was reached.

The MINI is a structured diagnostic interview for psychiatric disorders according to the Diagnostic and Statistical Manual of Mental Disorders, and the International Classification of Disease, 10th revision (24). It is widely used for research purposes in psychiatric and general populations, including primary care patients (25, 26). The MINI has been translated and adapted by authorship holders for use in 67 languages, including Latvian and Russian (27). It consists of 120 questions and screens 17 axis I disorders for 24 current and lifetime diagnoses. The interview was conducted over the telephone, which is acceptable and has been used in other studies (28). We administrated all modules of the MINI to identify current diagnoses of anxiety disorders, such as panic disorder, agoraphobia, social phobia, obsessive-compulsive disorder, posttraumatic stress disorder, and generalized anxiety disorder.

The participants' sex (male or female), age (18–34, 35–49, 50–64, or 65+ years), marital status (married/cohabiting, single, or living separately/divorced/widowed), employment status (employed, unemployed, or economically inactive), educational level (higher/unfinished higher education, general/vocational secondary/unfinished secondary education, or 9-year basic/unfinished basic education), and place of residence [urban: capital (Riga)/other city, or rural] were recorded.

Statistical analysis

The internal consistency of the GAD-7 and GAD-2 was assessed by Cronbach's alpha coefficient, while their criterion validity was assessed by receiver operating characteristic (ROC) analysis. The criterion validity was analyzed in terms of sensitivity (true positive), specificity (true negative), positive and negative predictive values [PPV, NPV; the probability that individuals with a positive (negative) test result truly have (do not have) the condition], a positive likelihood ratio (LR+; "probability that a positive test would be expected in a patient divided by the probability that a positive test would be expected in a patient without a disease"), and a negative likelihood ratio (LR–; "the probability of a patient testing negative who has a disease divided by the probability

of a patient testing negative who does not have a disease") for different cut-off scores (29). The Latvian and Russian versions of the MINI, which were used to diagnose GAD and other anxiety disorders, served as the criterion standard. Data analyses were performed using IBM- SPSS (Statistical Package for the Social Sciences), version 26.0. A separate analysis was conducted for the responders who answered the survey in Latvian, and those who used Russian translation of the survey. ROC curves were created for each instrument. The area under the curve (AUC) which is a measure that provides an overall summary of the utility of the scale to correctly identify GAD cases was determined. The statistical significance of the differences of demographic characteristics between groups of mental disorders was assessed using Chi-Squared test of Fisher's exact test. The results were considered as statistically significant if $p < 0.05$.

Results

Of the 1,756 patients who visited their GP, 152 refused to participate. At baseline, a sample of 1,604 patients was approached to complete the GAD-7 and GAD-2. Response rate among the patients was 91.3% and varied in the range 86.3–93.7% across 24 primary care settings all over the country. The questionnaires were completed by 1,585 participants. Of those who completed the screening questionnaire, 100 did not agreed to be interviewed with the MINI over phone or did not answer the telephone call three times within 2 weeks, and were excluded from the study. Those patients who were excluded from the study did not show statistically significant differences in sociodemographic status compared to those who were included. The remaining 1,485 patients were interviewed with the MINI over the telephone. The questionnaires of 18 patients had to be

TABLE 1 Demographic characteristics of study sample with respect to current mental disorders established by the Mini International Neuropsychiatric Interview ($n = 1,467$).

Variable	Total Sample		No anxiety disorders		GAD only		Anxiety disorders without GAD		GAD + other Anxiety Disorders		*p*
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	
Total	**1467**	**100.0**	**1236**	**84.3**	**61**	**4.2**	**142**	**9.7**	**28**	**1.9**	
Sex											
Female	1019	69.5	839	67.9	43	70.5	113	79.6	24	85.7	**0.008**
Male	448	30.5	397	32.1	18	29.5	29	20.4	4	14.3	
Age											
18–34	209	14.2	172	13.9	7	11.5	23	16.2	7	25.0	0.25
35–54	455	31.0	385	31.1	13	21.3	48	33.8	9	32.1	
55–64	349	23.8	288	23.3	19	31.1	34	23.9	8	28.6	
65+	454	30.9	391	31.6	22	36.1	37	26.1	4	14.3	
Education											
Higher and unfinished higher education	436	29.9	389	31.6	11	18.6	25	17.6	11	40.7	**0.001**
General or vocational secondary and unfinished secondary	838	57.4	692	56.2	34	57.6	97	68.3	15	55.6	
9-year basic, unfinished basic	185	12.7	150	12.2	14	23.7	20	14.1	1	3.7	
Employment status											
Employed	776	53.2	655	53.2	26	43.3	79	55.6	16	59.3	0.06
Unemployed	82	5.6	62	5.0	4	6.7	12	8.5	4	14.8	
Economically inactive	602	41.2	514	41.8	30	50.0	51	35.9	7	25.9	
Marital status											
Married. cohabiting	895	61.3	768	62.4	33	55.0	81	57.0	13	48.1	0.09
Single	144	9.9	116	9.4	6	10.0	15	10.6	7	25.9	
Live separately, divorced, widowed	421	28.8	347	28.2	21	35.0	46	32.4	7	25.9	
Place of residence											
Capital (Riga)	303	20.7	255	20.6	20	32.8	21	14.8	7	25.0	**0.005**
Other city	692	47.2	598	48.4	26	42.6	59	41.5	9	32.1	
Rural	472	32.2	383	31.0	15	24.6	62	43.7	12	42.9	

discarded due to insufficient data quality. Of the 1,467 patients, eight patients were missing because the language in which the GAD-7 was completed was not specified. Finally, 1,459 patients were included in the analysis.

The demographic characteristics of our study sample with respect to current anxiety disorders determined by the MINI are summarized in Table 1. According to the MINI, 61 patients (4.2%) were diagnosed with GAD, 142 patients (9.7%) with anxiety disorder without GAD and 28 patients (1.9%) had comorbidity of GAD and other anxiety disorders.

In the total sample ($n = 1,467$) the mean score of the GAD-7 was 4.1 [standard deviation (SD) 4.0] and of the GAD-2–1.5 (SD 1.4). Whereas in the group of patients with GAD as per the MINI ($n = 89$) the mean score of the GAD-7 and GAD-2 was 8.7 (SD = 5.1) and 3.0 (SD = 1.8), respectively.

Cronbach's alpha for the Latvian version of the GAD-7 and GAD-2 was 0.87 and 0.75, respectively, and for the Russian version of the GAD-7 was 0.85, indicating good internal consistency. However, Cronbach's alpha for the Russian version of the GAD-2 was found to be 0.68, demonstrating a questionable level of internal consistency.

All items in the GAD-7 for both languages were significantly and positively associated with the total GAD-7 scores, and Cronbach's alpha did not decrease if the items were deleted. The data presented in Tables 2, 3 demonstrate corrected item-total correlations, Cronbach's alpha, scale mean, and scale variance when an item is deleted from the GAD-7 scale in Latvian and Russian versions.

The ROC analysis of the GAD-7 and GAD-2 for the diagnosis of GAD, established by the MINI, is shown in Table 4. The ROC curves of the GAD-7 and GAD-2 are

TABLE 2 Corrected item-total correlations and Cronbach's alpha, scale mean, scale variance when an item is deleted from the GAD-7 and GAD-7 in Latvian ($n = 908$).

	Scale mean if an item deleted	Scale variance if an item deleted	Corrected item-total correlation	Cronbach's alpha if an item deleted
GAD-7 Latvian				
GAD7: 1. Feeling nervous. anxious or on edge	3.09	10.66	0.71	0.84
GAD7: 2. Not being able to stop or control worrying	3.63	11.07	0.74	0.84
GAD7: 3. Worrying too much about different things	3.31	10.78	0.68	0.85
GAD7: 4. Trouble relaxing	3.60	11.15	0.68	0.85
GAD7: 5. Being so restless that it is hard to sit still	3.80	12.28	0.60	0.86
GAD7: 6. Becoming easily annoyed or irritable	3.46	11.87	0.56	0.86
GAD7: 7. Feeling afraid as if something awful might happen	3.68	12.02	0.57	0.86
GAD-2 Latvian				
GAD7: 1. Feeling nervous. anxious or on edge	0.46	0.53	0.60	.
GAD7: 2. Not being able to stop or control worrying	1.00	0.69	0.60	.

TABLE 3 Corrected item-total correlations and Cronbach's alpha, scale mean, scale variance when an item is deleted from the GAD-7 in Russian ($n = 551$).

	Scale mean if item deleted	Scale variance if item deleted	Corrected item-total correlation	Cronbach's alpha if item deleted
GAD-7 Russian				
GAD7: 1. Feeling nervous, anxious or on edge	3.03	11.77	0.63	0.83
GAD7: 2. Not being able to stop or control worrying	3.72	12.45	0.69	0.82
GAD7: 3. Worrying too much about different things	3.36	11.54	0.66	0.82
GAD7: 4. Trouble relaxing	3.58	11.98	0.67	0.82
GAD7: 5. Being so restless that it is hard to sit still	3.85	13.69	0.55	0.84
GAD7: 6. Becoming easily annoyed or irritable	3.30	12.33	0.51	0.85
GAD7: 7. Feeling afraid as if something awful might happen	3.68	12.68	0.61	0.83
GAD-2 Russian				
GAD7: 1. Feeling nervous. anxious or on edge	0.37	0.50	0.52	.
GAD7: 2. Not being able to stop or control worrying	1.06	0.79	0.52	.

10.3389/fpsyt.2022.972628

TABLE 4 The ROC analyses of the GAD-7 and GAD-2 Latvian and Russian versions for the diagnosis of GAD established by the MINI (*n* GAD-7 and GAD-2 Latvian = 908; *n* GAD-7 and GAD-2 Russian = 551).

Cut of score	Sensitivity, %	Specificity, %	PPV, %	NPV, %	LR+	LR–
LATVIAN						
GAD-7						
≥3	86.0	44.4	9.4	97.9	1.55	0.32
≥4	80.7	57.9	11.4	97.8	1.92	0.33
≥5	75.4	68.9	14.0	97.7	2.42	0.36
≥6	68.4	75.3	15.7	97.3	2.77	0.42
≥7	57.9	81.3	17.2	96.6	3.10	0.52
≥8	49.1	86.3	19.3	96.2	3.58	0.59
≥9	38.6	89.4	19.6	95.6	3.64	0.69
≥10	29.8	91.7	19.3	95.1	3.59	0.77
≥11	24.6	93.2	19.4	94.9	3.62	0.81
≥12	21.1	94.9	21.8	94.7	4.14	0.83
≥13	17.5	95.8	21.7	94.5	4.17	0.86
≥14	14.0	96.6	21.6	94.4	4.12	0.89
≥15	8.8	98.2	25.0	94.1	4.89	0.93
GAD-2						
≥1	87.7	27.0	7.5	97.0	1.20	0.46
≥2	78.9	63.7	12.7	97.8	2.17	0.33
≥3	54.4	84.7	19.3	96.5	3.56	0.54
≥4	28.1	91.7	18.4	95.0	3.39	0.78
≥5	17.5	96.8	27.0	94.6	5.47	0.85
≥6	15.8	98.5	40.9	94.6	10.53	0.85
RUSSIAN GAD-7						
≥3	100.0	45.3	9.5	100.0	1.83	0.00
≥4	86.7	61.2	11.4	98.8	2.23	0.22
≥5	86.7	72.2	15.2	98.9	3.12	0.18
≥6	76.7	78.3	16.9	98.3	3.53	0.30
≥7	73.3	84.1	21.0	98.2	4.61	0.32
≥8	66.7	87.5	23.5	97.9	5.34	0.38
≥9	53.3	90.0	23.5	97.1	5.33	0.52
≥10	50.0	91.7	25.9	97.0	6.02	0.55
≥11	43.3	93.1	26.5	96.6	6.28	0.61
≥12	36.7	94.6	28.2	96.3	6.80	0.67
≥13	33.3	96.0	32.3	96.2	8.33	0.69
≥14	23.3	96.7	29.2	95.6	7.06	0.79
≥15	16.7	97.7	29.4	95.3	7.26	0.85
GAD-2						
≥1	96.7	26.3	7.0	99.3	1.31	0.13
≥2	83.3	69.1	13.4	98.6	2.70	0.24
≥3	60.0	85.0	18.8	97.4	4.00	0.47
≥4	43.3	92.5	25.0	96.6	5.77	0.61
≥5	20.0	96.7	26.1	95.5	6.06	0.83
≥6	16.7	98.1	33.3	95.3	8.79	0.85

PPV, positive predictive value; NPV, negative predictive value; LR+, positive likelihood ratio; LR–, negative likelihood ratio.

10.3389/fpsyt.2022.972628

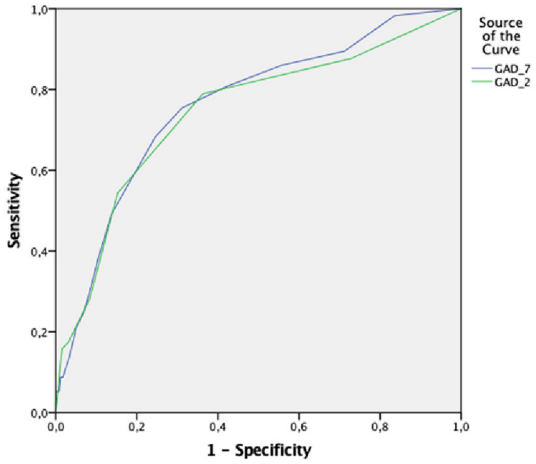


FIGURE 1
ROC (reciever operating characterstics curve of GAD-7 and GAD-2 in Latvian.

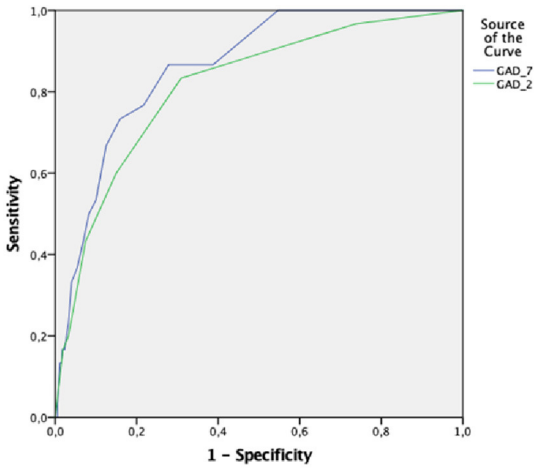


FIGURE 2
ROC (reciever operating characterstics curve of GAD-7 and GAD-2 in Russian.

illustrated in Figure 1 for Latvian versions and in Figure 2 for Russian versions.

The ROC analysis of the GAD-7 in Latvian exhibited an area under the curve (AUC) of 0.76 (SE = 0.03; $p = 0.000$; 95% CI = 0.70–0.83). Youden's index was highest with a cut-off score of 5 or above, and the GAD-7 sensitivity was 75.4%, specificity was 68.9% with a PPV of 14.0% and a NPV of 97.7%, a LR+ of 2.4 and a LR– of 0.36 for this cut-off score.

The ROC analysis of the GAD-2 in Latvian exhibited the AUC of 0.74 (SE = 0.04; $p = 0.000$; 95% CI = 0.67–0.82), and Youden's index was highest with a cut-off score of 2 or above (Figure 1). At this cut-off score the GAD-2 sensitivity was 78.9% and specificity was 63.7%, with the PPV (positive predictive value) of 12.7% and the NPV (negative predictive value) of 97.8, and the LR+ of 2.17 and the LR– of 0.33.

For the Russian version of the GAD-7 and GAD-2, the AUC (area under the ROC curve) in the ROC analysis was 0.86 (SE = 0.03; $p = 0.000$; 95% CI = 0.81–0.92) and 0.81 (SE = 0.04; $p = 0.000$; 95% CI = 0.74–0.89), respectively (Figure 2).

A cut-off score of 7 or above for the GAD-7 Russian language demonstrated sensitivity of 73.3% and specificity of 84.1%, with the PPV of 21.0% and the NPV of 98.2%, and the LR+ of 3.53 and the LR– of 0.30.

The GAD-2 Russian version indicate sensitivity of 83.3% and specificity of 69.1% at a cut-off score 2 or above. The PPV was 13.4% and the NPV was 98.6, and the LR+ was 4.61 and the LR– was 0.32 at this cut-off score.

Discussion

The present study aimed to investigate the validity of the GAD-7 and GAD-2 Latvian and Russian versions, for Latvia, and to identify a cut-off score to detect the symptoms of GAD in a nationwide sample of patients who visited their GP due to any medical reason. The reference standard in our study was a structured clinical interview (MINI) that was conducted by four trained psychiatrists. This screener so far is the only questionnaire that has been tested for anxiety symptoms in a primary care in Latvia.

Validation of the GAD-7 and GAD-2 scales has earlier been carried out in different settings, languages and populations worldwide, for example, among pregnant women, among patients with migraine, HIV, and epilepsy, and among high school students, indicating that these tools are valid and useful for screening GAD (21, 30–35).

The initial validation study for the GAD-7 that was performed in 15 primary care clinics in the United States, had a Cronbach's alpha of 0.92, and at a cut of score of 9, the GAD-7 had a sensitivity of 89%, and specificity of 82% (21).

In terms of reliability, the GAD-7 and GAD-2 Latvian versions and GAD-7 Russian version had good internal consistency (Cronbach's alpha of 0.87 for the GAD-7, 0.75 for the GAD-2 Latvian version, and 0.85 for the Russian version of the GAD-7). This result supports the homogeneity of the scale and the contribution of all the items to the measurement of anxiety symptoms. However, the Russian version of the GAD-2 had demonstrated lower level of internal consistency (Cronbach's alpha of 0.68) in comparison with the Latvian version.

Our study showed that at a cut-off score of 5 or above for the GAD-7-Latvian version, and at a cut- off score of 7 or over for the GAD-7-Russian version, had the highest sum of specificity and sensitivity. A recent systematic review of validated screening tools for anxiety disorders in low to middle

income countries identified six validation studies of the GAD-7 that were performed in different population groups with a similar methodological approach. In this review, a wide range of sensitivity (57–94%) and specificity (53–94%) was reported at cut-off scores 6 to 10, that varied depending on the regions where the studies were conducted, and sample size (20).

In another systematic review and diagnostic meta-analysis that aimed to evaluate the accuracy of the GAD-7 and GAD-2 questionnaires to identify anxiety disorders, 12 samples with 5,223 participants were analyzed. The authors suggested that the GAD-7 had acceptable properties for identifying GAD at cut-off scores ranging from 7 to 10 (36).

In a Finnish validation study of the GAD-7 carried out in primary care, it was found that the sensitivity and specificity for GAD with a cut-off point of 7 or more were 100.0 and 82.6%, respectively (34).

The identified cut-off score for the GAD in Latvian language was lower in comparison with previous studies carried out in primary care, however, the score for Russian version was consistent with a Finnish validation study of GAD-7 (34). Identified differences in the cut-off points of the GAD-7 across the studies support the suggestion that specific validation of scales is required for each country, population group and language.

The literature data on validation of the GAD-2 in primary care are limited, since it has not been as frequently validated as the GAD-7. The first validation study of the GAD-2 was done in 2007 on the primary care population of the United States of America, in which reported sensitivity and specificity were 86 and 83%, respectively, at a cut of score of 3 or greater (5). The systematic review and meta-analysis carried out in 2016, identified six samples that provided data on the accuracy of the GAD-2 for detecting GAD. The meta-analysis data suggested that pooled sensitivity and specificity values appeared acceptable at a cut-off point of 3 [sensitivity: 0.76 (95% CI 0.55–0.89), specificity: 0.81 (95% CI 0.60–0.92)] (36). The validation of a Finnish translation of the GAD-7 and GAD-2 screening tools in primary care population indicated a sensitivity of 0.83 and specificity of 0.90, at a cut-off point of 3 or more for the GAD-2 (34). Our study demonstrated that a cut-off score of 2 in the GAD-2 for both languages has the best sensitivity and specificity, and it was lower than in previous studies (5, 34, 36). Notably, the validation study of the GAD-7 and GAD-2 in patients with migraine demonstrated a cut-off score of 5 for the GAD-7, with sensitivity of 78.1% and specificity of 74.6% for the GAD-7, and a cut-off score of 1 for the GAD-2, with sensitivity of 44.6% and specificity of 94.3%, which is lower than in our study (35). These findings once again underline the necessity to validate scales in specific population groups and local languages.

Differences in cut-off scores across the countries can be explained with respect to study's settings, specific disease groups, sample size and characteristics (34, 37, 38). Another explanation includes cultural and language based differences in expression of psychopathology, and different interpretations of grading using the Likert scale (35, 39). Vast amount of literature is highlight the need for culturally and ethnically sensitive GAD screening tools (40).

Our data demonstrate that the Latvian and Russian (for Latvia) translations of the GAD-7 and GAD-2 are valid screening tools with acceptable sensitivity and specificity for GAD. Additional information is needed to further define the optimal cut-off point for Latvian and Russian versions. The GAD-7 and GAD-2 could be validated for other anxiety disorders in the future, as has been done in previous studies (5, 34).

The strengths of this study include the fact that all patients were from primary care and all of them received a MINI assessment as the reference standard. Our study included a large sample size of patients in primary care, which covered all regions of Latvia and was conducted in urban as well as rural areas. Moreover, in the study, only those patients were included, who visited their GP due to medical reasons. The respondents were assessed in the language of their preference. The patients were interviewed by four trained psychiatrists who were unaware of the GAD-7 estimates. Finally, GAD cases without any other comorbid mental disorders were included in the data analysis. Further studies in other clinical populations are necessary to evaluate its sensitivity and specificity as well as cut-off points to screen for GAD and other anxiety disorders.

This study has important practical implications. In early 2021, in response to the negative impact of the Covid-19 pandemic on the mental health of the population, the Ministry of Health of Latvia issued an information report on the dynamic follow-up of patients with mental and behavioral disorders conducted by GPs, that is, "Dynamic observation of patients with mental and behavioral disorders by a family doctor" (41). The Ministry of Health, together with mental health professionals and GPs, has developed easy-to-read algorithms using our validated GAD-2 and GAD-7 scales to help the GP assess patients with mental health issues, in order to make a diagnosis and select the appropriate treatment path and specialists to be consulted. Patients with prevalent anxiety disorders, for whom the GP does not consider referral to be necessary, can be adequately treated at the primary care level. Implementation of the GAD at the primary care level might contribute to improvement in recognition of anxiety spectrum disorders.

Conclusion

In summary, the Latvian and Russian versions of the GAD-7 and GAD-2 have moderate psychometric properties for screening for GAD. The optimal cut-off score of the GAD-7 Latvian and Russian version for Latvia, which had the best psychometric characteristics for detecting GAD, was 5 or above

and 7 or above, accordingly. The recommended cut-off score of the GAD-2 was 2 or above for both Latvian and Russian versions.

There are several limitations in our study. First, there was a rather small sample size of GAD cases according to the MINI. Meanwhile, small sample size might reflect the differences in sensitivity and specificity compared with other studies. Second, the data demonstrated the prevalence of anxiety disorders and validity of the GAD-7 and GAD-2 for determining of GAD in a primary care population, which eliminates the potential to characterize individuals and use the GAD-7 and GAD-2 observed in specialized psychiatric outpatient departments, clinical settings, and the general population. However, our target population involved persons visiting primary care settings. The GAD-7 and GAD-2 consist of a self-report questionnaire. These screening instruments only provide a probable diagnosis of GAD that has to be investigated by further evaluation. Another limitation of our study is meaning of the LR; at a cut-off point of 5 or over for the GAD-7 Latvian version, LR+ of 2.42 and the LR– of 0.36 were found; in the Russian version, at a cut-off point of 7 or higher LR+ of 4.61 and LR– of 0.32 were found. The GAD-2 Latvian version, at a cut-off point of 2 or over, demonstrated LR of 2.17+ and LR– of 0.33, and the Russian version, at a cut-off point of 2 or over, had the LR+ of 2.70 and LR– of 0.24. These rates of the LR reflect rather small probability and sometimes useful test levels for all versions of the scales. The GAD-7 measures anxiety over the past 2 weeks, however, the MINI measures the GAD over the past 6 months. The difference in the observation period between the two instruments may affect probability of the usefulness of the GAD-7 and GAD-2. Additionally, one of the limitations is cross-sectional design of the study; there is a need for larger number of patients with GAD to improve the statistical significance of our findings, longitudinal studies are needed to establish the sensitivity to change. Future research should consider exploring psychometric properties using exploratory factor analysis and confirmatory factor analysis of the GAD-7 and the GAD-2 Latvian and Russian versions. Inclusion of currently diagnosed and treated patients may increase bias by inflating estimates of screening accuracy.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of Research in Riga Stradins University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JV, LR, AK-U, and ER conceived the presented idea and study design and analyzed the data. AK-U was responsible for the statistical data analysis. JV and LR wrote the first version of the manuscript. All authors participated in interpreting the data and developing further stages and the final version of the paper. All authors contributed to the article and approved the submitted version.

Funding

The study was supported by The National Research Programme BIOMEDICINE 2014–2017 (Nr. 5.8.1.).

Acknowledgments

We are grateful to all patients who participated in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Yang X, Fang Y, Chen H, Zhang T, Yin X, Man J, et al. Global, regional and national burden of anxiety disorders from 1990 to 2019: results from the Global Burden of Disease Study 2019. *Epidemiol Psych Sci.* (2021) 30:e36-e. doi: 10.1017/S2045796021000275

2. Casey BJ, Lee FS. Optimizing treatments for anxiety by age and genetics. *Ann N Y Acad Sci.* (2015) 1345:16–24. doi: 10.1111/nyas.12746

3. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. *Dialog Clin Neurosci.* (2015) 17:327–35. doi: 10.31887/DCNS.2015.17.3/bbandelow

4. Bosman RC, Ten Have M, de Graaf R, Muntingh AD, van Balkom AJ, Batelaan NM. Prevalence and course of subthreshold anxiety disorder in the general population: a three-year follow-up study. *J Affect Disord.* (2019) 247:105–13. doi: 10.1016/j.jad.2019.01.018

5. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med.* (2007) 146:317–25. doi: 10.7326/0003-4819-146-5-200703060-00004

6. Martín-Merino E, Ruigómez A, Wallander MA, Johansson S, García-Rodríguez LA. Prevalence, incidence, morbidity and treatment patterns in a cohort of patients diagnosed with anxiety in UK primary care. *Fam Pract.* (2010) 27:9–16. doi: 10.1093/fampra/cmp071

7. Rancans E, Renemane L, Kivite-Urtane A, Ziedonis D. Prevalence and associated factors of mental disorders in the nationwide primary care population in Latvia: a cross-sectional study. *Ann Gen Psychiatry.* (2020) 19:25. doi: 10.1186/s12991-020-00276-5

8. Somers JM, Goldner EM, Waraich P, Hsu L. Prevalence and incidence studies of anxiety disorders: a systematic review of the literature. *Can J Psychiatry revue Canadienne de Psychiatrie.* (2006) 51:100–13. doi: 10.1177/070674370605100206

9. Wittchen HU. Generalized anxiety disorder: prevalence, burden, and cost to society. *Depress Anxiety.* (2002) 16:162–71. doi: 10.1002/da.10065

10. Kasper S. Anxiety disorders: under-diagnosed and insufficiently treated. *Int J Psychiatry Clin Pract.* (2006) 10 Suppl 1:3–9. doi: 10.1080/13651500600552297

11. Santomauro DF, Mantilla Herrera AM, Shadid J, Zheng P, Ashbaugh C, Pigott DM, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet.* (2021) 398:1700–12. doi: 10.1016/S0140-6736(21)02143-7

12. Santabárbara J, Lasheras I, Lipnicki DM, Bueno-Notivol J, Pérez-Moreno M, López-Antón R, et al. Prevalence of anxiety in the COVID-19 pandemic: an updated meta-analysis of community-based studies. *Prog Neuropsychopharmacol Biol Psych.* (2021) 109:110207. doi: 10.1016/j.pnpbp.2020.110207

13. Jakovljevic M, Bjedov S, Jaksic N, Jakovljevic I. COVID-19 Pandemia and public and global mental health from the perspective of global health securit. *Psychiatr Danub.* (2020) 32:6–14. doi: 10.24869/psyd.2020.6

14. De La Vega D, Giner L, Courtet P. Suicidality in subjects with anxiety or obsessive-compulsive and related disorders: recent advances. *Curr Psychiatry Rep.* (2018) 20:26. doi: 10.1007/s11920-018-0885-z

15. Wittchen HU, Jacobi F. Size and burden of mental disorders in Europe–a critical review and appraisal of 27 studies. *Eur Neuropsychopharmacol J Eur Coll Neuropsychopharmacol.* (2005) 15:357–76. doi: 10.1016/j.euroneuro.2005.04.012

16. Sapra A, Bhandari P, Sharma S, Chanpura T, Lopp L. Using Generalized Anxiety Disorder-2 (GAD-2) and GAD-7 in a Primary Care Setting. *Cureus.* (2020) 12:e8224-e. doi: 10.7759/cureus.8224

17. Šica K, Pulmanis, Taube M. *Slimibu profilakses un kontroles centrs; Psihiskā veseliba Latvijā 2016. gadā. Tematiskais zinojums. Center for Disease Prevention and Control; Mental health in Latvia in 2016*. Thematic report. (2017). Available online at: https://www.spkc.gov.lv/lv/media/2722/download (accessed September 2022).

18. Vinogradova VV, Kivite-Urtane A, Vrublevska J, Rancans E. Anxiety screening among the general population of latvia and associated factors. *Medicina.* (2022) 58:1163. doi: 10.3390/medicina58091163

19. NICE. Common mental health problems: identification and pathways to care. Clinical guideline [CG123] (https://www.nice.org.uk/guidance/cg123/resources) 2011

20. Mughal AY, Devadas J, Ardman E, Levis B, Go VF, Gaynes BN, et al. systematic review of validated screening tools for anxiety disorders and PTSD in low to middle income countries. *BMC Psychiatry.* (2020) 20:338. doi: 10.1186/s12888-020-02753-3

21. Spitzer RL, Kroenke K, Williams JB, Löwe B, A. brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* (2006) 166:1092–7. doi: 10.1001/archinte.166.10.1092

22. CSBRoL. Available online at: https://www.csp.gov.lv/lv/demografija (accessed April 2022).

23. Rancans E, Trapencieris M, Ivanovs R, Vrublevska J. Validity of the PHQ-9 and PHQ-2 to screen for depression in nationwide primary care population in Latvia. *Ann Gen Psychiatry.* (2018) 17:33. doi: 10.1186/s12991-018-0203-5

24. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry.* 1998;59 Suppl 20:22-33;quiz 4–57.

25. Pettersson A, Modin S, Wahlström R, Af Winklerfelt Hammarberg S, Krakau I. The Mini-International Neuropsychiatric Interview is useful and well accepted as part of the clinical assessment for depression and anxiety in primary care: a mixed-methods study. *BMC Fam Pract.* (2018) 19:19. doi: 10.1186/s12875-017-0674-5

26. Vrublevska J, Trapencieris M, Snikere S, Grinberga D, Velika B, Pudule I, et al. The 12-month prevalence of depression and health care utilization in the general population of Latvia. *J Affect Disord.* (2017) 210:204–10. doi: 10.1016/j.jad.2016.12.031

27. Boudrot A, Sheehan D, Acquadro C. Lost in Translation: Translatability of Psychiatric Terms – The Example of the Mini– International Neuropsychiatric Interview (M. *INI) Value Health.* (2013) 16:A599. doi: 10.1016/j.jval.2013.08.1692

28. Duburcq A, Blin P, Charpak Y, Blachier C, Allicar MP, Bouhassira M, et al. Use of a structured diagnostic interview to identify depressive episodes in an epidemiologic study: a posteriori internal validation. *Revue d'epidemiologie et de sante publique.* (1999) 47:455–63.

29. Celentano DD, Szklo M. Gordis Epidemiology 6th ed: Saunders; 2019. p. 94-122.

30. Alkhadhari S, Alsabrrie AO, Ohaeri JU, Varghese R, Zahid MA, Mulsant BH. Mental and physical comorbidity in an Arab primary health care setting. BMC Psychiatry. 2018;18:313. doi: 10.1186/s12888-018-1903-8

31. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Med Care.* (2008) 46:266–74. doi: 10.1097/MLR.0b013e318160d093

32. Chibanda D, Verhey R, Gibson LJ, Munetsi E, Machando D, Rusakaniko S, et al. Validation of screening tools for depression and anxiety disorders in a primary care population with high HIV prevalence in Zimbabwe. *J Affect Disord.* (2016) 198:50–5. doi: 10.1016/j.jad.2016.03.006

33. Gong Y, Zhou H, Zhang Y, Zhu X, Wang X, Shen B, et al. Validation of the 7-item Generalized Anxiety Disorder scale (GAD-7) as a screening tool for anxiety among pregnant Chinese women. *J Affect Disord.* (2021) 282:98–103. doi: 10.1016/j.jad.2020.12.129

34. Kujanpää T, Ylisaukko-Oja T, Jokelainen J, Hirsikangas S, Kanste O, Kyngäs H, et al. Prevalence of anxiety disorders among Finnish primary care high utilizers and validation of Finnish translation of GAD-7 and GAD-2 screening tools. *Scand J Prim Health Care.* (2014) 32:78–83. doi: 10.3109/02813432.2014.920597

35. Seo JG, Park SP. Validation of the Generalized Anxiety Disorder-7 (GAD-7) and GAD-2 in patients with migraine. *J Headache Pain.* (2015) 16:97. doi: 10.1186/s10194-015-0583-8

36. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. *Gen Hosp Psychiatry.* (2016) 39:24–31. doi: 10.1016/j.genhosppsych.2015.11.005

37. García-Campayo J, Zamorano E, Ruiz MA, Pardo A, Pérez-Páramo M, López-Gómez V, et al. Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. *Health Qual Life Outcomes.* (2010) 8:8. doi: 10.1186/1477-7525-8-8

38. Donker T, van Straten A, Marks I, Cuijpers P. Quick and easy self-rating of Generalized Anxiety Disorder: validity of the Dutch web-based GAD-7, GAD-2 and GAD-SI. *Psychiatry Res.* (2011) 188:58–64. doi: 10.1016/j.psychres.2011.01.016

39. Asnaani A, Richey JA, Dimaite R, Hinton DE, Hofmann SG, A. Cross-ethnic comparison of lifetime prevalence rates of anxiety disorders. *J Nerv Ment Dis.* (2010) 198:551–5. doi: 10.1097/NMD.0b013e3181ea169f

40. Parkerson HA, Thibodeau MA, Brandt CP, Zvolensky MJ, Asmundson GJG. Cultural-based biases of the GAD-7. *J Anxiety Disord.* (2015) 31:38–42. doi: 10.1016/j.janxdis.2015.01.005

41. National Health Service Republic of Latvia. Available online at: https://www.vmnvd.gov.lv/lv/gimenes-arstiem-0 (accessed June 2022. 2020).

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 10 October 2022
DOI 10.3389/fpsyt.2022.973134

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University, Germany

REVIEWED BY
Alejandra Dominguez Espinosa,
Ibero American University, Mexico
Marinés Mejía Alvarez,
AIGLÉ Foundation, Argentina

*CORRESPONDENCE
Filiberto Toledano-Toledano
filiberto.toledano.phd@gmail.com

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 19 June 2022
ACCEPTED 09 September 2022
PUBLISHED 10 October 2022

CITATION
Astudillo-García CI, Austria-Corrales F, Rivera-Rivera L, Reynales-Shigematsu LM, Gómez-García JA, Séris-Martinez M, Jiménez-Tapia A, Robles R, Morales-Chainé S, López-Montoya A, Cuevas-Renaud C and Toledano-Toledano F (2022) Measurement invariance of the GAD-5 Generalized Anxiety Disorder Scale in a Mexican general population sample. *Front. Psychiatry* 13:973134. doi: 10.3389/fpsyt.2022.973134

Measurement invariance of the GAD-5 Generalized Anxiety Disorder Scale in a Mexican general population sample

Claudia I. Astudillo-García[1†], Fernando Austria-Corrales[2†], Leonor Rivera-Rivera[3], Luz Myriam Reynales-Shigematsu[3], José Alberto Gómez-García[4], Marina Séris-Martinez[3], Alberto Jiménez-Tapia[5], Rebeca Robles[5], Silvia Morales-Chainé[6], Alejandra López-Montoya[6], Corina Cuevas-Renaud[6] and Filiberto Toledano-Toledano[7,8,9*]

[1]Servicios de Atención Psiquiátrica (SAP), Secretaría de Salud, Ciudad de México, México, [2]Comisión Nacional para la Mejora Continua de la Educación (MEJOREDU), Ciudad de México, México, [3]Centro de Investigación en Salud Poblacional, Instituto Nacional de Salud Pública (INSP), Cuernavaca, Morelos, México, [4]Secretariado Técnico del Consejo Nacional de Salud Mental (STCONSAME), Secretaría de Salud, Ciudad de México, México, [5]Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz (INPRFM), Ciudad de México, México, [6]Facultad de Psicología, Universidad Nacional Autónoma de México, Ciudad de México, México, [7]Unidad de Investigación en Medicina Basada en Evidencias, Hospital Infantil de México Federico Gómez,Instituto Nacional de Salud, Ciudad de México, Mexico, [8]Unidad de Investigación Sociomédica, Instituto Nacional de Rehabilitación Luis Guillermo Ibarra Ibarra, Ciudad de México, México, [9]Dirección de Investigación y Diseminación del Conocimiento, Instituto Nacional de Ciencias e Innovación para la Formación de Comunidad Científica, INDEHUS, Ciudad de México, México

The primary objective of this study was to evaluate the measurement of invariance by sex, age, and educational level of an online version of the Generalized Anxiety Disorder Scale in a five-item version (GAD-5). Configural, metric, scalar, and strict invariance were evaluated using data from 79,473 respondents who answered a mental health questionnaire during the COVID-19 pandemic in Mexico. The sex variable was classified as male or female; age was categorized as minors, youth, young adults, adults, and older adults; and educational level was divided into basic, upper secondary, higher, and graduate education. To test for configural invariance, confirmatory factor models were constructed. For metric invariance, equality restrictions were established for the factor loadings between the construct and its items; for scalar invariance, equality restrictions were established between the intercepts; strict variance implied the additional restriction of the residuals. Statistical analysis was performed in R software with the lavaan package. The results show that with respect to sex, age, and educational level, configural and metric measurement invariance was confirmed (ΔCFI < 0.002; ΔRMSEA < 0.015). However, with respect to scalar and strict invariance, the results showed significant differences regarding the fit model (ΔCFI > 0.002; ΔRMSEA > 0.015). We conclude that the GAD-5 presents configural and metric invariance for sex, age, and educational level, and scalar invariance for sex and age groups. However, the scale does not demonstrate strict invariance. We discuss the implications and suggest that this result could be related to the evaluation of sociodemographic variables.

KEYWORDS

anxiety, Generalized Anxiety Disorder Scale (GAD), measurement invariance, multiple-group analysis, factor analysis, statistical, mass screening

Introduction

Anxiety disorders account for a large proportion of the global burden of disease and disability. A systematic review published in 2022 (1) reported that 301.4 million people worldwide had some type of anxiety disorder, with an age-standardized prevalence rate of 3779.5 (3181.1–4473.3) per 1,00,000 population. However, in Latin America and the Caribbean, this rate is 5502.3 (4625.9–6588.7). The global prevalence of generalized anxiety disorder (GAD) was 4.5% in 2021; although a higher prevalence has been reported in high-income countries (5.3%) than in low-income countries (2.8%), the proportion of people who have received treatment is lower in the latter (19.2 vs. 38.4%) (2). In low- and middle-income countries, most people with these disorders will never see a mental health specialist (3). It has also been reported that subthreshold anxiety disorders may have twice the frequency of the full syndrome, and are more persistent, cause greater suffering and functional impairment, and have a higher risk of onset and aggravation of other mental health conditions, such as pain and comorbid somatic disorders, increasing care costs (4).

The existing differences by sex and age must be added to this care gap. Women present greater anxiety than men. According to the 2022 GBD review, 187.5 million women suffer from anxiety disorders vs. 109.3 million men, in addition to the fact that the number of disability-adjusted life years (DALYs) increases steadily during childhood and adolescence, reaching a maximum between the ages of 25 and 34 and decreasing steadily after the age of 35 (1). In contexts such as the COVID-19 pandemic, evidence shows that there are significant differences by sex and age, with women and younger people scoring significantly higher in anxiety, and these differences are present also by educational level (5). In order to make judgments across conditions of age, sex, or educational level, scales are needed that operate equivalently for these different groups of interest (6), and that are available in non-specialized care settings.

Primary care is the ideal setting for the identification and appropriate treatment of the most common mental disorders. Screening for their early detection and treatment in primary care can improve quality of life, help contain health care costs, and limit complications from medical and mental health comorbidities (7). The application of screening scales is a useful alternative in primary care in low- and high-income countries, given existing time and resource pressures (8). These scales have the potential to improve case detection through procedures that could be incorporated into primary care practice. They direct attention to anxiety symptoms, and help to determine the current status of the individual and offer a specific diagnosis and treatment (8). Population-based screening requires that such tools have psychometric properties that allow for valid comparisons.

The factorial invariance of a scale is the statistical property that indicates whether it measures the same latent construct among the subgroups of a sample, which is a prerequisite for making valid group comparisons. The presence of non-variance could be indicative of bias due to differences in the interpretation of the items included in a scale (9). To determine whether a measure presents factorial invariance, factor loadings, intercepts, and residual variances are tested to ensure that they are equivalent in a factorial model that evaluates a latent concept. To this end, a set of increasingly restricted structural equation models are run to test whether differences between these models are significant (10). Failure to test for invariance means that different groups or subjects may respond differently to the items and that factor means cannot be reasonably compared (10).

The GAD Scale was developed as a screening tool for primary care settings (11). Its initial version consisted of nine items reflecting all of the DSM-IV diagnostic criteria for the disorder, as well as four items based on a review of existing anxiety scales (11). A seven-item version (GAD-7) has reported good to excellent sensitivity and specificity for most of the relevant DSM-5 disorders (5) in both the general population and in primary care patients (12). Measures of invariance have been reported for the GAD-7 (6, 9, 13), but not for the GAD-5, a five-item version obtained from studies of the primary care population (3, 8). The five items are directly linked to the ICD-11 diagnostic guidelines for depression and anxiety, in which a total score of 3 or more predicted 89.6% of above-threshold cases with generalized anxiety (11). This brief assessment of anxiety minimizes the time required in the patient encounter and obviates the need for paper and pencil tests and instrument scoring (3). It therefore offers a substantially more practical alternative for implementation in low-resource settings, and it may also be of considerable value in high-income countries (3).

The confirmation of parameter invariance helps to verify that the items and measures are free of biases that produce differences, which could be the result of differences in age, gender, and educational level. For example, the use of certain words may create a difference between those who fully understand an item and those who do not. In addition, gender bias in the wording of items can generate systematic error variances that may affect measurement precision. Confirming the invariance of parameters across different ages, sexes, and educational levels will help to understand whether the five attributes measured by the GAD-5 are relatively constant across groups and whether the groups analyzed share the same metric: whether the construct being measured is equivalent across groups (14). The aim of this study was thus to assess measurement invariance through the estimates of configural, metric, scalar, and strict invariance of the five-item version of the Generalized Anxiety Disorder Scale (GAD-5), across sex, age group, and educational level.

Materials and methods

Participants and procedure

We used a convenience sampling strategy to recruit 79,473 people who were analyzed for this study. Participants answered the GAD-5 questionnaire from April 1 to December 31, 2020, as part of the survey Atención Psicológica a Distancia para la Salud Mental por la contingencia por COVID-19 (Remote Mental Health Care during the COVID-19 Pandemic). This survey was part of the Mexican effort, led by the Secretary of Health, the Universidad Nacional Autónoma de México (UNAM), the Instituto Nacional de Psiquiatría, and civil society organizations to meet the mental health needs of the population and reduce the stress caused by the pandemic. The survey was administered by a team from the UNAM Faculty of Psychology through the federal government's coronavirus.gob.mx website. On this website, people were invited to participate voluntarily and confidentially and offered care resources according to the risk levels detected for different mental health problems. The questionnaire was self-administered online. A description of the survey and the variables assessed is available in a previous publication (15).

Study variables

The sociodemographic variables considered were sex, age group, and educational level. Sex was classified as male or female. Age was categorized as minors (13–17 years), youth (18–25 years), young adults (26–35 years), adults (36–59 years), and older adults (60 years and older). Educational level was divided into basic (elementary and junior high school), upper secondary (high school or equivalent), higher education (undergraduate degree) and graduate (specialty, master's, and doctoral degrees). The age categories are consistent with Mexican law that considers adulthood to begin at age 18 and senior citizens to be those over 60. The intermediate ages were divided into three groups that represent the life trajectories of adults in Mexico. However, it should be noted that the complexity of life trajectories makes it difficult to construct a universal division of different life stages (16). The categories of educational level were based on the organization of the educational system in Mexico, which includes basic (elementary and junior high school), middle (high school), and higher education (university); the latter was divided into separate categories for undergraduate and graduate education.

The GAD-5 consists of five items: "I feel nervous, anxious, or about to burst," "I have felt unable to control my worries," "I have felt so worried, I have been unable to keep still," "I have found it hard to relax," and "I have felt afraid that something terrible was going to happen." Participants were asked to what extent each of these items described them in the past 2 weeks. The standard response form was modified to match the rest of the instruments used in order to avoid having to provide different instructions and response options for each part of the questionnaire. The response options for the entire survey were a 10-point Likert scale, where 0 indicated "does not describe me" and 10 "describes me exactly." With five items, the range of possible scores was thus 0–50 points. There is evidence suggesting that increasing the number of response options increases validity coefficients by 0.04 (17). This evidence also suggests that the coefficients do not rise artificially as the number of response options increases; however, the validity does consistently improve. In another study, Alwin (18) conducted a confirmatory factor analysis to compare the performance of the versions with seven and eleven response options and found that the latter had better validity and reliability and lower invalidity indices.

Data analysis

We first performed a confirmatory factor analysis (CFA), using R software and the lavaan package (19), to test the theoretical structure of the scale as well as its unidimensionality. The covariance matrix was analyzed using the maximum likelihood method, applying the Satorra–Bentler correction (20), since the data do not assume multivariate normality. The fit of the model was assessed with four fit indices. The comparative fit index (CFI) takes possible values between 0 and 1, with a value of at least 0.90 denoting adequate fit and a value greater than or equal to 0.95 a very good fit. The Tucker-Lewis index (TLI) also has a range from 0 to 1 with the same interpretation criteria. The Root Mean Square Error Approximation (RMSEA) should ideally have values of <0.06, although values of 0.08 are considered acceptable. Finally, the Standardized Root Mean Square Residual (SRMR) is considered acceptable with a value <0.10 and a good fit with a value <0.05 (21).

We next assessed measurement invariance using multi-group confirmatory factor analysis; this technique makes it possible to gradually impose restrictions in order to test different levels of parameter invariance: configural, metric, scalar, and strict. The first step was to test configural invariance; this model was used as a baseline for comparison with models that gradually incorporated more equality constraints. To assess configural invariance, it was necessary to keep the factor loading structure constant between the different comparison groups, although the values of the loadings, factor variances, and covariances could vary because they were not restricted to being equal. Metric invariance was subsequently determined by establishing equality restrictions on the values of the factor loadings. We then proceeded to test scalar invariance through the establishment of equality restrictions between the intercepts, and finally strict invariance, where equality was also restricted among residuals. We evaluated changes in the comparative fit index (CFI) to assess the measurement invariance between the different groups:

TABLE 1 Sociodemographic characteristics of the sample.

Variable	*n* (%)
Sex	
Women	48,308 (60.79)
Men	31,165 (39.21)
Age group	
Minors	6,392 (8.04)
Youth	14,967 (18.83)
Young adults	22,267 (28.02)
Adults	32,760 (41.22)
Older adults	3,087 (3.89)
Educational level	
Basic education	11,703 (14.73)
Upper secondary education	23,444 (29.50)
Higher education	35,318 (44.44)
Graduate	9,008 (11.33)

TABLE 2 Factor loads of GAD-5 items.

Item	Standardized coefficient (β)*
I feel nervous, anxious, or about to burst	0.910
I have felt unable to control my worries	0.919
I have felt so worried I have been unable to keep still	0.865
I have found it hard to relax	0.899
I have felt afraid that something terrible was going to happen	0.823

*All values are significant, $p < 0.001$.

a change in CFI of −0.01 or more from the baseline was used to reject the between-group invariance hypothesis (22). We also evaluated ΔSRMR and ΔRMSEA as alternative fit indices, as suggested by Chen (23).

Results

Data were analyzed from 79,473 people who participated voluntarily and answered the questionnaire. The sample included 60.79% women and 39.21% men, with an average age of 35.11 years ($SD = 12.74$). The distribution by age group and educational level is shown in Table 1.

GAD-5 factor analysis

The resulting model showed an adequate fit between the theoretical model and the empirical data, as shown by the following fit indices: CFI = 0.993; TLI = 0.987; RMSEA = 0.07, CI [0.075, 0.081]; SMRM = 0.009. Table 2 shows the factor loads of items in the GAD-5. The resulting model, as well as the standardized parameters, can be seen in Figure 1.

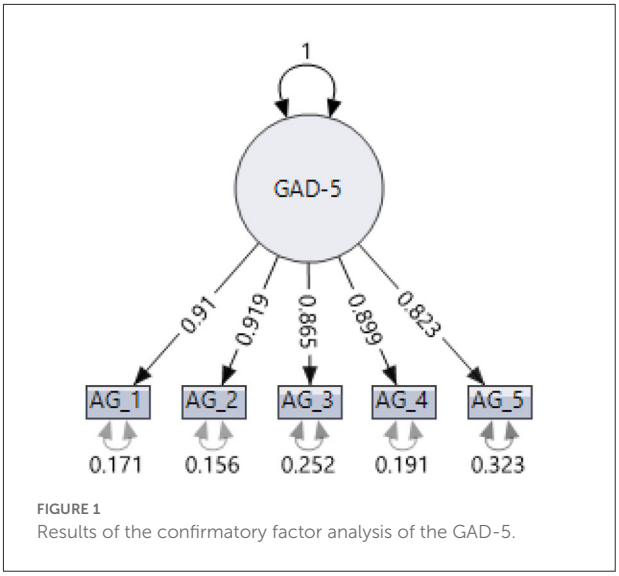


FIGURE 1
Results of the confirmatory factor analysis of the GAD-5.

Multi-group CFA and measurement invariance

Once the unidimensionality of the GAD-5 and its parametric stability were demonstrated, variances were divided by sex, age group, and educational level, according to the categorizations described above. Equality restrictions were then gradually imposed, using the configural model as the baseline.

As regards invariance by sex, the configural invariance showed a good fit with respect to the general model, indicating a lack of significant differences in the factorial structure between women and men. When equality restrictions were placed on the factor loadings (metric invariance), no differences were observed in the comparative fit index (ΔCFI = 0.000). This evidence suggests that the GAD-5 is metrically invariant by sex. Equality restrictions were then imposed on the intercepts (scalar invariance), reducing the ΔCFI by −0.001, suggesting a lack of significant differences. Finally, after imposing equality restrictions on residuals (strict invariance), a change of −0.008 was observed in the ΔCFI, a value of <0.01, the traditional criterion for assessing the invariance of parameters. As regards age, five groups were compared: minors, youth, young adults, adults, and older adults. Table 3 shows that differences in the ΔCFI in the metric, scalar, and strict invariance are in all cases less than the criteria established by Cheung and Rensvold (22), suggesting that the GAD-5 is invariant at the configural, metric, scalar, and strict levels. In relation to educational level, we observed that changes in the ΔCFI in metric, scalar, and strict invariance do not

TABLE 3 Results of tests of measurement invariance.

Model	X^2	*df*	CFI	TLI	SRMR	RMSEA	Model comparison	ΔCFI
By sex								
Configural	997.856	10	0.993	0.987	0.010	0.079 (0.074-0.083)	-	-
Metric	1,223.088	14	0.993	0.990	0.013	0.068 (0.065-0.071)	Configural – Metric	0.000
Scalar	1,659.291	18	0.992	0.991	0.015	0.065 (0.063-0.068)	Metric – Scalar	−0.001
Strict	3,180.108	23	0.984	0.986	0.020	0.082 (0.080-0.085)	Scalar – Strict	−0.008
By age								
Configural	1,126.013	25	0.993	0.986	0.010	0.082 (0.078-0.086)	-	-
Metric	1,646.677	41	0.992	0.990	0.020	0.069 (0.066-0.072)	Configural – Metric	−0.001
Scalar	2,190.414	57	0.991	0.992	0.022	0.063 (0.063-0.065)	Metric – Scalar	−0.001
Strict	3,165.974	77	0.983	0.989	0.023	0.073 (0.071-0.075)	Scalar – Strict	−0.008
By educational level								
Configural	959.343	20	0.994	0.987	0.009	0.077 (0.073-0.081)	-	-
Metric	1,194.531	32	0.994	0.992	0.011	0.061 (0.058-0.064)	Configural – Metric	0.000
Scalar	2,054.791	44	0.991	0.991	0.017	0.063 (0.061-0.066)	Metric – Scalar	−0.003
Strict	2,713.069	59	0.986	0.990	0.019	0.067 (0.065-0.069)	Scalar – Strict	−0.005

exceed the −0.01 criterion, suggesting that the GAD-5 is invariant across educational levels. The results are shown in Table 3.

To confirm these results based on the traditional criteria for assessing the invariance of parameters, the change in CFI (ΔCFI), additional assessments were made using two alternative indices suggested by Chen (23): changes in the RMSEA of 0.015 and the SMRM of 0.030 for metric invariance, and changes in the scalar and strict invariance of 0.015. The results are summarized in Table 4 for each of the comparison variables: sex, age group, and educational level.

The results by sex and age showed that ΔSRMR and ΔRMSEA have values of <0.030 and 0.015 respectively in assuming metric and scalar invariance, suggesting that these invariances might be present, but not strict invariance. However, the values observed for ΔRMSEA indicate significant differences in the model, so this possibility is not empirically supported. As for educational level, there is only metric, not scalar or strict invariance, since the ΔRMSEA value is −0.015.

Taken together, these findings suggest that the GAD-5 has psychometric properties that provide invariant measurements for the sociodemographic characteristics of sex, age, and educational level. However, the invariance is not complete in all cases. The traditional ΔCFI and alternative indexes of ΔSRMR and ΔRMSEA coincide to show the following: (a) by sex, GAD-5 has configural, metric, and scalar invariance; (b) by age group, it has configural, metric, and scalar invariance; and (c) by educational level, it has configural and metric invariance.

Discussion

Using data drawn from a large Mexican general population sample, we assessed measurement invariance of the GAD-5 by sex, age, and educational level. Our findings indicate that the GAD-5 conforms to the proposed theoretical structure, since a unidimensional construct of generalized anxiety symptomatology was obtained, which presented configural and metric invariance in the comparison by sex, age, and educational

TABLE 4 Alternative fit indices to evaluate measurement invariance by sex, age, and education.

Model	SRMR	RMSEA	Model comparison	ΔCFI	ΔSRMR	Δ RMSEA
By sex						
Configural	0.010	0.079 (0.074–0.083)	-	-	-	-
Metric	0.013	0.068 (0.065–0.071)	Configural – Metric	−0.001	0.004	−0.011
Scalar	0.015	0.065 (0.063–0.068)	Metric – Scalar	−0.001	0.002	−0.002
Strict	0.020	0.082 (0.080–0.085)	Scalar – Strict	−0.008	0.004	−0.017
By age						
Configural	0.010	0.082 (0.078–0.086)	-	-	-	-
Metric	0.020	0.069 (0.066–0.072)	Configural – Metric	0.000	0.011	−0.013
Scalar	0.022	0.063 (0.063–0.065)	Metric – Scalar	−0.001	0.001	−0.006
Strict	0.020	0.082 (0.080–0.085)	Scalar – Strict	−0.008	0.002	0.010
By educational level						
Configural	0.009	0.077 (0.073–0.081)	-	-	-	-
Metric	0.011	0.061 (0.058–0.064)	Configural – Metric	0.000	0.002	−0.015
Scalar	0.017	0.063 (0.061–0.066)	Metric – Scalar	−0.003	0.006	0.002
Strict	0.019	0.067 (0.065–0.069)	Scalar – Strict	−0.005	0.003	0.004

level, and scalar invariance in the comparison by sex and age. This provides evidence that the use of the GAD-5 as a screening instrument in the general population allows for adequate comparisons between men and women and between age groups.

The results of the measures of configural, metric, and scalar invariance, both by sex and by age group, show that the construct (factor loadings) and the levels of the underlying items (intercepts) are equal in all the groups tested. Accordingly, these groups attribute the same meaning to the latent construct studied, and their scores on the latent variable can be compared. Although strict variance was not achieved, indicating that the explained error variances are not equal in all groups, they can still be compared with respect to the latent variable. It should be noted that the latent variable is measured with different degrees of error between groups (10). However, provided that at least two loadings and intercepts are the same across groups, valid inferences can be made about the differences between the means of the latent factors in the model (10).

Since there is still a significant debate concerning the fit indices to be used to assess parameter invariance, this study used traditional indices (ΔCFI) and alternative indices that have been proposed in recent years (ΔSRMR and Δ RMSEA) to obtain additional evidence. It was therefore possible to observe that some scalar invariance hypotheses were rejected when more than one fit index was compared. Likewise, we should note that the confirmation of certain measurement invariance hypotheses does not mean there are no variations between the attributes of the different groups under comparison. What it means is that the instrument is able to efficiently measure, and with less error, between the different groups, without affecting the measurements, which increases the internal validity of the inferences that can be drawn. The results showed, for example, that the hypotheses of configural invariance and metric invariance are sustained across educational levels, whereas the scalar and strict hypotheses are rejected. This evidence suggests that the anxiety characteristics measured by the GAD-5 are present at all four educational levels (configural invariance) and that the metric for measuring anxiety in each level is identical (metric invariance). However, the latent averages (intercepts) obtained from the measurements between the different levels

vary significantly, as does the degree of error in the estimation process (residuals).

At the same time, it is important to recognize that although measures of configural, metric, scalar, and strict invariance are enormously useful in the construction and evaluation of psychological theories, their validity and existence in the real world of psychological measurement and research can never be definitively established in practice: they remain more of an ideal (24). The challenge for researchers who allow for partial invariance (in other words, that evidence is not obtained for all types of invariance) is to determine how much non-invariance can be tolerated while still claiming to measure the same construct across groups: they must make a decision based on the anticipated threat to the validity of their findings in each course of action (25). Novel approaches have been proposed for the use of partial invariance analysis through simulations, and it has been suggested that these can outperform total and partial invariance approaches when there are many small differences in item parameters (26).

Despite these considerations, the GAD-5 is a useful alternative in the general population that can be used in primary care settings, like the GAD-7 (11, 12, 27–31), and during health emergencies such as the COVID-19 pandemic. In this respect, the GAD-5 offers the practicality of web-based application in addition to the novelty of the response format used. These features contribute to the current debate on how the number of response options affects the psychometric properties of Likert-type scales (32, 33): it has been reported that reliability increases and excessive interpolation is avoided when response options increase from five to seven (34, 35), a result that could be more evident in online surveys.

Finally, it is important to consider the need to identify anxiety-like symptomatology even if it has only been present for a short time, and the GAD-5 refers to the previous 2 weeks. Short periods of anxiety have been reported to be predictive of subsequent psychopathology and may present as much associated disability at 6-month follow-up as longer periods (3). Including these screening options in routine care settings could therefore be a highly effective preventive action for the detection of common mental disorders in primary care, and improve the level of detection and diagnosis of these disorders in public health systems (3, 36).

Limitations

Although our data represent a robust sample of the Mexican population, it should be noted that data collection was conducted entirely online, which may lead to participation as well as information bias. At the same time, by considering only the categories of male and female, we omitted transgender, nonbinary, and gender-diverse individuals, who experience more mental health issues than their cisgender peers, including higher rates of depression, suicide, violence, and drug use (37). By achieving parameter invariance in these groups, we could confirm whether variations are due to the level of anxiety presented by the person, irrespective of group membership. There is thus a need to obtain scientific evidence regarding this sexually diverse population to support its mental health by strengthening the competencies of health system professionals, and also for the formulation of public policy (38).

We must also recognize that cross-sectional measurement does not allow for the exploration of invariance over time, which is also important (39). To do so, it would be necessary to conduct follow-up measurements to assess long-term effects in the population, which was beyond the initial scope of the mental health strategy during the COVID-19 pandemic. Future studies could evaluate the partial invariance of the GAD-5 parameters at levels that could not be confirmed in this study, for example at the educational level, and for scalar and strict invariance in all cases. We also think it is important to evaluate other variables of interest, but given that our study was a secondary analysis, this was not possible. Finally, the absence of additional validation criteria and comparative studies of the validity and usefulness of the GAD-5 could also be considered a limitation of the study requiring future research.

Despite these limitations, our results show that the scale performs quite satisfactorily, and this allows us to make several observations. First, it is possible to use the scale without the need for any special adjustment or scoring to detect anxiety levels in the population, in contrast to other measures that are used indiscriminately without knowing their psychometric properties or whether they require specific scoring to accurately place examinees on a continuum. Second, the scale allows for comparisons between examinees, regardless of their age, educational level, or sex, since the data show invariance across these variables, facilitating direct comparisons without the need for linear transformations to compare populations. Third, the five attributes measured by the GAD-5 are sufficiently general as to be present in all of the groups compared, which in itself constitutes evidence of external validity.

Conclusion

The GAD-5 shows a unidimensional theoretical structure and configural, metric, and scalar invariance in its comparisons by sex and by age group, which supports its use as a screening instrument in the general population. Since it is a short, easily administered instrument, its use could make a crucial contribution to the identification and treatment of mental health

problems in both the general population and the primary care setting. This study adds to the growing evidence about the concise and simple GAD-7 questionnaire, demonstrating that its five-item version, the GAD-5, could facilitate its application in primary care settings. The brevity and predictive value of this scale suggest its potential value as an initial assessment tool for clinicians that facilitates timely intervention to treat these disorders.

Data availability statement

The data presented in this study are available on request from the corresponding author.

Ethics statement

The study was reviewed and approved by the Research Ethics Committee of the UNAM Psychology Faculty, registration number FPSI/422/CEIP/157/2020. All data are protected under the computer security standards of Mexican personal data protection laws. The participants provided their written informed consent to participate in this study.

Author contributions

CA-G and FA-C: conceptualization, methodology, and formal analysis. CA-G, FA-C, LR-R, LR-S, JG-G, MS-M, AJ-T, and FT-T: research. RR, SM-C, AL-M, and CC-R: data curation. CA-G, FA-C, and AJ-T : writing of original draft. All authors participated in the writing, review, editing, read and approved the final manuscript.

Funding

This work presents some results of the research project HIM/2015/017/SSA.1207 Effects of mindfulness training on psychological distress and quality of life of the family caregiver, Main researcher: Filiberto Toledano-Toledano, Ph.D. The present research was funded with federal funds for health research and was approved by the Commissions of Research, Ethics and Biosafety [Comisiones de Investigación, Ética y Bioseguridad], Hospital Infantil de México Federico Gómez, Instituto Nacional de Salud. The funding agency had no control over the design of the study; the collection, analysis and interpretation of the data; or the writing of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry*. (2022) 9:137–50. Available online at: https://linkinghub.elsevier.com/retrieve/pii/S2215036621003953. doi: 10.1016/S2215-0366(21)00395-3

2. Stein DJ, Kazdin AE, Ruscio AM, Chiu WT, Sampson NA, Ziobrowski HN, et al. Perceived helpfulness of treatment for generalized anxiety disorder: a World Mental Health Surveys report. *BMC Psychiatry*. (2021) 21:1–14. doi: 10.1186/s12888-021-03363-3

3. Goldberg DP, Reed GM, Robles R, Minhas F, Razzaque B, Fortes S, et al. Screening for anxiety, depression, and anxious depression in primary care: a field study for ICD-11 PHC. *J Affect Disord*. (2017) 213:199–206. doi: 10.1016/j.jad.2017.02.025

4. Haller H, Cramer H, Lauche R, Gass F, Dobos GJ. The prevalence and burden of subthreshold generalized anxiety disorder: a systematic review. *BMC Psychiatry*. (2014) 14:1–13. doi: 10.1186/1471-244X-14-128

5. Santabárbara J, Lasheras I, Lipnicki DM, Bueno-Notivol J, Pérez-Moreno M, López-Antón R, et al. Prevalence of anxiety in the COVID-19 pandemic: an updated meta-analysis of community-based studies. *Prog Neuropsychopharmacol Biol Psychiatry*. (2021)109:110207. doi: 10.1016/j.pnpbp.2020.110207

6. Shevlin M, Butter S, McBride O, Murphy J, Gibson-Miller J, Hartman TK, et al. Measurement invariance of the Patient Health Questionnaire (PHQ-9) and generalized Anxiety Disorder scale (GAD-7) across four European countries during the COVID-19 pandemic. *BMC Psychiatry*. (2022) 22:1–9. doi: 10.1186/s12888-022-03787-5

7. Mulvaney-Day N, Marshall T, Downey Piscopo K, Korsen N, Lynch S, Karnell LH, et al. Screening for behavioral health conditions in primary care settings: a systematic review of the literature. *J Gen Intern Med*. (2018) 33:335–46. doi: 10.1007/s11606-017-4181-0

8. Goldberg DP, Prisciandaro JJ, Williams P. The primary health care version of ICD-11: the detection of common mental disorders in general medical settings. *Gen Hosp Psychiatry*. (2012) 34:665–70. doi: 10.1016/j.genhosppsych.2012.06.006

9. Romano I, Ferro MA, Patte KA, Leatherdale ST. Measurement invariance of the GAD-7 and CESD-R-10 among adolescents in Canada. *J Pediatr Psychol*. (2021) 47:1–10. doi: 10.1093/jpepsy/jsab119

10. van de Schoot R, Lugtig P, Hox J, A. Checklist for testing measurement invariance. *Eur J Develop Psychol.* (2012) 9:486–92. doi: 10.1080/17405629.2012.686740

11. Spitzer RL, Kroenke K, Williams JBW, Löwe B, A. brief measure for assessing generalized anxiety disorder: The GAD-7. *Arch Intern Med.* (2006) 166:1092–7. doi: 10.1001/archinte.166.10.1092

12. Moreno E, Muñoz-Navarro R, Medrano LA, González-Blanch C, Ruiz-Rodríguez P, Limonero JT, et al. Factorial invariance of a computerized version of the GAD-7 across various demographic groups and over time in primary care patients. *J Affect Disord.* (2019) 252:114–21. doi: 10.1016/j.jad.2019.04.032

13. Teymoori A, Real R, Gorbunova A, Haghish EF, Andelic N, Wilson L, et al. Measurement invariance of assessments of depression (PHQ-9) and anxiety (GAD-7) across sex, strata and linguistic backgrounds in a European-wide sample of patients after Traumatic Brain Injury. *J Affect Disord.* (2020) 262:278–85. doi: 10.1016/j.jad.2019.10.035

14. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Develop Rev.* (2016) 41:71–90. doi: 10.1016/j.dr.2016.06.004

15. Morales Chainé S, López Montoya A, Bosch Maldonado A, Beristain Aguirre A, Robles García R, Garibay Rubio CR, et al. Mental health symptoms, binge drinking, and the experience of abuse during the COVID-19 lockdown in Mexico. *Front Public Health*. (2021) 9:979. doi: 10.3389/fpubh.2021.656036

16. CONAPO. *Índices de desarrollo social, 2000 [Internet]*. (2003). Available online at: http://www.conapo.gob.mx/work/models/CONAPO/Resource/1342/1/images/03Indicedesarrollo.pdf

17. Matell MS, Jacoby J. Is there an optimal number of alternatives for likert scale items? Study 1: Reliability and validity. *Educ Psychol Meas.* (1971) 31:657–74. doi: 10.1177/001316447103100307

18. Alwin DF. Feeling thermometers versus 7-point scales: Which are better? *Sociol Methods Res.* (1997) 25:318–40. doi: 10.1177/0049124197025003003

19. Rosseel Y. lavaan: An R package for structural equation modeling. *J Stat Softw*. (2012) 48:1–36. Available online at: http://www.jstatsoft.org/v48/i02/. doi: 10.18637/jss.v048.i02

20. Satorra A, Bentler PM, A. Scaled difference chi-square test statistic for moment structure analysis. *Psychometrika.* (2001) 66:507–14. doi: 10.1007/BF02296192

21. Kline RB. *Principles and practice of structural equation modeling (3rd ed)*. New York: The Guilford Press. (2011).

22. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Model.* (2002) 9:233–55. doi: 10.1207/S15328007SEM0902_5

23. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Modeling.* (2007) 14:464–504. doi: 10.1080/10705510701301834

24. Meredith W. Measurement invariance, factor analysis and factorial invariance. *Psychometrika.* (1993) 58:525–43. doi: 10.1007/BF02294825

25. Guenole N, Brown A. The consequences of ignoring measurement invariance for path coefficients in structural equation models. *Front Psychol.* (2014) 5:980 doi: 10.3389/fpsyg.2014.00980

26. van de Schoot R, Kluytmans A, Tummers L, Lugtig P, Hox J, Muthén B. Facing offwith Scylla and Charybdis: a comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Front Psychol.* (2013) 1–15. doi: 10.3389/fpsyg.2013.00770

27. Staples LG, Dear BF, Gandy M, Fogliati V, Fogliati R, Karin E, et al. Psychometric properties and clinical utility of brief measures of depression, anxiety, and general distress: the PHQ-2, GAD-2, and K-6. *Gen Hosp Psychiatry.* (2019) 56:13–8. doi: 10.1016/j.genhosppsych.2018.11.003

28. García-Campayo J, Zamorano E, Ruiz MA, Pérez-Páramo M, López-Gómez V, Rejas J. The assessment of generalized anxiety disorder: psychometric validation of the Spanish version of the self-administered GAD-2 scale in daily medical practice. *Health Qual Life Outcomes.* (2012) 10:1–10. doi: 10.1186/1477-7525-10-114

29. Moriana JA, Jurado-González FJ, García-Torres F, Contreras A, Muñoz-Navarro R, González-Blanch C, et al. Exploring the structure of the GAD-7 scale in primary care patients with emotional disorders: a network analysis approach. *J Clin Psychol.* (2022) 78:283–97. doi: 10.1002/jclp.23217

30. McDonald AJ, Wickens CM, Bondy SJ, Elton-Marshall T, Wells S, Nigatu YT, et al. Age differences in the association between loneliness and anxiety symptoms during the COVID-19 pandemic. *Psychiatry Res.* (2022) 310:114446. doi: 10.1016/j.psychres.2022.114446

31. Lin S. (Lamson). Generalized anxiety disorder during COVID-19 in Canada: gender-specific association of COVID-19 misinformation exposure, precarious employment, and health behavior change. *J Affect Disord.* (2022) 302:280–92. doi: 10.1016/j.jad.2022.01.100

32. Dawes J. Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *Int J Market Res.* (2008) 50:61–77. doi: 10.1177/147078530805000106

33. Lozano LM, García-Cueto E, Muñiz J. Effect of the number of response categories on the reliability and validity of rating scales. *Methodology.* (2008) 4:73–9. doi: 10.1027/1614-2241.4.2.73

34. Dillman DA, Smyth JD, Christian LM. *Internet, Mail and Mixed-Mode Surveys*. Hoboken, New Jersey: Wiley. (2014).

35. Finstad K. Response interpolation and scale sensitivity: evidence against 5-point scales. *J Usability Stud.* (2010) 5:104–10.

36. Mughal AY, Devadas J, Ardman E, Levis B, Levis B, Go VF, et al. A systematic review of validated screening tools for anxiety disorders and PTSD in low to middle income countries. *BMC Psychiatry BioMed Central.* (2020) 20:1–18. doi: 10.1186/s12888-020-02753-3

37. Newcomb ME, Hill R, Buehler K, Ryan DT, Whitton SW, Mustanski B. High burden of mental health problems, substance use, violence, and related psychosocial factors in transgender, non-binary, and gender diverse youth and young adults. *Arch Sex Behav.* (2020) 49:645–59. doi: 10.1007/s10508-019-01533-9

38. Dolotina B, Turban JL, A. Multipronged, evidence-based approach to improving mental health among transgender and gender-diverse youth. *JAMA Network Open Am Med Assoc.* (2022) 5:e220926. doi: 10.1001/jamanetworkopen.2022.0926

39. van de Schoot R, Schmidt P, de Beuckelaer A. Measurement invariance [Internet]. In van de Schoot R, Schmidt P, de Beuckelaer A, editors. *Frontiers Media S.A.* (2015). (Frontiers Research Topics). doi: 10.3389/978-2-88919-650-0

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 13 October 2022
DOI 10.3389/fpsyt.2022.897175

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University, Germany

REVIEWED BY
Rebecca Pearson,
Manchester Metropolitan University, United Kingdom
Grace Lim,
University of Pittsburgh, United States

*CORRESPONDENCE
Anh Phuong Hoang
anh.hp@vinuni.edu.vn

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 15 March 2022
ACCEPTED 02 September 2022
PUBLISHED 13 October 2022

CITATION
Nguyen HTT, Do LTK, Pham HTT, Hoang AP, Truong HT and Nguyen HTH (2022) The symptoms of postpartum depression observed by family members: A pilot study. *Front. Psychiatry* 13:897175. doi: 10.3389/fpsyt.2022.897175

The symptoms of postpartum depression observed by family members: A pilot study

Huong Thi Thanh Nguyen[1], Ly Thi Kim Do[2], Huong Thi Thu Pham[3,4,5], Anh Phuong Hoang[1,3]*, Hoa Thi Truong[1] and Huyen Thi Hoa Nguyen[1]

[1]College of Health Sciences, VinUniversity, Hanoi, Vietnam, [2]General Surgical Department, University Medical Center Schleswig-Holstein, Lübeck, Germany, [3]Faculty of Nursing and Midwifery, Hanoi Medical University, Hanoi, Vietnam, [4]School of Nursing, College of Medicine, National Taiwan University, Taipei, Taiwan, [5]National Institute of Mental Health, Bach Mai Hospital, Hanoi, Vietnam

Postpartum Depression (PPD) is a burden on women's mental health after delivery, predominantly occurring in the 1st year. PPD poses a threat to the mother's life and affects the quality of childcare. Early detection by family members of depressive symptoms is critical. This study aimed to examine the role of family members in reporting depressive symptoms of PPD among new mothers. A cross-sectional study was conducted, where 56 family members were asked to report depressive symptoms observed in new mothers. At the same time, the new mothers were also screened for PPD using the Edinburgh Postpartum Depression Scale (EPDS). Binary logistic regression was performed. Depressive symptoms of new mothers reported by family members, including emotional and behavioral disturbance, being under stress, high anxiety, isolation, changing lifestyle, and inability to take care of their children, were found as predictors of PPD.

KEYWORDS

symptoms, postpartum depression, role of family members, postpartum, family role attitudes

Background

Postpartum Depression (PPD) is a major maternal health problem. The prevalence of PPD ranges from 1.9 to 82.2% in developing countries and from 5.3 to 74% in developed countries (1). In Viet Nam, a study discovered that the prevalence of women suffering from PPD was 20.4% in urban areas and 15.8% in rural areas (2). PPD is associated with a reduction in women's physical and mental health, and depressed women experience a lower quality of life (3). If a mother is depressed, anxious, or stressed, her children are more likely to have a wide range of adverse outcomes, including emotional problems, attention deficit hyperactivity disorder (ADHD), or impaired cognitive development (4). Social and/or family support plays a significant role in detecting PPD. When women receive support and care from their husbands, the percentage of those with depression reduces significantly (5). However, limited studies are available on the role of family members in screening for PPD. The sooner the PPD is detected, the better the outcome is achieved. This study, therefore, aims to explore the feasibility of family members of women with PPD helping detect early symptoms of PPD.

Methods

This study is cross-sectional, piloted in a small group of families with new mothers whose child anywhere from birth to 1 year old. A convenience sampling method was used. The research team conducted home visits with new mothers whose children were under 1 year old to invite them to participate in this study. A brief introduction of this study and a consent form was sent out to all new mothers who gave birth within one previous year during the home visits in a commune of Thuong Tin district, Hanoi. In case mothers were not available at home during visiting time, the researchers would leave a message and return within that day. If the new mothers agreed to participate in the study, they were asked to complete the Edinburgh Postpartum Depression Scale (EPDS). The EPDS includes 10 items measured on a Likert scale of 0–3. A score of 12 and above indicates the risk of depression. The sensitivity and specificity of EPDS were 65–100% and 49–100% respectively (6). The validated Vietnamese version of EPDS is the most common screening tool for perinatal common mental disorders used in Vietnam with an internal consistency of 0.77 (7).

During home visits, researchers paid attention to the living environments and family members involved in caring for new mothers to identify the closest caregiver to that mother. The family members were asked to complete the second part of the questionnaire with 9 items to report any depressive symptoms they could observe in the new mothers. This part of the questionnaire was developed by the research team based on guidelines in DSM-V, including common items relating to the possibility of observing signs and symptoms of PPD among these new mothers (8). The evidenced-based items were developed based on DSM-V by the American Psychiatric Association and previous studies (5, 9, 10).

Although 116 families were reached out in total, only 56 families responded with both the answers of the new mothers and her family members.

SPSS software version 23.0 was employed for data management and analysis. Binary logistic regression was also performed to investigate the relationship between PPD among new mothers and depressive symptoms observed by family members. The proposal, including ethical considerations, was approved by Hanoi Medical University Research Proposal Committee, Decision No 5042/Q-HYHN.

Results

Fifty-six new mothers participated in the screening with EPDS, and 20 (35.7%) of them were detected to be at risk of PPD (with an EPDS score of 12 and above). A family member of these mothers was also invited to an interview to report depressive symptoms observed from the mothers. Of the 56 family members, husbands were the majority ($n = 29$, accounting for 51.8%), followed by relatives and close friends ($n = 17$, accounting for 30.4%). There were only 5 maternal parents (8.9%) and 5 parents-in-law (8.9%) that participated in the interview (Table 1).

TABLE 1 The percentage of family members participating in the study.

Family member	Number of participants (n)	Percentage (%)
Husbands	29	51.8
Maternal parents	5	8.9
Parents in law	5	8.9
Relatives and close friends	17	30.4
Total	56	100

The relationships between observed depressive symptoms by family members and the risk of PPD among new mothers were reported. The results (Table 2) indicated that many symptoms and signs of PPD could be observed by family members. "Being sad unreasonably" was the most prevalent sign of PPD. As shown below, new mothers whom family members observed with the symptom "being sad unreasonably" experienced the risk of PPD (EPDS $\geq$ 12) 11.3 times higher than that of a new mother without this symptom (OR = 11.3; $p < 0.05$). Other significant depressive symptoms/signs observed by family members included "Separated to the outside" (OR = 8.8; $p < 0.05$); "Changing lifestyle" (OR = 7.6; $p < 0.05$); "Under stress and anxiety" (OR = 6.1; $p < 0.05$); "Uncontrolled emotion and behavior" (OR = 5.3; $p < 0.05$); and "Over-taking care of and being concerned about the child(ren)" (OR = ; $p < 0.05$).

Discussion

Depressive symptoms/ signs observed by family members in our study were consistent with the literature on signs and symptoms of PPD, including sleep disturbances, emotional disorders, and separating from society (2, 11, 12). Emotional and behavioral disturbance, being under stress and anxiety, isolation, changing lifestyle, and inability to take care of their child(ren) are likely predictors of PPD. These are similar to studies summarizing PPD psychosocial predictors such as stress, social support, and family connection (13), or factors of food intake patterns, sleep status, exercise, and physical activities (14) were also commonly reported. An explanation could be that the biological changes after delivery lead to fatigue and changes in emotions, behaviors, and daily activities (15). However, without early detection and/or proper treatment and care, psychoses occur in 1 to 2 per 1,000 postpartum women, and they may present as schizophrenic or affective disorders or as confused states (16).

Abbreviations: PPD, postpartum depression; EPDS, Edinburgh Postpartum Depression Scale; DSM-V, Diagnostic and Statistical Manual of Mental Disorder V.

TABLE 2 Depression symptoms reported by family members.

Symptoms reported by a family member		EPDS		OR
		≥12	<12	
Uncontrolling emotion and behavior	Yes	15	13	5.3*
	No	5	23	Reference
Being sad unreasonably	Yes	17	12	11.3*
	No	3	24	Reference
Under stress and anxiety	Yes	11	6	6.1*
	No	9	30	Reference
Separated to the outside	Yes	4	1	8.8*
	No	16	35	Reference
Changing lifestyle	Yes	11	5	7.6*
	No	9	31	Reference
Over-taking care of and concerning about the child(ren)	Yes	11	8	4.3*
	No	9	28	Reference
Ignoring child	Yes	1	0	2.9
	No	19	36	Reference
Having illusion	Yes	1	0	2.9*
	No	19	36	Reference

*$p \leq 0.05$.

Changing lifestyles reported by family members, such as eating habits or sleeping patterns, which were significantly associated with the risk of PPD, were also commonly found in women with PPD in other previous studies. It is not denied that we frequently observed the signs or symptoms of eating disorders related to depression among postpartum women. Typically, our results found a clear association between PPD and changing lifestyles ($p < 0.001$). Similarly, in another study, taking unhealthy food, and performing an unhealthy lifestyle, were found to have a significant relationship with depression, with 26.1% changing their daily lifestyle (17). Sleeping disorder is not only a predictor of PPD but also is a consequence of increasing PPD (18). Recently researchers examined the links between maternal sleep, maternal depressive symptoms, and mothers' perceptions of their emotional relationship with their infant in a self-recruited sample of mothers (11). Some studies described the association between serotonin and anxiety and depressive symptoms would be consistent with numerous observations indicating abnormal functioning of the serotoninergic system in depression for people experiencing anorexia or overeating (19).

Conclusion

A larger-scale study with a bigger sample size is recommended to provide more substantial evidence that many symptoms/ signs of PPD can be reported by family members. It will then be followed by strategies to raise awareness that family members can play a crucial role in early screening for PPD.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Hanoi Medical University Research Proposal Committee, Decision No. 5042/Q-HYHN. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Norhayati M, Hazlina NN, Asrenee A, Emilin WW. Magnitude and risk factors for postpartum symptoms: a literature review. *J Affect Disord.* (2015) 175:34–52. doi: 10.1016/j.jad.2014.12.041

2. Murray L, Dunne MP, Van Vo T, Anh PNT, Khawaja NG, Cao TN. Postnatal depressive symptoms amongst women in Central Vietnam: a cross-sectional study investigating prevalence and associations with social, cultural and infant factors. *BMC Pregnancy Childbirth.* (2015) 15:234. doi: 10.1186/s12884-015-0662-5

3. Tungchama FP, Piwuna CG, Armiya'u AY, Maigari YT, Davou FJ, Umar MU, et al. Relationship between quality of life and postpartum depression among women in North-Central, Nigeria. *Highland Med Res J.* (2017) 17:11–8.

4. Glover V. Maternal depression, anxiety and stress during pregnancy and child outcome; what needs to be done. *Best Practice Res Clin Obstetrics Gynaecol.* (2014) 28:25–35. doi: 10.1016/j.bpobgyn.2013.08.017

5. Stewart DE, Vigod S. Postpartum depression. *N Engl J Med.* (2016) 375:2177–86. doi: 10.1056/NEJMcp1607649

6. Kozinszky Z, Dudas RB. Validation studies of the edinburgh postnatal depression scale for the antenatal period. *J Affect Disord.* (2015) 176:95–105. doi: 10.1016/j.jad.2015.01.044

7. Tran TD, Tran T, Fisher J. Validation of three psychometric instruments for screening for perinatal common mental disorders in men in the north of Vietnam. *J Affect Disord.* (2012) 136:104–9. doi: 10.1016/j.jad.2011.08.012

8. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*: American Psychiatric Pub (2013). doi: 10.1176/appi.books.9780890425596

9. Kettunen P, Koistinen E, Hintikka J. Is postpartum depression a homogenous disorder: time of onset, severity, symptoms and hopelessness in relation to the course of depression. *BMC Pregnancy Childbirth.* (2014) 14:402. doi: 10.1186/s12884-014-0402-2

10. Marshall EM, Simpson JA, Rholes WS. Personality, communication, and depressive symptoms across the transition to parenthood: A dyadic longitudinal investigation. *Eur J Pers.* (2015) 29:216–34. doi: 10.1002/per.1980

11. Haddad S, Brown D, Dennis C, Lee K, Murray B, Pullenayegum E, et al. 1104 Maternal depression and sleep quality in early postpartum: Do maternal sleep-related cognitions and nighttime behaviours mediate the relationship? *J Sleep Sleep Disord Res.* (2017) 40(suppl_1):A411–2. doi: 10.1093/sleepj/zsx050.1103

12. Hatton DC, Harrison-Hohner J, Coste S, Dorato V, Curet LB, McCarron DA. Symptoms of postpartum depression and breastfeeding. *J Human Lactation.* (2005) 21:444–9. doi: 10.1177/0890334405280947

13. Yim IS, Stapleton LRT, Guardino CM, Hahn-Holbrook J, Schetter CD. Biological and psychosocial predictors of postpartum depression: systematic review and call for integration. *Annu Rev Clin Psychol.* (2015) 11:99. doi: 10.1146/annurev-clinpsy-101414-020426

14. Ghaedrahmati M, Kazemi A, Kheirabadi G, Ebrahimi A, Bahrami M. Postpartum depression risk factors: A narrative review. *J Educ Health Promot.* (2017) 6.

15. Sacher J, Rekkas PV, Wilson AA, Houle S, Romano L, Hamidi J, et al. Relationship of monoamine oxidase-A distribution volume to postpartum depression and postpartum crying. *Neuropsychopharmacology.* (2015) 40:429–35. doi: 10.1038/npp.2014.190

16. VanderKruik R, Barreix M, Chou D, Allen T, Say L, Cohen LS. The global prevalence of postpartum psychosis: a systematic review. *BMC Psychiatry.* (2017) 17:272. doi: 10.1186/s12888-017-1427-7

17. Wan Mohamed Radzi CWJB, Salarzadeh Jenatabadi H, Samsudin N. Postpartum depression symptoms in survey-based research: a structural equation analysis. *BMC Public Health.* (2021) 21:1–12. doi: 10.1186/s12889-020-09999-2

18. Park EM, Meltzer-Brody S, Stickgold R. Poor sleep maintenance and subjective sleep quality are associated with postpartum maternal depression symptom severity. *Arch Women's Mental Health.* (2013) 16:539–47. doi: 10.1007/s00737-013-0356-9

19. Gauthier C, Hassler C, Mattar L, Launay JM, Callebert J. Symptoms of depression and anxiety in anorexia nervosa: Links with plasma tryptophan and serotonin metabolism. *Psychoneuroendocrinology.* (2014) 39:170–8. doi: 10.1016/j.psyneuen.2013.09.009

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 03 November 2022
DOI 10.3389/fpsyt.2022.1013108

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University, Germany

REVIEWED BY
Tamara J. Sussman,
Stony Brook University, United States
Jing Wang,
Fudan University, China

*CORRESPONDENCE
Zhengzhi Feng
fzz@tmmu.edu.cn
Zhiyi Chen
chenzhiyi@email.swu.edu.cn

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 06 August 2022
ACCEPTED 12 October 2022
PUBLISHED 03 November 2022

CITATION
Miao K, Liu X, Zhang X, Li Y, Liao X, Zhang R, Feng Z and Chen Z (2022) Revision and psychometric properties of the negative cognitive processing bias scale.
Front. Psychiatry 13:1013108.
doi: 10.3389/fpsyt.2022.1013108

Revision and psychometric properties of the negative cognitive processing bias scale

Kuan Miao[1], Xuerong Liu[1], Xiaoling Zhang[1], Yuanyuan Li[2], Xingya Liao[1], Rui Zhang[1], Zhengzhi Feng[1]* and Zhiyi Chen[1,3]*

[1]School of Psychology, Army Medical University, Chongqing, China, [2]Western Medical Branch of PLA General Hospital, Beijing, China, [3]School of Psychology, Experimental Research Center for Medical and Psychological Science (ERC-MPS), Army Medical University, Chongqing, China

Negative cognitive processing bias (NCPB) is a cognitive trait that makes individuals more inclined to prioritize negative external stimuli (cues) when processing information. Cognitive biases have long been observed in mood and anxiety disorders, improving validation of tools to measure this phenomenon will aid us to determine whether there is a robust relationship between NCPB and major depressive disorder, anxiety disorders and other clinical disorders. Despite the development of an initial measure of this trait, that is, the negative cognitive processing bias questionnaire (NCPBQ), the lack of psychometric examinations and applications in large-scale samples hinders the determination of its reliability and validity and further limits our understanding of how to measure the NCPB traits of individuals accurately. To address these issues, the current study evaluated the psychometric properties of the NCPBQ in a large-scale sample (n = 6,069), which was divided into two subsamples (Subsample 1, n = 3,035, serving as the exploratory subsample, and Subsample 2, n = 3,034, serving as the validation subsample), and further revised it into a standardized scale, that is the negative cognitive processing bias scale (NCPBS), based on psychometric constructs. The results show that NCPBS possesses good construct reliability, internally consistent reliability, and test-retest reliability. Furthermore, by removing two original items from NCPBQ, NCPBS was found to have good criterion-related validity. In conclusion, the present study provides a reliable and valid scale for assessing negative cognitive processing bias of individuals.

KEYWORDS

negative cognitive processing bias, mental health, depression, scale revision, psychometric properties, measurement

Introduction

Negative cognitive processing bias (NCPB) is a cognitive trait which not only directs attention to negative internal or external stimuli but also leads to misinterpretation of this information in a more negative way (1). Thus, the negative impacts that NCPB may have on psychological health and psychiatric conditions have sparked great interest in the scientific community. For instance, a study has demonstrated the prominent predictive role of NCPB on short-sighted judgments and decisions (2). Furthermore, NCPB has been revealed as one of the most common phenomena of major depressive disorder (MDD) and anxiety disorders (3, 4). Moreover, NCPB has been established to not only play an important role in the onset of depression, but also maintain depressed mood states (5). To reveal why a close association between NCPB and these mood disorders exists, Beck (6) explained that hypersensitivity in processing negative information (i.e., NCPB) in daily life is the key cognitive pedestal for depression symptom development and maintenance (6). Numerous studies have shown a robust link between NCPB and cognitive-related mental health problems (7, 8). Although NCPB is important for mental health, reliable and valid tools to measure this trait accurately are still scarce.

Different aspects of NCPB in depression and anxiety have been examined, including attention bias (9), memory bias (10), and interpretation bias (11, 12). Negative attention bias, which acts as the first filter for information selection, shows an attentional preference for negative stimuli and deviation from positive stimuli (13). Negative memory bias is the inclination to recall negative materials more often than positive materials. A supporting evidence is that patients with MDD show poorer recall performance for positive stimuli in memory tests than healthy controls (14). Moreover, interpretation bias involves prominent preferences for interpreting information or materials in negative ways (15). Recently, the response styles theory (RST) proposed a new framework for explaining NCPB for repetitive rethinking of negative memories and information that predominates transdiagnostic hallmark across mood disorders (16). Thus, predominating negative information in repetitive rethinking – that is, rumination – has been increasingly indicated to be an additional profile for NCPB (17).

Despite the lack of a reliable and valid scale to measure NCPB in terms of the conceptual structure mentioned above, several tools have been developed to partly assess cognitive bias. For example, the Dysfunctional Attitudes Scale (DAS) (18) was built to assess individuals' maladaptive attitudes and beliefs about life and contains three main facets: perfectionism, utilitarianism, and criticism, which are not fully equal to one's cognitive traits (i.e., NCPB). In addition, the Automatic Thoughts Questionnaire (ATQ) (19) was developed to investigate the frequency of negative automatic thoughts in self-statements. However, it focuses on individuals' automatic thoughts with multiple cognitive components rather than cognitive traits. Furthermore, the Cognitive Bias Questionnaire (CBQ) (20) and Cognitive Style Questionnaire (CSQ) (21) were developed to measure one's cognitive processing styles, such as negative self-evaluation, cognitive distortion, and outcome expectation. However, both questionnaires aim to measure one's daily life behaviors and quantify the NCPB trait indirectly. Recently, the Negative Cognitive Processing Bias Questionnaire (NCPBQ) (22) was initially proposed to measure NCPB directly following four subdimensions. However, due to the lack of psychometric examinations and limited applications in small-scale samples, such as military personnel (23) and elderly individuals (24), the reliability and validity of the NCPBQ remain unclear. Furthermore, disparities in the construct structure of NCPBQ were found in previous studies (23, 24). Thus, it is necessary to examine the psychometric properties of NCPBQ and revise it into a standardized scale.

To address these issues, we recruited a large-scale sample (n = 6,069) across mainland China. In Subsample 1 (n = 3,035), we examined the reliability of the original version of NCPBQ using internal consistency analysis and test-retest analysis. In addition, validity was examined using criterion-related analysis and exploratory factor analysis (EFA) model. Furthermore, the NCPBQ was revised using standardized pipelines for building the NCPB scale. Finally, confirmatory factor analysis (CFA) was used for construct validity examination in Subsample 2 (n = 3,034).

Materials and methods

Participants

A large-scale sample (n = 6,069) was recruited using a hierarchical random sampling method (45.18% females and a mean age of 31.44 years, SD = 8.99, range = 18–65 years). Participants were recruited based on the provincial population distribution in mainland China, with a larger number of participants recruited in provinces with a larger population (e.g., Guangdong, Shandong and Henan). This sample pool covered the vast majority of occupations in mainland China (e.g., students, farmers, and businessmen). The sociodemographic features and results of statistical test are shown in **Table 1**.

All the participants were instructed to complete an online survey *via* a webpage and received payment for their participation. All the participants provided written informed consent preceding access to the online questionnaire. In addition, items for lie detection were included in the survey for quality control. A third-party platform (WJX Platform, Ran-Xin Technique Co. Ltd., Changsha, China) was involved in the

TABLE 1 Descriptive characteristics of the participants (n = 6,069).

Variables	Grouping	Frequency	Percent (%)	χ^2	p
Gender	Female	2,736	45.18	58.73	0.000***
	Male	3,333	54.82		
Age	18–25	1,578	26.00	3,967.95	0.000***
	26–35	2,879	47.44		
	36–45	1,109	18.27		
	46–55	406	6.69		
	56–65	97	1.60		
Educational attainment	Primary school	305	5.03	6,092.57	0.000***
	Middle school	497	8.19		
	High school	1,344	22.15		
	Bachelor's degree	3,530	58.16		
	Master's degree or above	393	6.47		
Occupation	Student	711	11.72	3,580.89	0.000***
	Farmer	379	6.24		
	Manual worker	648	10.68		
	Military personnel	346	5.70		
	Public servant	1,070	17.63		
	Businessman/Office worker	2,402	39.58		
	Intellectual/Scientific researcher	513	8.45		

***$p < 0.001$.

sampling, data acquisition and quality control. This study was approved by the IRB of Army Medical University (China).

Measures

Sociodemographic characteristics

To ensure sample representativeness, a sociodemographic investigation was conducted. This part included items as follows: gender, age, educational attainment, and occupation.

Negative cognitive processing bias questionnaire

The NCPBQ was initially developed by Yan et al. (22) to assess cognitive processing traits. The original version of the NCPBQ contained 23 self-reported items rated on a 4-point Likert-scale style ("1" for "never"; "4" for "always") in four dimensions: negative attention bias (NAB, e.g., My attention is easily drawn to the tragic images on TV and is difficult to shift.), negative memory bias (NMB, e.g., I can easily remember the negative comments people make about me.), negative interpretation bias (NIB, e.g., If a new leader or teacher is hard on me, I think it is because he sees me in a bad light and wants to get me in trouble.), and negative rumination bias (NRB, e.g., I often think about why I am so sad). Each dimension included five items, except these, there are three lie detection items in the original measure (items 4, 16, and 23. e.g., I have never told a lie.).

Dysfunctional attitude scale

The DAS consists of 40 items evaluating respondents' attitudes toward daily life, such as "undesirable life attitudes or beliefs," "black-and-white attitudes for moral judgment" and "perfectionism." This measure uses a 7-point Likert scale ranging from "completely disagree" to "completely agree," with higher scores indicating more maladaptive attitudes. This scale has been found to have high reliability (Cronbach's $\alpha = 0.93$) in psychometric examinations (25).

Beck depression inventory II

Developed and revised by Beck (26), Beck Depression Inventory II (BDI-II) contains 21 self-reported items and has long been acknowledged as one of the most broadly certified tools for assessing the severity of MDD (27). This inventory has been validated for good reliability and validity in clinical practice (28). The internal and test-retest reliability of the BDI-II of the Chinese version was found to be good (Cronbach's $\alpha = 0.94$; $r_{test-retest} = 0.55$) (29).

Statistical analyses

The full sample was divided into two subsamples, with one serving as an exploratory subsample (Subsample 1, n = 3,035) and the other serving as the validation subsample (Subsample 2, n = 3,034). Descriptive statistics were first reported for both subsamples. Item analysis was performed to examine the

items' suitability. Furthermore, to reveal the factor structures of the original version of the NCPBQ, EFA was conducted with principal component analysis (PCA), dimension reduction and varimax rotation in Subsample 1. By visual inspection of the scree plot and psychometric criteria (factor eigenvalues > 1.0), the number of factor structures was determined. Moreover, to revise the original version of the NCPBQ for a better factor structure, items with loadings under 0.50 were removed (30). Finally, CFA was carried out on the revised negative cognitive processing bias scale (NCPBS) in Subsample 2 to validate the factor structure. Seven metrics assessing goodness-of-fit were drawn to evaluate this model, including root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), goodness-of-fit index (GFI), comparative fit index (CFI), normed fit index (NFI), incremental fit index (IFI), and the Tucker-Lewis Index (TLI). An RMSEA and SRMR of <0.05; a GFI and CFI > 0.95; and a NFI, IFI, and TLI > 0.90, together, would suggest a good model fit (31).

The reliability was evaluated mainly *via* internal consistency reliability and test-retest reliability. Cronbach's alpha (α) and McDonald's omega (ω) were used to measure the internal consistency reliability. Since there is an argument regarding which is the best measure for assessing internal consistency reliability (32), α and ω were both calculated, with values of 0.70 or higher considered acceptable (33). In addition to internal consistency reliability, the 2-month test-retest reliability was also evaluated. A Pearson *r* value >0.50 indicated good test-retest reliability for a given scale (34).

The criterion-related validity of the revised NCPBS was estimated by the Pearson bivariate correlations across the NCPBS, BDI-II, and DAS, with significantly positive correlations between the NCPBS and both the BDI-II and DAS for high validity.

To gain further insights into the validity of this revised scale, between-group differences were examined for demographic features, which were compared by using independent sample *t*-tests or one-way ANOVA (Bonferroni correction for *post hoc* test), including gender, age, and education. The participants were classified into three age groups: early-adult group (aged 18–30), mid-adult group (aged 31–45), and old-adult group

TABLE 2 Descriptive statistics and the normality of data.

Item	Item score		Kolmogorov–Smirnov test		Skewness and kurtosis	
	M	SD	*D*-value	*P*-value	*S*	*K*
P1	2.48	0.782	0.249	0.000***	−0.075	−0.425
P2	2.59	0.938	0.205	0.000***	−0.043	−0.897
P3	2.15	0.864	0.244	0.000***	0.334	−0.581
P5	2.07	0.759	0.279	0.000***	0.345	−0.206
P6	2.26	0.850	0.235	0.000***	0.159	−0.645
P7	2.42	0.913	0.214	0.000***	0.064	−0.808
P8	2.78	0.848	0.258	0.000***	−0.295	−0.504
P9	2.33	0.867	0.250	0.000***	0.221	−0.602
P10	2.12	0.963	0.223	0.000***	0.451	−0.790
P11	2.68	0.922	0.229	0.000***	−0.200	−0.797
P12	2.43	0.864	0.227	0.000***	0.063	−0.659
P13	2.10	0.860	0.258	0.000***	0.454	−0.417
P14	2.41	0.926	0.209	0.000***	0.070	−0.853
P15	2.60	0.939	0.220	0.000***	−0.128	−0.870
P17	2.24	0.799	0.258	0.000***	0.175	−0.465
P18	2.23	0.867	0.252	0.000***	0.298	−0.566
P19	2.12	0.897	0.239	0.000***	0.406	−0.620
P20	2.74	0.870	0.257	0.000***	−0.296	−0.561
P21	2.39	0.915	0.218	0.000***	0.100	−0.811
P22	2.01	0.856	0.247	0.000***	0.516	−0.397
NAB	11.47	3.056	0.086	0.000***	0.108	−0.403
NMB	10.601	2.674	0.081	0.000***	−0.119	−0.443
NIB	11.206	2.870	0.087	0.000***	0.177	−0.240
NRB	8.66	2.712	0.103	0.000***	0.285	−0.451
Total	41.94	8.767	0.035	0.000***	−0.051	−0.092

M, mean; SD, standard deviation; NAB, negative attention bias; NMB, negative memory bias; NIB, negative interpretation bias; NRB, negative rumination bias.
***$p < 0.001$.

TABLE 3 Results of the item analysis.

	Extreme groups analysis	Item-total correlation		Homogeneity test		
Item	Critical ratio value	Item-total correlation	Corrected item-total correlations	Cronbach's α if item omitted	Communalities	Factor loading
P1	41.284***	0.519***	0.456	0.873	0.278	0.527
P2	46.056***	0.553***	0.480	0.872	0.299	0.547
P3	39.364***	0.501***	0.429	0.874	0.247	0.497
P5	39.606***	0.504***	0.441	0.873	0.258	0.508
P6	45.844***	0.553***	0.487	0.872	0.315	0.561
P7	52.139***	0.593***	0.526	0.870	0.350	0.592
P8	33.026***	0.439***	0.364	0.876	0.178	0.422
P9	49.757***	0.588***	0.524	0.871	0.350	0.592
P10	47.672***	0.564***	0.490	0.872	0.313	0.559
P11	50.289***	0.585***	0.517	0.871	0.340	0.583
P12	40.065***	0.508***	0.437	0.873	0.253	0.503
P13	42.641***	0.539***	0.471	0.872	0.298	0.546
P14	38.725***	0.501***	0.423	0.874	0.238	0.488
P15	46.622***	0.567***	0.495	0.871	0.316	0.562
P17	47.257***	0.582***	0.523	0.871	0.355	0.596
P18	42.299***	0.517***	0.447	0.873	0.264	0.514
P19	49.821***	0.590***	0.524	0.870	0.352	0.593
P20	49.701***	0.598***	0.535	0.870	0.359	0.599
P21	46.778***	0.568***	0.498	0.871	0.328	0.573
P22	52.035***	0.601***	0.539	0.870	0.368	0.607

***$p < 0.001$.

TABLE 4 Standardized factor loading of the negative cognitive processing bias scale.

Item		NAB	NMB	NIB	NRB
1	My attention is easily drawn to tragic images on TV and is difficult to shift	0.728			
17	My attention is easily drawn to the sad expressions of others and is difficult to shift	0.719			
6	My attention is easily drawn to harrowing sounds and is difficult to shift	0.690			
21	My attention is easily drawn to the tragic storylines of the novels and is difficult to shift	0.653			
13	My attention is easily drawn to the hesitant eyes of others and is difficult to shift	0.602			
20	I can easily remember the negative comments people make about me		0.731		
15	Even if I think I have done nothing wrong, I remember the criticism of others for a long time		0.711		
11	In the process of interacting with others, if I say the wrong thing, I will not forget it for a long time		0.665		
2	I still vividly remember a time when I was ridiculed		0.562		
5	If I meet a friend for the first time and he (she) says very little to me, I will think he or she doesn't like me			0.730	
3	If an acquaintance walks across the street and does not say hello to me, I will think he or she has a problem with me			0.643	
18	If a new leader or teacher is hard on me, I think it is because he sees me in a bad light and wants to get me in trouble			0.633	
9	If I were to go on stage and give a speech in public, and when I come down, I see a few people next to me whispering, I think they are laughing at my bad speech			0.594	
12	If I participated in a job applications and the interviewer had a serious expression throughout the process, I would think that the application would most likely fail.			0.525	
22	I often think about why I am so sad				0.702
19	I often think about why my mood is low and those of others are not				0.701
10	I often think about why I am so lonely				0.675
14	I often think about why I lack interest and motivation to do things				0.652
Percent of variance		14.787%	12.171%	13.017%	12.894%

NAB, negative attention bias; NMB, negative memory bias; NIB, negative interpretation bias; NRB, negative rumination bias.

TABLE 5 Confirmatory factor analysis model fit indexes.

	RMSEA	SRMR	GFI	CFI	NFI	IFI	TLI
Criteria	<0.05	<0.05	>0.95	>0.95	>0.90	>0.90	>0.90
Fit indexes	0.04	0.03	0.97	0.96	0.95	0.96	0.95

RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; GFI, goodness-of-fit index; CFI, comparative fit index; NFI, normed fit index; IFI, incremental fit index; TLI, Tucker-Lewis Index.

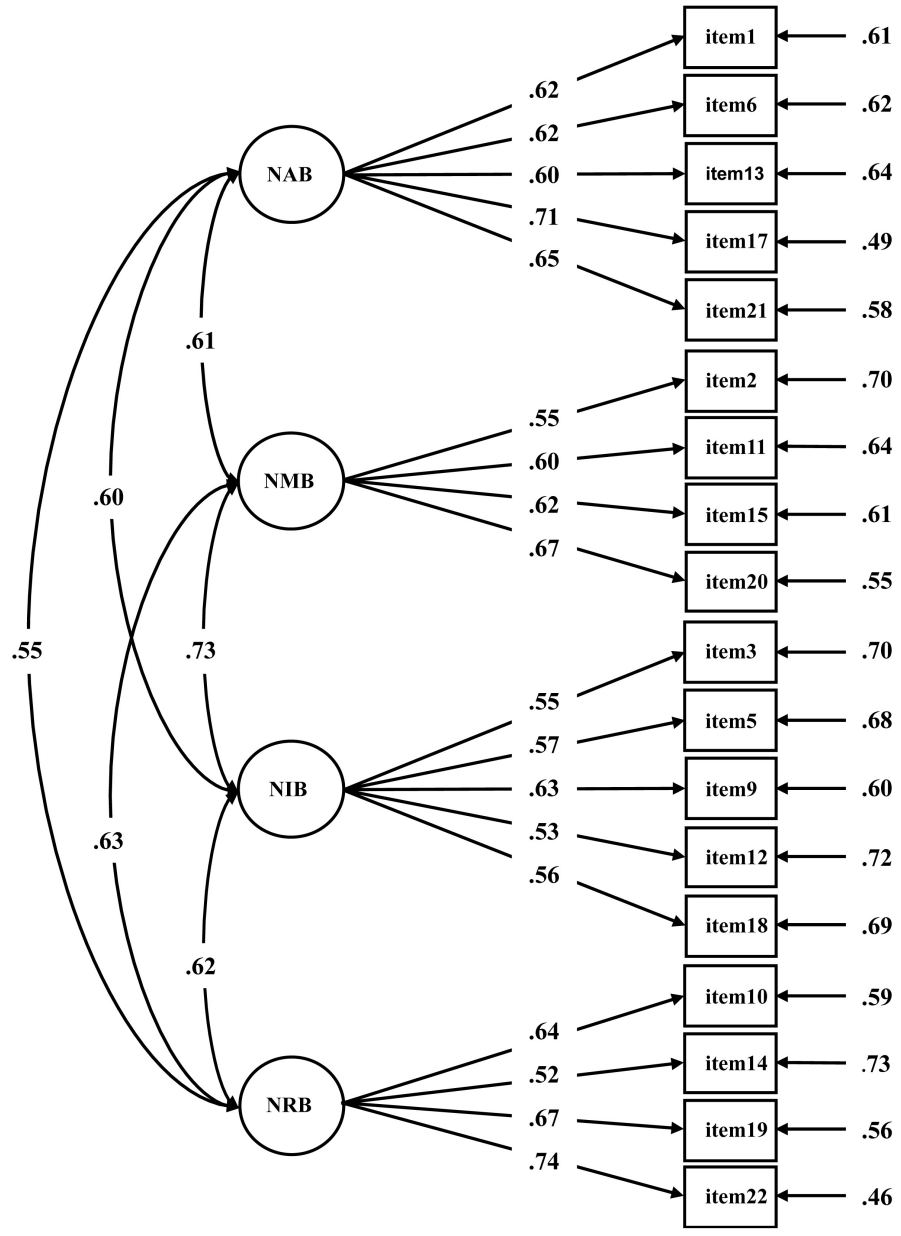


FIGURE 1
The four-factor confirmatory factor analysis model of the negative cognitive processing bias scale for the validation set (n = 3,034). Each number alongside the lines indicates standardized factor loading. NAB, negative attention bias; NMB, negative memory bias; NIB, negative interpretation bias; NRB, negative rumination bias.

(aged 46–65). Additionally, education level was divided into two group: well-educated group for educational experiences >13 years and a less-educated group for educational experiences <13 years.

The data were analyzed by IBM SPSS Statistics 26.0 (Armonk, NY: IBM Corp.), Amos 21.0 programs, and JASP 0.16.2.[1]

Results

Descriptive statistics

The descriptive statistics for each item and subdimensions were tabulated (see Table 2), including the mean value, standard deviation and normality (estimated by Kolmogorov–Smirnov test, skewness, and kurtosis). Although, the results showed each item failed the Kolmogorov–Smirnov test, we found no prominent irregular skewness or kurtosis for this sample (The distribution of item scores was listed in Supplementary Figures). Furthermore, a similar pattern was found in each dimension, including negative attention bias (NAB, 11.47 ± 3.06 for items 1, 6, 13, 17, and 21), negative memory bias (NMB, 10.60 ± 2.67, for items 2, 8, 11, 15, and 20), negative interpretation bias (NIB, 11.21 ± 2.87, for item 3, 5, 9, 12, and 18), and negative rumination bias (NRB, 8.66 ± 2.71, for items 7, 10, 14, 19, and 22).

1 https://jasp-stats.org/

TABLE 6 Pearson's correlations between the negative cognitive processing bias scale and the criterion measures.

	Overall scale	NAB	NMB	NIB	NRB
DAS	0.551***	0.357***	0.376***	0.494***	0.466***
BDI-II	0.447***	0.315***	0.346***	0.328***	0.377***

NAB, negative attention bias; NMB, negative memory bias; NIB, negative interpretation bias; NRB, negative rumination bias.
***$p < 0.001$.

TABLE 7 Cronbach's α, McDonald's ω, and test-retest reliability of the negative cognitive processing bias scale.

	Cronbach's *a*	McDonald's *ω*	Test-retest reliability
Overall scale	0.866	0.866	0.943
NAB	0.733	0.734	0.705
NMB	0.698	0.701	0.785
NIB	0.703	0.704	0.761
NRB	0.732	0.735	0.748

NAB, negative attention bias; NMB, negative memory bias; NIB, negative interpretation bias; NRB, negative rumination bias.

Item analysis

To examine the validity of each item, critical ratio (CR) method, Pearson correlation method and a homogeneity test were applied to Subsample 1 for item analysis. The results revealed significant score differences between the high-total-score (top 27%) and low-total-score groups (last 27%) for all items, irrespective of CR values, indicating that all the items possessed high discrimination power (see Table 3). Further analysis also illustrated statistically significant item-total correlations. Finally, the results of the homogeneity test found acceptable communalities and factor loadings, except for item 8 (see Table 3, communalities = 0.178, factor loading = 0.422). As a result, item 8 (I always remember my mistakes clearly) was removed in this step.

Exploratory factor analysis

The results demonstrated the suitability of EFA for the current dataset with an acceptable KMO coefficient (=0.93) and significant skewness from a spherical distribution (χ^2 = 15,492.751, $p < 0.001$). Furthermore, principal axis factoring (PAF) was adopted for factor extraction and loading estimation, with <0.5 used as the exclusion criterion. The results indicated a four-dimensional structure for the NCPBQ, in which four common factors with eigenvalues >1 in orthogonal rotation from the maximum variance method), accounting for 52.87% of the total variance (see Table 4). Although the four-facet construct structure was validated here, item 7 was found to be unacceptable, as its factor loading was less than 0.50 (factor loading = 0.492). Finally, item 7 (I often think about why I am always inferior to others) was removed in this step.

On balance, the original version of the NCPBQ has been revised by removing two items (item 7 and item 8) based on the above results.

Validity analysis

Confirmatory factor analysis

To estimate the construct validity for the revised version, CFA was carried on this four-dimensional structure in independent Subsample 2. The results revealed good goodness-of-fit metrics for the revised NCPBS (RMSEA = 0.04, SRMR = 0.03, and GFI = 0.97, more details in Table 5). In addition, all items were found to have acceptable factor loadings (β = 0.52–0.74, see Figure 1).

Criterion-related validity

To test the validity of the revised NCPBS, the criterion-related validity was estimated by correlating its scores to those of the DAS and BDI-II. The results showed that the scores for both

DAS (r = 0.551, $p < 0.001$) and BDI-II (r = 0.447, $p < 0.001$) were significantly correlated with the total score of the revised NCPBS and even its subdimensions (see Table 6).

Reliability analysis

Cronbach's α and McDonald's ω were calculated to estimate the internal consistency of the revised NCPBS. The results showed good internal consistency reliability for this revised version, including the whole scale and its subdimensions (both α and ω = 0.866, see Table 7). Furthermore, significant correlations were found in the 2-month test-retest consistency analysis, demonstrating good test-retest consistency reliability for the revised NCPBS (r = 0.943 for overall scale scores, more results can be found in Table 7).

Differential analysis for sociodemographic features

To examine whether the revised NCPBS was valid, differential analysis was conducted for sociodemographic features in the whole sample. The results revealed significant differences between genders for NCPBS scores (total for males: 41.34 ± 8.84, total for females: 42.67 ± 8.62, t = −5.936, $p < 0.001$; BF_{10} = 1.18 × 10^6 at JSY Cauchy distribution, see Table 8). Furthermore, we also found difference between educational levels in the NCPBS scores, with low scores for low educational level (total for less-educated group: 39.80 ± 8.49, total for well-educated group: 43.10 ± 8.70, t = −14.258, $p < 0.001$; BF_{10} = 6.26 × 10^{41} at JSY Cauchy distribution, see Table 9). Finally, the NCPBS scores varied between age-related groups (total for early-adult group: 42.83 ± 8.96, total for mid-adult group: 40.97 ± 8.43, total for old-adult group: 40.53 ± 8.32, F = 38.082, $p < 0.001$, see Table 10).

Discussion

The purpose of this study was to evaluate the psychometric properties of the original version of the NCPBQ in a large sample and to revise it into a reliable and valid scale. As we expected, the results showed good four-dimensional construct validity and reliability for the revised NCPBS. Furthermore, small differences in NCPBS scores across sociodemographic features, including gender, educational level, and age, were found, with old less-educated males exhibiting low NCPB. On balance, the current study revealed the psychometric properties of the initial NCPBQ and further revised it into a reliable and valid scale for measuring individuals' cognitive trait in negative information processing.

Here, a weakly skewed distribution was found in the item analysis for NCPBS, which seemed to align with general prevalence of cognition-biased mood disorder (e.g., MDD). The lifetime prevalence of these psychiatric conditions that reported in previous literature, such as depressive disorders (6.8%) and anxiety disorders (7.6%), were low in China (35, 36). Accordingly, we found no prominent irregular skewness or kurtosis for this sample by canonical criteria (i.e., skewness < 3; kurtosis < 8) (37). In this vein, this finding indicated that

TABLE 8 Results for gender differences in negative cognitive processing bias scale scores.

	Male (n = 3,333)	Female (n = 2,736)	T-value	P-value	BF_{10}
NAB	11.25 ± 3.06	11.75 ± 3.03	−6.348	0.000***	1.458 × 10^6
NMB	10.50 ± 2.67	10.72 ± 2.68	−3.239	0.001**	5.422
NIB	11.01 ± 2.88	11.45 ± 2.84	−6.027	0.000***	2.034 × 10^6
NRB	8.58 ± 2.74	8.75 ± 2.68	−2.455	0.014*	0.586
Total	41.34 ± 8.84	42.67 ± 8.62	−5.936	0.000***	1.184 × 10^6

BF, Bayesian factor; NAB, negative attention bias; NMB, negative memory bias; NIB, negative interpretation bias; NRB, negative rumination bias.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

TABLE 9 Results for education level differences in negative cognitive processing bias scale scores.

	Less-educated (n = 2,146)	Well-educated (n = 3,904)	T-value	P-value	BF_{10}
NAB	10.86 ± 2.88	11.80 ± 3.10	−11.648	0.000***	3.505 × 10^{27}
NMB	10.07 ± 2.57	10.89 ± 2.68	−11.586	0.000***	1.740 × 10^{27}
NIB	10.48 ± 2.80	11.59 ± 2.83	−14.618	0.000***	4.400 × 10^{43}
NRB	8.38 ± 2.59	8.81 ± 2.76	−5.925	0.000***	1.155 × 10^{43}
Total	39.80 ± 8.49	43.10 ± 8.70	−14.258	0.000***	6.263 × 10^{41}

BF, Bayesian factor; NAB, negative attention bias; NMB, negative memory bias; NIB, negative interpretation bias; NRB, negative rumination bias.
***$p < 0.001$.

TABLE 10 Results for age differences in negative cognitive processing bias scale scores.

	18–30 (n = 3,278)	31–45 (n = 2,288)	46–65 (n = 503)	ANOVA		*Post hoc* test
				F-value	*P*-value	*P*-value (corrected)
NAB	11.60 ± 3.10	11.31 ± 2.97	11.41 ± 3.11	6.400	0.002**	1–2:0.001** 1–3:0.566 2–3: 1.000
NMB	10.87 ± 2.72	10.32 ± 2.59	10.15 ± 2.50	36.622	0.000***	**1–2:0.000***** **1–3:0.000***** 2–3:0.625
NIB	11.45 ± 2.91	10.97 ± 2.81	10.70 ± 2.71	26.956	0.000***	**1–2:0.000***** **1–3:0.000***** 2–3:0.148
NRB	8.92 ± 2.76	8.37 ± 2.62	8.27 ± 2.62	33.795	0.000***	**1–2:0.000***** **1–3:0.000***** 2–3: 1.000
Total	42.83 ± 8.96	40.97 ± 8.43	40.53 ± 8.32	38.082	0.000***	**1–2:0.000***** **1–3:0.000***** 2–3:0.918

NAB, negative attention bias; NMB, negative memory bias; NIB, negative interpretation bias; NRB, negative rumination bias.
**$p < 0.01$, ***$p < 0.001$. Bold font in *post hoc* test indicates $p < 0.05$.

the NCPBS may be a valid tool for measuring cognitive processing trait.

Furthermore, the present study addressed a long-standing debate over the factor structure of NCPB. Both EFA and CFA revealed that the four-factor structure possessed good construct validity for the revised NCPBS, which strongly supports RST (16). The conventional theoretical basis for NCPB frames individuals' cognitive biases in terms of fundamental cognitive components, such as attention, memory, and interpretation (38). However, it should be borne in mind that these cognitive faculties bias individual's behaviors through "processing." Thus, as a typical processing style, rumination functions to boost cognitive biases based on these cognitive components, which may determine the extent to which NCPB increases the risk of mood disorders (39). Thus, this study clarified the potential structure of NCPB by psychometric methods. In addition, NCPBS showed better reliability and validity than the previous version. Thus, the major goal of the current study was to provide a standardized NCPBS to accurately measure individuals' negative cognitive trait.

In addition to revising the NCPBS, some between-group differences were also found. Although the differences were very small, the exploratory explanations would be inferred here. Firstly, a small gender difference in the NCPBS scores was observed, with slightly higher scores in females. This result may be supported by both theoretical and empirical evidence, as gender-related environment susceptibility theory proposes that females detect more subtle negative cues from daily life events and the environment due to genetic imprinting (40–42). In addition, previous studies have validated this theoretical model, showing increased neural activity and behavioral reactions to negative information in females compared with that of males (43, 44). Thus, in the current study, these indirect evidence may imply higher negative information susceptibility in females compared to males. Furthermore, we observed slightly lower NCPB for individuals with a less educational level, which is consistent with previous evidence. Daraei and Ghaderi (45) documented the association of a low education level with optimism and well-being (45). Besides, compared to elders, young adults exhibited a little higher NCPB as measured by the revised NCPBS. This could be explained well by the differences in their emotional regulation ability. Existing studies have revealed that, as predicted by emotional regulation ability (including regulation resource and regulation strategy), older adults exhibit better decision-making ability in both positive and negative emotional conditions than young adults (46–48). Thus, we inferred that such age-related effects in emotional processing may cause different NCPB for distinct age groups. Together, these evidences suggest that the current study may provide a valid and reliable measure to quantify individuals' cognitive trait, with potentials for application in psychological and psychiatric domains.

Limitations

Several limitations in the current study should be borne in mind before applying the NCPBS. Despite claiming it to be a robust predictor of depression, little is known about whether this scale can be used in clinical practice because no depression patients were recruited in the present study. Thus, a cohort study for investigating the association between depression and NCPB

in clinical practice is needed in the future. In addition, this large-scale sample was taken only from the Chinese population, which hampers the cross-cultural generalizability of this scale. To address this issue, future studies should examine the reliability and validity of the scale by using a broad sample.

Conclusion

The current study recruited a large-scale sample to validate the psychometric properties of the NCPBQ, and followed a standardized pipeline to revise the scale. The results show that NCPBS has better reliability and validity than the original version, with higher internal consistency reliability and construct validity. Furthermore, we found the statistical differences in NCPB across sociodemographic features by using NCPBS, which provided further evidence of external validity of this scale. Taken together, this study provides a reliable and valid measure to estimate individuals' cognitive inclinations toward negative content accurately and advanced our understanding of the core components of NCPB.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/ucmw4/.

Ethics statement

The studies involving human participants were reviewed and approved by the IRB of Army Medical University (China). The patients/participants provided their written informed consent to participate in this study.

Author contributions

KM: writing – original draft and review, and editing and data curation. XUL and XZ: writing – review and editing. YL, XIL, and RZ: data curation. ZF and ZC: conceptualization. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (NSFC 81971278).

Acknowledgments

We hereby appreciated Professor Qin Dai (Army Medical University) for her advice with regards to conceptional construction and statistical model, Professor Xi Luo (Army Medical University) and Xingmei Gu (Army Medical University) for their language editing, and Chang Liu (Army Medical University) for her assistance in preparation of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2022.1013108/full#supplementary-material

References

1. Cacioppo JT, Cacioppo S, Gollan JK. The negativity bias: conceptualization, quantification, and individual differences. *Behav Brain Sci.* (2014) 37:309–10. doi: 10.1017/S0140525X13002537

2. Pronin E. Perception and misperception of bias in human judgment. *Trends Cogn Sci.* (2007) 11:37–43. doi: 10.1016/j.tics.2006.11.001

3. Beevers CG, Mullarkey MC, Dainer-Best J, Stewart RA, Labrada J, Allen JJB, et al. Association between negative cognitive bias and depression: a symptom-level approach. *J Abnorm Psychol.* (2019) 128:212–27. doi: 10.1037/abn0000405

4. Clark DA, Beck AT. Cognitive theory and therapy of anxiety and depression: convergence with neurobiological findings. *Trends Cogn Sci.* (2010) 14:418–24. doi: 10.1111/1467-8721.01242

5. Browning M, Blackwell SE, Holmes EA. *The use of cognitive bias modification and imagery in the understanding and treatment of depression*. Berlin: Springer Press (2013). p. 243–60.

6. Beck AT. Cognitive models of depression. *J Cogn Psychother.* (1987) 1:5–37.

7. Smith EM, Reynolds S, Orchard F, Whalley HC, Chan SWY. Cognitive biases predict symptoms of depression, anxiety and wellbeing above and beyond neuroticism in adolescence. *J Affect Disord.* (2018) 241:446–53. doi: 10.1016/j.jad.2018.08.051

8. Tairi T, Adams B, Zilikis N. Cognitive errors in greek adolescents: the linkages between negative cognitive errors and anxious and depressive symptoms. *Int J Cogn Ther.* (2016) 9:261–78. doi: 10.1521/ijct_2016_09_11

9. Keller AS, Leikauf JE, Holt-Gosselin B, Staveland BR, Williams LM. Paying attention to attention in depression. *Transl Psychiatry.* (2019) 9:279. doi: 10.1038/s41398-019-0616-1

10. Burt DB, Zembar MJ, Niederehe G. Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity. *Psychol Bull.* (1995) 117:285–305. doi: 10.1037/0033-2909.117.2.285

11. Hirsch CR, Meeten F, Krahé C, Reeder C. Resolving ambiguity in emotional disorders: the nature and role of interpretation biases. *Annu Rev Clin Psychol.* (2016) 12:281–305. doi: 10.1146/annurev-clinpsy-021815-093436

12. Everaert J, Podina IR, Koster EH. A comprehensive meta-analysis of interpretation biases in depression. *Clin Psychol Rev.* (2017) 58:33–48. doi: 10.1016/j.cpr.2017.09.005

13. Gotlib IH, Joormann J. Cognition and depression: current status and future directions. *Annu Rev Clin Psychol.* (2010) 6:285–312. doi: 10.1146/annurev.clinpsy.121208.131305

14. Matt GE, Vázquez C, Campbell WK. Mood-congruent recall of affectively toned stimuli: a meta-analytic review. *Clin Psychol Rev.* (1992) 12:227–55. doi: 10.1016/0272-7358(92)90116-P

15. Beck AT, Haigh EAP. Advances in cognitive theory and therapy: the generic cognitive model. *Annu Rev Clin Psychol.* (2014) 10:1–24. doi: 10.1146/annurev-clinpsy-032813-153734

16. Nolen-Hoeksema S. The response styles theory. In: Papageorgiou C, Wells A, editors. *Depressive rumination: nature, theory and treatment*. Hoboken, NJ: Wiley (2003). p. 105–23. doi: 10.1002/9780470713853.ch6

17. Nolen-Hoeksema S, Wisco BE, Lyubomirsky S. Rethinking rumination. *Perspect Psychol Sci.* (2008) 3:400–24. doi: 10.1111/j.1745-6924.2008.00088.x

18. Weissman AN, Beck AT. Development and Validation of the dysfunctional attitude scale: a preliminary investigation. In: *Paper presented at the meeting of the American educational research association*. Toronto, ON (1978).

19. Hollon SD, Kendall PC. Cognitive self-statements in depression: development of an automatic thoughts questionnaire. *Cognit Ther Res.* (1980) 4:383–95. doi: 10.1007/BF01178214

20. Krantz S, Hammen CL. Assessment of cognitive bias in depression. *J Abnorm Psychol.* (1979) 88:611–9. doi: 10.1037/0021-843X.88.6.611

21. Abramson LY, Metalsky GI, Alloy LB. Hopelessness depression: a theory-based subtype of depression. *Psychol Rev.* (1989) 96:358–72. doi: 10.1037/0033-295X.96.2.358

22. Yan X, Zhang R, Feng Z. Development of negative cognitive processing bias questionnaire. *J Third Mil Med Univ.* (2017) 39:2329–34.

23. Yan X, Feng Z. Assessment of reliability and validity of negative cognitive processing bias scale in military personnel. *J Third Mil Med Univ.* (2018) 40:1619–23.

24. Zhou Y, Li Z. Reliability and validity of negative cognitive processing bias questionnaire in older adults. *Chin Nurs Res.* (2021) 31:44–7.

25. Rogers GM, Park J-H, Essex MJ, Klein MH, Silva SG, Hoyle RH, et al. The dysfunctional attitudes scale: psychometric properties in depressed adolescents. *J Clin Child Adolesc Psychol.* (2009) 38:781–9. doi: 10.1080/15374410903259007

26. Beck AT, Steer RA, Ball R. Comparison of beck depression inventories -IA and -II in psychiatric outpatients. *J Pers Assess.* (1996) 67:588-97.

27. Kind P. Measuring health: a guide to rating scales and questionnaires. *Public Health.* (2008) 122:217. doi: 10.1016/j.puhe.2007.04.003

28. Wang YP, Gorenstein C. Psychometric properties of the beck depression inventory-II: a comprehensive review. *Braz J Psychiatry.* (2013) 35:416–31. doi: 10.1590/1516-4446-2012-1048

29. Wang Z, Yuan C, Huang J, Li Z, Chen Y, Zhang H, et al. Reliability and validity of the Chinese version of beck depression inventory-II among depression patients. *Chin Mental Health J.* (2011) 25:476–80.

30. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract Assess Res Eval.* (2005) 10:1–9. doi: 10.7275/JYJ1-4868

31. Marsh HW, Hau K-T, Wen Z. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Struct Equ Model.* (2004) 11:320–41. doi: 10.1207/s15328007sem1103_2

32. Hayes AF, Coutts JJ. Use Omega rather than Cronbach's Alpha for estimating reliability. *Commun Methods Meas.* (2020) 14:1–24. doi: 10.1080/19312458.2020.1718629

33. George D, Mallery P. *SPSS for windows step by step: a simple guide and reference, 11.0 update*. Boston, MA: Allyn and Bacon (2003).

34. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. 3rd ed. Upper Saddle River, NJ: Pearson Prentice Hall (2008).

35. Lu J, Xu X, Huang Y, Li T, Ma C, Xu G, et al. Prevalence of depressive disorders and treatment in China: a cross-sectional epidemiological study. *Lancet Psychiatry.* (2021) 8:981–90. doi: 10.1016/S2215-0366(21)00251-0

36. Huang Y, Wang Y, Wang H, Liu Z, Yu X, Yu Y, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. *Lancet Psychiatry.* (2019) 3:211–24. doi: 10.1016/S2215-0366(18)30511-X

37. Kline RB. *Principles and practice of structural equation modeling.* 4th ed. New York, NY: Guilford Press (2016).

38. Beck AT. The evolution of the cognitive model of depression and its neurobiological correlates. *Am J Psychiatry.* (2008) 165:969–77. doi: 10.1176/appi.ajp.2008.08050721

39. Nolen-Hoeksema S. The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *J Abnorm Psychol.* (2000) 109:504–11. doi: 10.1037/0021-843X.109.3.504

40. Belsky J. Variation in susceptibility to environmental influence: an evolutionary argument. *Psychol Inq.* (1997) 8:182–6. doi: 10.1207/s15327965pli0803_3

41. Belsky J, Bakermans-Kranenburg MJ, van Ijzendoorn MH. For better and for worse: differential susceptibility to environmental influences. *Curr Dir Psychol Sci.* (2007) 16:300–4. doi: 10.1111/j.1467-8721.2007.00525.x

42. Hentges RF, Davies PT, Sturge-Apple ML. Domain specificity of differential susceptibility: testing an evolutionary theory of temperament in early childhood. *Dev Psychopathol.* (2022):1–14. [Epub ahead of print]. doi: 10.1017/S0954579422000256

43. Yuan J, Luo Y, Yan JH, Meng X, Yu F, Li H. Neural correlates of the females' susceptibility to negative emotions: an insight into gender-related prevalence of affective disturbances. *Hum Brain Mapp.* (2009) 30:3676–86. doi: 10.1002/hbm.20796

44. Levkovich I, Shinan-Altman S. The impact of gender on emotional reactions, perceived susceptibility and perceived knowledge about COVID-19 among the Israeli public. *Int Health.* (2021) 13:555–61. doi: 10.1093/inthealth/ihaa101

45. Daraei M, Ghaderi AR. Impact of education on optimism/pessimism. *J Indian Acad Appl Psychol.* (2012) 38:339–43.

46. Urry HL, Gross JJ. Emotion regulation in older age. *Curr Dir Psychol Sci.* (2010) 19:352–7. doi: 10.1177/0963721410388395

47. You XQ, Ju CT, Wang M, Zhang BS, Liu P. Age differences in the influence of induced negative emotion on decision-making: the role of emotion regulation. *J Gerontol B Psychol Sci Soc Sci.* (2019) 74:796–805. doi: 10.1093/geronb/gbx137

48. Chow S-M, Hamagani F, Nesselroade JR. Age differences in dynamical emotion-cognition linkages. *Psychol Aging.* (2007) 22:765–80. doi: 10.1037/0882-7974.22.4.765

frontiers | Frontiers in Psychiatry

TYPE Systematic Review
PUBLISHED 04 January 2023
DOI 10.3389/fpsyt.2022.1041770

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University, Germany

REVIEWED BY
Paolo Meneguzzo,
University of Padua, Italy
Xi Jiang,
University of Electronic Science and Technology of China, China

*CORRESPONDENCE
Taolin Chen
tlchen@scu.edu.cn

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 11 September 2022
ACCEPTED 15 November 2022
PUBLISHED 04 January 2023

CITATION
Guo H, Feng B, Ma Y, Zhang X, Fan H, Dong Z, Chen T and Gong Q (2023) Analysis of the screening and predicting characteristics of the house-tree-person drawing test for mental disorders: A systematic review and meta-analysis. *Front. Psychiatry* 13:1041770. doi: 10.3389/fpsyt.2022.1041770

Analysis of the screening and predicting characteristics of the house-tree-person drawing test for mental disorders: A systematic review and meta-analysis

Huibing Guo[1,2,3], Bin Feng[1,3], Yingqiao Ma[1,4], Xueyi Zhang[1,5], Huiyong Fan[6], Zaiquan Dong[7], Taolin Chen[1,4,5]* and Qiyong Gong[1,8,9]

[1]Huaxi MR Research Center (HMRRC), Department of Radiology, West China Hospital of Sichuan University, Chengdu, Sichuan, China, [2]Department of Student Affairs Management, West China School of Medicine, West China Hospital, Sichuan University, Chengdu, China, [3]West China School of Medicine, Sichuan University, Chengdu, China, [4]Research Unit of Psychoradiology, Chinese Academy of Medical Sciences, Chengdu, Sichuan, China, [5]College of Medical Technology, West China Hospital of Sichuan University, Chengdu, China, [6]Institute of Education, Bohai University, Jinzhou, China, [7]Mental Health Center, West China Hospital of Sichuan University, Chengdu, China, [8]Functional and Molecular Imaging Key Laboratory of Sichuan Province, West China Hospital of Sichuan University, Chengdu, Sichuan, China, [9]Department of Radiology, West China Xiamen Hospital of Sichuan University, Xiamen, Fujian, China

Background: The house-tree-person (HTP) drawing test has received growing attention from researchers as a common projective test. However, the methods used to select and interpret drawing indicators still lack uniformity.

Objective: This study aims to integrate drawing indicators into the process of screening for or classifying mental disorders by conducting a systematic review and meta-analysis of the application of the HTP test.

Methods: A search of the following electronic databases was performed in May 2022: PubMed, Web of Science, Embase, EBSCO, CNKI, VIP, and Wanfang. Screening and checking of the literature were performed independently by two researchers. The empirical studies published on the use of the HTP test in mental disorders and studies providing specific data on the occurrence frequency of drawing characteristics were analyzed. A total of 30 studies were included in the meta-analysis, including 665 independent effect sizes and 6,295 participants. The strength of the association between drawing characteristics of the HTP test and the prevalence of mental disorders was measured by the ratio (OR) with a *95% CI*. Publication bias was assessed using a funnel plot, Rosenthal's fail-safe number (N_{fs}), and the trim and fill method.

Results: The results revealed 50 drawing characteristics that appeared at least three times in previous studies, of which 39 were able to significantly predict mental disorders. The HTP test can be divided into the following four

dimensions: house, tree, person, and the whole. These dimensions reflect the structure, size, and other characteristics of the picture. The results showed that the greatest predictor of mental disorders was the whole (OR = 4.20, $p < 0.001$), followed by the house (OR = 3.95, $p < 0.001$), the tree (OR = 2.70, $p < 0.001$), and the person (OR = 2.16, $p < 0.001$). The valid predictors can be categorized into the following four types: item absence, bizarre or twisted, excessive details, and small or simplified. The subgroup analysis showed that the affective-specific indicators included *no motion, leaning house, and decorated roof;* thought-specific indicators included *excessive separation among items*, *no window*, *loss of facial features,* and *inappropriate body proportions*; and common indicators of mental disorders included *no additional decoration*, *simplified drawing*, *very small house*, *two-dimensional house,* and *very small tree*.

Conclusion: These findings can promote the standardization of the HTP test and provide a theoretical reference for the screening and clinical diagnosis of mental disorders.

KEYWORDS

house-tree-person drawing test, mental disorders, screening, aiding diagnosis, meta-analysis

Introduction

Mental disorders are usually associated with distress or cognitive function, emotional regulation, or behavioral impairment. The prevalence of mental disorders has been increasing annually in recent years and has become a major contributor to the global disease burden (1, 2). One in every 8, or 970 million people in the world, lives with a mental disorder (3). There are many different types of mental disorders, which can be classified into thought-type disorders and affective-type disorders according to the main symptoms. Affective-type disorders include depression and anxiety (4), while thought-type disorders mainly include schizophrenia, paranoia, rumination, etc. (5).

Accurate screening and diagnosis should be performed before treating mental disorders to reduce their prevalence. However, current assessments mainly rely on scales, and these traditional measures have many drawbacks (6, 7). For example, it is difficult to assess the symptoms of patients with unclear self-perceptions based on scale questions. Moreover, due to social desirability, subjects may deliberately choose positive answers to hide their symptoms, resulting in a lack of valid results (8, 9).

As one of the three major testing techniques in psychology, projective testing can be used to compensate for the shortcomings of scales (10). The comprehensive use of various testing techniques is a future trend and can aid in the development of a projective test with better validity (11, 12). The house-tree-person (HTP) drawing test was proposed by Buck in 1948 and is currently one of the most widely used projective tests (13). According to a survey of 102 commonly used psychological tests conducted by the American Psychological Association, HTP ranks 8th in usage (14). The HTP test has the following advantages: initiative, structure, and non-verbal. On the one hand, it can conceal the test purpose and overcome the defensive psychology of subjects. On the other hand, painting is not affected by a subject's culture and expression and thus can more accurately reflect personality traits and potential psychological problems (11, 15).

Many studies have applied the HTP test in screening and aiding the diagnosis of mental disorders. For example, one of the earliest studies examined whether the drawing characteristics in the HTP test were related to mental disorders, and they found a significant correlation between "line strength" and EEG; thus, line strength was thought to be a predictor of psychopathology (16). In addition, a psychological survey of 1906 college freshmen showed that the combined usage of HTP and SCL-90 increased the accuracy of screening for mental problems (17). HTP has also been found to be an effective tool for classifying personality disorders, depression, anxiety, and other mental disorders (18–20).

However, there are some shortcomings in the existing studies. One is that the scoring and interpretation of the HTP test are not standardized and lack consistency. The drawing characteristics selected by researchers are subjective, which makes it difficult to compare the results of different studies (21, 22). Moreover, the interpretations of some drawing characteristics are inconsistent. For instance, some researchers believe that drawing a "chimney" is a negative expression of family or internal conflict (23), while others believe that it is a

positive characteristic indicating open communication channels with the outside world and the seeking of support and warmth (15, 24).

Although the above issues have received extensive attention from researchers, most of them have been presented in systematic reviews or research prospects (25–27). It is difficult to solve the problem through a theoretical summary alone. Therefore, we will integrate the drawing indicators of the HTP test of mental disorders through meta-analysis. Specifically, this study will answer the following three questions: (1) Which drawing characteristics have been frequently selected as screening indicators for mental disorders in previous studies? (2) How well do these drawing characteristics predict mental disorders? (3) Are there any differences in the predictive effects of the same drawing characteristics for affective-type disorders and thought-type disorders?

Materials and methods

This study was conducted according to the Preferred Reporting Items for a Systematic Review and Meta-analysis (PRISMA) statement (28).

Search strategy

To obtain studies to be used in the analysis, four English (PubMed, Web of Science, Embase, and EBSCO) and three Chinese (China National Knowledge Infrastructure (CNKI), Chinese Scientific Journal Database (VIP), and Wanfang) databases were utilized. The core search terms were "house-tree-person," "HTP test," "S-HTP," "K-HTP," "projective test," and "drawing test." The search period was initially from 1 January 1948 to 20 May 2022. Further details of the search strategy are displayed in the **Supplementary material**.

Inclusion and exclusion criteria

For the retrieved literature, the inclusion or exclusion of meta-analysis was further judged according to the following criteria: (1) published empirical studies of the HTP test of mental disorders, excluding purely theoretical and literature review articles, were included; (2) the included studies distinguished between subjects with and without mental disorders using recognized and credible scales; (3) the included studies contrasted participants with mental disorders with those without mental disorders, and studies that focused only on those who had mental disorders were excluded; (4) the included studies provided specific data on the occurrence frequency of drawing characteristics, and studies where the original data were calculated in other forms or where the effect sizes could not be converted were excluded; and (5) all duplicate publications were excluded.

Screening and checking of the literature were performed independently by two investigators (HBG and BF) based on the inclusion and exclusion criteria, and the consensus was negotiated in case of discrepancies. The final review was conducted by the corresponding author (TLC). A total of 1,498 potentially relevant studies were identified in different databases through the search strategy, with an initial screening achieved by scanning the titles and abstracts, followed by a full-text screening, resulting in the inclusion of 30 studies. The literature search and screening process is shown in **Figure 1**.

Quality assessment

Quality was assessed using the Cross-Sectional Study Quality Assessment Forms (CSSQAF) recommended by the Agency for Healthcare Research and Quality. The form has 11 items, which receive a score of 0 if the result is "no" or "unclear" and 1 if the answer is "yes." Publications with scores of 8-11, 4-7, and 0-3 were considered high-quality literature, moderate-quality literature, and low-quality literature, respectively. Two investigators (HBG and BF) independently rated the included literature and calculated the rater agreement coefficient, which was found to be good, with a Kappa value of 0.85.

Coding procedures

As various studies used different names to describe the drawing characteristics, we standardized and unified the names. Three different naming principles were adopted according to the following cases: (1) when the same meaning but different wording was used, for example, *the house, tree, and person are clearly spaced,* and *excessive separation among items*; we adopted the expressions more frequently used by the predecessors; (2) when the meaning was the same but different directions were used, for example, *roots* and *no roots of trees*, the expressions more often by the predecessors were retained, and the opposite characteristics were scored in reverse; (3) when the meanings were similar but different wording was used, for example, *paintings without additional objects*, *no flowers and grass*, *painting without clouds*, as summary expression was utilized, such as *no additional decoration*. It should be noted that such characteristics should be carefully categorized. This process was carried out independently by two researchers (HBG and BF), and after completion, the agreement was reached after deliberation and discussion. In case of dispute, it was consulted by a third researcher (HYF) to resolve the issue.

In addition, the translation procedures used to translate Chinese drawing characteristics were as follows. First, two

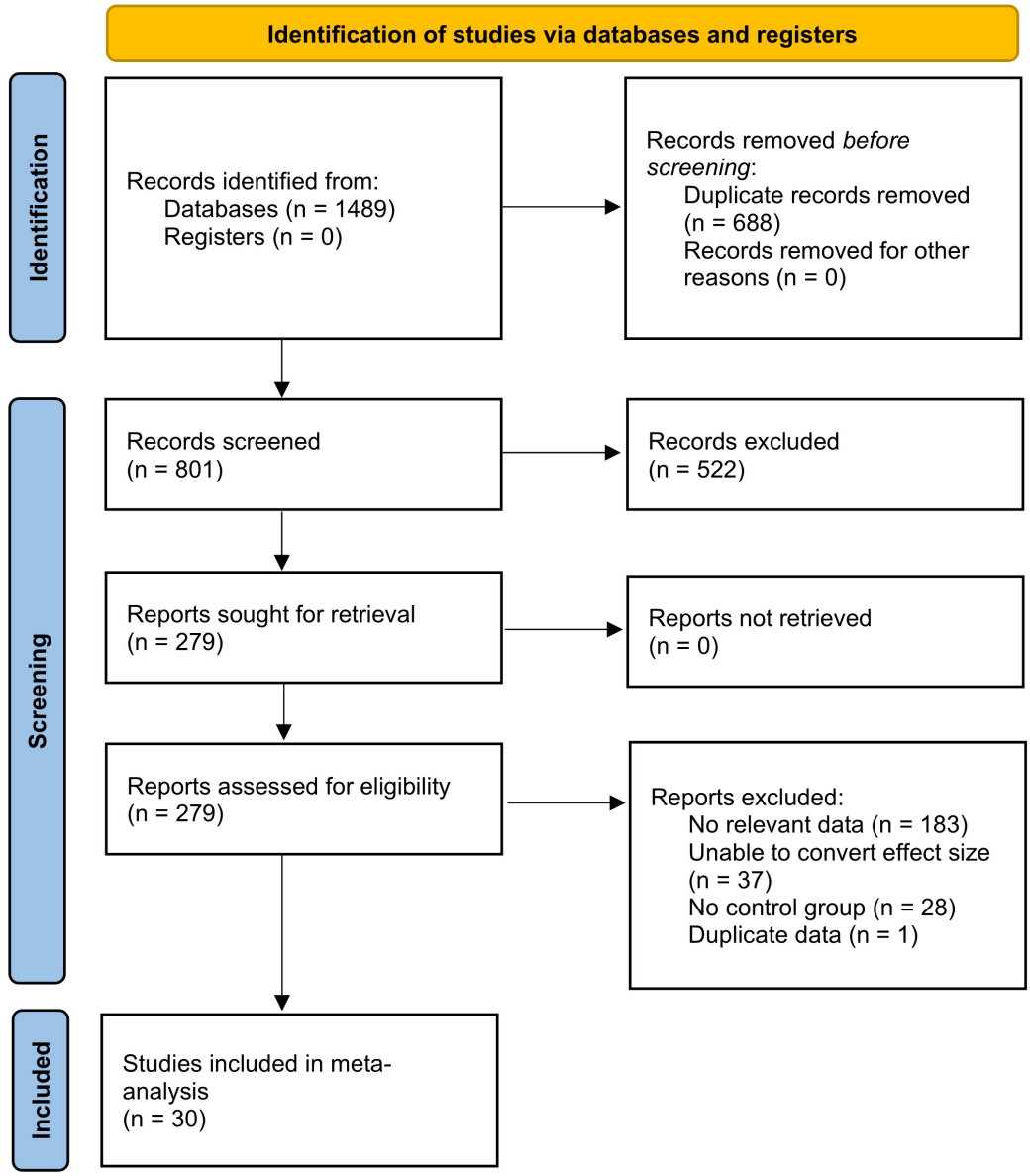


FIGURE 1
Flow chart of the identified studies.

researchers (HBG and BF) independently translated the drawing characteristics into English, then discussed differences and merged them into version 1. Second, one researcher (ZQD) modified the grammar as well as the words and formed version 2. Third, back-translation was performed by two other researchers (HYF and YQM) and the translation was modified accordingly to ensure accuracy. Finally, the final version was formed by considering the three previous coding principles and maintaining consistency with English characteristic names. After completion, the corresponding author (TLC) reviewed it. Discussions and revisions continued if there was disagreement until all researchers reached a consensus.

Statistical analysis

The strength of the association between drawing characteristics of the HTP test and the prevalence of mental disorders was measured by the ratio (OR) with a *95% CI*. Significance was determined by the *Z*-test, and $p < 0.05$ was considered statistically significant. The included literature was tested for heterogeneity and evaluated comprehensively using the Q test and the I^2 statistic according to the Cochrane Handbook for Systematic Reviews of Interventions. The Q test obeys a Chi-square test distribution, and when $Q \leq 0.10$, the heterogeneity test is considered statistically significant, while I^2

reflects the proportion of between-study variation attributable to heterogeneity, rather than random error or chance. I^2 values of 25, 50, and 75% represent low, moderate, and high heterogeneity, respectively, and a random effects model is more appropriate when heterogeneity is high (29).

Publication bias

Publication bias was assessed using a funnel plot, Rosenthal's fail-safe number (N_{fs}), and the trim and fill method. If the effect values were concentrated at the top of the funnel plot and clustered roughly symmetrically around the mean, there was no publication bias. In addition, the larger the fail-safe number is, the less likely bias is, which means that it is less likely that the conclusions will be overturned. If $N_{fs} < 5k + 10$ (k is the original number of studies), publication bias should be considered (30). The trim and fill methods distribute the studies as symmetrically as possible on the left and right sides of the mean effect size by first cutting and then complementing and re-estimating the true value of the combined effect size (31). If the effect size does not change significantly after cutting and complementing, then publication bias can be considered not to exist. All statistics for this meta-analysis were calculated by CMA 3.0 software.

Results

Study selection and characteristics

A total of 30 studies were included in this meta-analysis, all of which were cross-sectional studies, including 10 in English and 20 in Chinese. A total of 665 independent effect sizes were included, and 6,295 subjects participated in the survey. The results of the quality assessment showed that 19 of the publication included in this study scored between 7 and 9, which indicates high quality, and 11 scored between 4 and 6, which indicates moderate quality. For the studies included in the meta-analysis, the following information was extracted: (1) first author and time of publication; (2) version of the HTP test used for the study; (3) total sample size, including the number of subjects in the mental disorder and control groups; and (4) the type of mental disorder and screening tool. The results are shown in **Table 1**.

Predictive effect of mental disorders

Of the 30 included studies, 341 different drawing characteristics of the HTP drawing test were found; of which, 5 appeared more than or equal to 10 times, 20 appeared 5 to 10 times, 25 appeared 3 to 5 times, and 289 appeared 1 to 3 times. A total of 50 drawing characteristics with a high frequency (more than or equal to 3 times) were selected for inclusion in the analysis to explore their validity in screening for mental disorders (32). The HTP test can be divided into the following four dimensions that reflect the structure, size, and other characteristics of the picture: house, tree, person, and the whole drawing. The predictive effects of the four dimensions regarding mental disorders were in the following order: the whole drawing (OR = 4.20, $p < 0.001$) had the highest effect, followed by the house drawing (OR = 3.95, $p < 0.001$), the tree drawing (OR = 2.70, $p < 0.001$), and finally, the person drawing (OR = 2.16, $p < 0.001$). The predictive effects of specific drawing characteristics for each dimension are shown in **Table 2**.

Whole drawing characteristics

Of the 15 whole drawing characteristics, 12 were significant predictors of mental disorders. The significant characteristics in order of OR size were the *emphasis on straight lines* (OR = 11.75, $p = 0.004$), *simplified drawing* (OR = 9.64, $p < 0.001$), *no theme* (OR = 9.36, $p < 0.001$), *small drawing size* (OR = 5.71, $p < 0.001$), *excessive separation among items* (OR = 3.84, $p < 0.001$), *weak or intermittent lines* (OR = 3.19, $p < 0.001$), *no motion* (OR = 2.96, $p = 0.003$), *shadow* (OR = 2.88, $p = 0.006$), *omitted house, tree or person* (OR = 2.81, $p = 0.001$), *shaded or blackened drawing* (OR = 2.72, $p < 0.001$), *no additional decoration* (OR = 2.59, $p = 0.041$), and *scribbled drawing* (OR = 2.56, $p < 0.001$). In contrast, characteristics such as *emphasizing the horizon*, *weighted or repeated lines*, and the *fence* did not significantly predict mental disorders ($p > 0.05$).

House drawing characteristics

Of note, 9 of the 12 house drawing characteristics that significantly predicted mental disorders were ranked according to OR: *bizarre house* (OR = 4.64, $p < 0.01$), *no door* (OR = 4.52, $p < 0.001$), *very small house* (OR = 4.24, $p < 0.001$), *no window* (OR = 43.09, $p < 0.01$), *leaning house* (OR = 2.68, $p = 0.01$), *decorated roof* (OR = 2.32, $p = 0.006$), *smoking chimney* (OR = 2.27, $p = 0.001$), *two-dimensional house* (OR = 1.76, $p < 0.001$), and *shaded or blackened wall* (OR = 1.66, $p = 0.044$). However, *chimney* and *closed door or window* were not statistically significant in predicting mental disorders ($p > 0.05$).

Tree drawing characteristics

Of the 9 tree drawing characteristics, 7 were significant predictors for mental disorders. The significant characteristics in order of OR were as follows: *roots* (OR = 4.35, $p < 0.001$), *truncated tree* (OR = 2.90, $p < 0.001$), *flattened crown*

10.3389/fpsyt.2022.1041770

TABLE 1 Studies included in the meta-analysis.

Author	Year	HTP type	Sample size	Disease group	Control group	Mental disorders	Scales	Score
Chen	2015	S-HTP	60	30	30	Schizophrenia	SCL-90, BPRS	8
Chen	2015	S-HTP	562	38	524	Dependent personality disorder	PDQ-4+	7
Deng	2014	S-HTP	64	32	32	Schizophrenia	BPRS	8
Deng	2017	S-HTP	60	30	30	Depression	–	4
Eisel	1978	HTP	138	69	69	Schizophrenia	DSM	8
Fukunishi	2002	S-HTP	192	50	142	Alexithymia	TAS-20	7
Kirchner	1974	HTP	195	49	146	Substance addiction disorder	–	4
Koide	1992	HTP	126	16	110	Organic mental disorders	–	5
Kwark	2017	S-HTP	100	50	50	Schizophrenia	–	5
Lee	2019	S-HTP	186	23	163	Depression	EPQ, PHQ-9	6
Lee	2020	S-HTP	186	60	126	Substance addiction disorder	NDSS	6
Li	2014	S-HTP	105	35	70	High-functioning-autism	DSM	8
Li	2016	S-HTP	65	30	35	Depression	HAMD	4
Li	2020	S-HTP	324	190	134	Anxiety	MSSMHS	9
Li	2021	S-HTP	60	30	30	Depression	SCL-90	5
Ning	2015	S-HTP	676	148	528	Depression	CDI	8
Sheng	2019	S-HTP	167	27	140	Anxiety	SAS	6
Wang	2007	S-HTP	55	25	30	Mental disease	SCL-90	7
Wang	2017	S-HTP	177	74	103	Anxiety	MHT	6
Wang	2019	S-HTP	71	-	-	Anxiety, depression, paranoia	SCL-90	8
Xiang	2020a	HTP	358	22	336	Attention deficit/hyperactivity disorder	CBCL	7
Xiang	2020b	HTP	358	68	290	Depression	CBCL	7
Xie	1994	S-HTP	220	110	110	Schizophrenia	–	5
Yan	2014	S-HTP	540	277	263	Depression	SDS	8
Yang	2019	S-HTP	167	57	110	Depression	SDS	9
Zhao	2015	HTP	170	37	133	Somatization disorder	CSI, CBCL	8
Zhou	2019	S-HTP	39	17	22	Schizophrenia	SAPS, SANS	7
Zhou	2021	S-HTP	200	100	100	Rumination	RRS	9
Zhu	2011	S-HTP	112	59	53	Post-traumatic stress disorder	PCL-C	8
Zhu	2020	S-HTP	562	140	422	Narcissistic personality disorder	PDQ-4+	7

SCL-90, symptom checklist 90; BPRS, brief psychiatric rating scale; PHQ-9, patient health questionnaire-9 items; DSM, diagnostic and statistical manual of mental disorders; TAS-20, Toronto Alexithymia scale; EPQ, Eysenck personality questionnaire; NDSS, nicotine dependence syndrome scale; HAMD, Hamilton depression scale; MSSMHS, middle school student mental health scale; CDI, children's depression inventory; SAS, self-rating anxiety scale; MHT, mental health test; CBCL, Achenbach Child Behavior Checklist; SDS, self-rating depression scale; CSI, children's somatization inventory; SAPS, scale for assessment of positive symptoms; SANS, scale for assessment of negative symptoms; RRS, ruminative responses scale; PCL-C, PTSD checklist-civilian version; PDQ-4+, personality diagnostic questionnaire-4+.

(OR = 2.82, $p < 0.001$), *bizarre tree* (OR = 2.78, $p < 0.001$), *dead tree* (OR = 2.67, $p < 0.001$), *very small tree* (OR = 2.65, $p = 0.002$), and *sharp branch* (OR = 2.35, $p < 0.001$). In contrast, *scars of trees* and *shaded or blackened trees* were not significant predictors of mental disorders ($p > 0.05$).

Person drawing characteristics

Notably, 11 of the 14 person drawing characteristics were significant predictors for mental disorder, ranked according to OR as follows: *incomplete person* (OR = 4.90, $p < 0.001$), *shaded or blackened person* (OR = 4.63, $p = 0.010$), *fist* (OR = 3.66, $p = 0.001$), *negative expression* (OR = 3.59, $p < 0.001$), *bizarre person* (OR = 3.18, $p = 0.003$), *very small person* (OR = 3.02, $p < 0.001$), *loss of facial features* (OR = 2.71, $p = 0.002$), *poker face* (OR = 2.09, $p < 0.001$), *inappropriate body proportions* (OR = 1.99, $p < 0.001$), *single line limbs* (OR = 1.93, $p = 0.001$), and *complete or partial loss of limbs* (OR = 1.82, $p = 0.001$). However, *simple person*, *fingers,* and *not frontal portrait* did not significantly predict mental disorders ($p > 0.05$).

Subgroup analysis

As Table 2 shows, there was very high heterogeneity ($I^2 > 75\%$) in 11 drawing characteristics, and 6 characteristics with I^2 values between 70 and 75% also had high heterogeneity. We conducted a subgroup analysis of the above characteristics. The results are shown in Table 3.

TABLE 2 The predictive effect of drawing characteristics on mental disorders.

	Drawing characteristics	*k*	Heterogeneity		OR	95% CI	*P*	N_{fs}
			Q(p)	I^2(%)				
Whole	No additional decoration	15	0.000	93.70	2.59	1.25∼10.29	0.041	197
12 items	Excessive separation among items	10	0.000	88.36	3.84	1.95∼7.56	0.000	199
	Simplified drawing	6	0.001	76.15	9.64	4.08∼22.75	0.000	170
	Weak or intermittent lines	5	0.196	33.79	3.19	2.03∼5.01	0.000	35
	No motion	5	0.000	85.41	2.96	1.46∼6.00	0.003	63
	Omitted house, tree or person	5	1.174	33.61	2.81	1.52∼5.18	0.001	14
	Small drawing size	4	0.182	38.32	5.71	3.37∼9.68	0.000	43
	Shaded or blackened drawing	4	0.013	71.99	2.72	1.60∼4.62	0.000	17
	Scribbled drawing	4	0.117	49.12	2.56	1.52∼4.32	0.000	13
	Emphasis on straight lines	3	0.000	88.19	11.75	2.16∼64.02	0.004	41
	No theme	3	0.805	0.00	9.36	3.74∼23.4	0.000	14
	Shadow	3	0.411	0.00	2.88	1.36∼6.11	0.006	5
	Total dimensional score	–	0.000	88.47	4.20	3.06∼5.77	0.000	–
House	Very small house	7	0.000	88.88	4.24	1.91∼9.44	0.000	193
9 items	No door	6	0.035	58.25	4.52	2.96∼6.92	0.000	71
	No window	6	0.001	74.89	3.09	2.02∼4.72	0.000	51
	Decorated roof	6	0.000	75.25	2.32	1.27∼4.25	0.006	65
	Leaning house	5	0.002	76.32	2.68	1.49∼4.81	0.001	52
	Two-dimensional house	5	0.003	74.48	1.76	1.38∼2.24	0.000	23
	Smoking chimney	4	0.078	55.96	2.27	1.43∼3.61	0.001	14
	Shaded or blackened wall	4	0.014	71.57	1.66	1.01∼2.71	0.044	0
	Bizarre house	3	0.092	58.12	4.64	2.56∼8.40	0.000	19
	Total dimensional score	–	0.000	76.94	3.09	2.42∼3.95	0.000	–
Tree	Very small tree	11	0.000	82.91	2.65	1.41∼4.97	0.002	123
7 items	Roots	6	0.139	39.93	4.35	2.96∼6.39	0.000	70
	Truncated tree	6	0.003	72.67	2.90	1.62∼5.18	0.000	24
	Sharp branch	6	0.007	68.60	2.35	1.60∼3.46	0.000	6
	Bizarre tree	4	0.025	67.94	2.78	1.91∼4.04	0.000	31
	Dead tree	4	0.362	6.14	2.67	1.59∼4.47	0.000	14
	Flattened crown	3	0.038	69.54	2.82	1.91∼4.17	0.000	13
	Total dimensional score	–	0.000	69.76	2.70	2.34∼3.11	0.000	–
Person	Loss of facial features	10	0.000	82.86	2.71	1.46∼5.04	0.002	94
11 items	Shaded or blackened person	5	0.002	76.79	4.63	1.45∼14.85	0.010	27
	Poker face	5	0.150	40.65	2.09	1.40∼3.12	0.000	16
	Inappropriate body proportions	5	0.006	72.10	1.99	1.37∼2.88	0.000	21
	Single line limbs	5	0.117	45.77	1.93	1.32∼2.81	0.001	18
	Negative expression	4	0.300	18.15	3.59	1.96∼6.59	0.000	8
	Bizarre person	4	0.105	51.16	3.18	1.49∼6.77	0.003	6
	Complete or partial loss of limbs	4	0.084	54.92	1.82	1.26∼2.63	0.001	6
	Incomplete person	3	0.139	49.30	4.90	3.05∼7.88	0.000	30
	Very small person	3	0.094	57.64	3.02	2.04∼4.45	0.000	25
	Fist	3	0.972	0.00	3.66	1.70∼7.85	0.001	6
	Total dimensional score	–	0.000	66.53	2.16	2.22∼3.49	0.000	–

In the subgroup analysis, mental disorders were classified into affective-type disorders (depression and anxiety) and thought-type disorders (schizophrenia, paranoia, and rumination). Drawing characteristics that appeared more than twice in both disorders (12 items in total) were extracted based on the number of studies after classification. The results showed that some drawing characteristics were significant predictors of affective-type disorders, but not of thought-type disorders, including *no motion* (OR = 3.34, $p = 0.019$), *leaning house* (OR = 2.13, $p < 0.001$), and *decorated roof*

TABLE 3 Subgroup analysis of mental disorder types.

	Drawing characteristics	Type	*k*	Heterogeneity		OR	95% CI	*P*
				Q(p)	I^2(%)			
ASI	No motion	AD	3	0.000	86.99	3.34	1.22∼9.16	0.019
		TD	2	0.001	91.06	2.63	0.63∼11.02	0.185
	Leaning house	AD	2	0.082	66.96	2.13	1.48∼3.07	0.000
		TD	2	0.002	89.06	3.77	0.87∼16.43	0.077
	Decorated roof	AD	3	0.000	85.16	2.49	1.50∼4.13	0.000
		TD	2	0.179	44.63	2.00	0.70∼5.77	0.197
TSI	Excessive separation among items	AD	4	0.000	92.49	2.47	0.79∼7.71	0.119
		TD	4	0.000	87.21	6.09	1.40∼26.41	0.016
	No window	AD	2	0.019	81.68	3.14	0.24∼41.20	0.385
		TD	3	0.928	0.00	6.41	3.53∼11.65	0.000
	Loss of facial features	AD	5	0.000	88.61	1.81	0.75∼4.35	0.185
		TD	3	0.025	72.82	2.60	1.62∼4.16	0.000
	Inappropriate body proportions	AD	2	0.865	0.00	1.29	0.71∼2.33	0.398
		TD	2	0.473	0.00	9.29	3.77∼22.90	0.000
MDC	No additional decoration	AD	6	0.000	94.23	1.38	0.45∼4.19	0.000
		TD	5	0.004	74.40	14.19	8.72∼23.08	0.000
	Simplified drawing	AD	2	0.000	94.60	13.06	1.12∼152.54	0.041
		TD	3	0.424	0.00	7.23	3.66∼14.30	0.000
	Very small house	AD	3	0.000	95.75	5.29	1.16∼24.21	0.032
		TD	3	0.149	47.44	3.87	2.09∼7.16	0.000
	Two-dimensional house	AD	2	0.024	80.40	3.00	0.99∼9.14	0.050
		TD	2	0.257	22.14	2.08	1.30∼3.36	0.002
	Very small tree	AD	6	0.003	72.01	2.70	1.96∼3.72	0.000
		TD	4	0.128	47.27	3.92	2.34∼6.59	0.000

ASI, affect-specific indicators; TSI, thought-specific indicators; MDC, mental disorders coindicators; AD, affective-type disorders; TD, thought-type disorders.

(OR = 2.49, $p < 0.001$), which could be known as affective-specific indicators. Conversely, drawing characteristics that significantly predicted thought-type disorders, but not affective-type disorders, included *excessive separation among items* (OR = 6.09, $p = 0.016$), *no window* (OR = 6.41, $p < 0.001$), *loss of facial features* (OR = 2.60, $p < 0.001$), and *inappropriate body proportions* (OR = 9.29, $p < 0.001$), which are thought-specific indicators. Furthermore, *no additional decoration*, *simplified drawing*, *very small house*, *two-dimensional house*, and *very small tree* were significant predictors of both mental disorders ($p < 0.01$) and could be described as mental disorder coindicators.

Analysis of publication bias

The funnel plot (Figure 2) showed that most of the effect sizes were located at the top of the funnel plot and were largely evenly clustered on either side of the mean effect values. It can be preliminarily judged that the possibility of publication bias in this meta-analysis is low. However, since the funnel plot evaluation was relatively subjective, the publication bias of each drawing characteristic was further judged according to the loss of safety factor (N_{fs}), and the results are shown in Table 2.

There was no publication bias, and the conclusion was more reliable for the drawing characteristics with $N_{fs} > 5k + 10$. For the painting features that did not meet this criterion, the results were further examined by the trim and fill method and are shown in Table 4. All drawing characteristics showed significant effect sizes except for the *shaded or blackened person* characteristic, which can be considered not to have significant publication bias. The significance of this characteristic should be interpreted with caution, probably due to the small number of published studies or the small effect size.

Discussion

Projection theory suggests that drawing is an expression of the subconscious, and the size and other characteristics of the drawing reflect an individual's emotions, personality, etc. (21). Many studies have demonstrated the use of the HTP drawing test to screen for mental disorders. However, there was a serious lack of consistency in the drawing characteristics

FIGURE 2
Funnel plot for the study of the relationship between drawing characteristics and mental disorders.

selected in previous studies, which made it difficult to compare the different findings. Additionally, the predictive effects of some drawing characteristics were inconsistent. In this study, we found through meta-analysis that the predictive effects of the HTP test's four dimensions on mental disorders were in the following order: the effect of the whole drawing was the greatest, followed by house drawing, tree drawing, and person drawing. Furthermore, we focused on integrating drawing characteristics that were used more frequently in previous studies and identified 39 significant predictors of mental disorders.

Psychodynamic theory suggests that behavior is driven or motivated by internal forces, with a focus on the unconscious, defense mechanisms, projections, etc. (33). Referring to this, HTP drawing characteristics can be categorized into the following four types: *Item absence*, *bizarre or twisted*, *excessive details,* and *small or simplified*. First, *item absence* reflects the loss of self-awareness, or strong psychological defenses, and can be thought of as an individual's repression of self. Second, *bizarre or twisted* implies psychological conflict or a sense of unreality that inner and external environments are inconsistent. Third, *excessive details* suggest that internal conflicts have led to obvious anxiety, which manifests as nervousness, sensitivity, and irritability. Finally, *small or simplified* reflects avoidance and retreat due to low mental motivation and energy. In the following, we discuss the drawing characteristics of each type and summarize them in **Table 5**.

Drawing characteristics of the house-tree-person test

Whole drawing characteristics

We found that the whole drawing characteristics were the best predictors of mental disorders, and the significant specific characteristics were as follows: (1) Item absence: this characteristic represents the absence of something in the picture that should be included, and two drawing characteristics are included, namely, *omitted house, tree, or person,* and *excessive separation among items.* The present study found that the absence of a house, tree, or person, or distance between them was a significant predictor of a mental disorder, which is consistent with the results of previous studies (23, 27). A drawing in which the house, tree, and person are complete and at appropriate distances reflects regularity and high personal reality satisfaction (34). In contrast, the absence of items from the whole drawing indicates strong defensiveness or lack of social support. (2) Excessive details: this characteristic indicates that some unnecessary characteristics have been drawn, including *shaded or blackened drawings* and *shadows*. Jung highlighted that shadows represent the hidden or unconscious psychology within the individual, and the presence of shadows and blackening indicates that the illustrator is autistic and or is experiencing (35). We found that the drawing

TABLE 4 Analysis of publication bias.

	Drawing characteristics	Trim and fill imputed studies	OR	Adjusted OR	95% CI
Whole	Omitted house, tree or person	3	2.81	1.87	1.09∼3.20
	Shaded or blackened drawing	1	2.72	2.43	1.44∼4.08
	Scribbled drawing	2	2.56	1.71	1.08∼2.71
	No theme	1	9.36	8.62	3.59∼20.69
	Shadow	0	2.88	2.88	1.36∼6.11
House	Two-dimensional house	1	1.76	1.61	1.28∼2.04
	Smoking chimney	2	2.27	1.73	1.13∼2.64
	Shaded or blackened wall	0	1.66	1.66	1.01∼2.71
	Bizarre house	0	4.64	4.64	2.56∼8.40
Tree	Truncated tree	2	2.90	2.13	1.23∼3.71
	Sharp branch	1	2.35	2.18	1.50∼3.18
	Dead tree	2	2.67	2.36	1.44∼3.88
	Flattened crown	0	2.82	2.82	1.91∼4.17
Person	Shaded or blackened person	3	4.63	1.32	0.95∼1.83
	Poker face	2	2.09	1.65	1.14∼2.39
	Inappropriate body proportions	1	1.99	1.77	1.23∼2.54
	Single line limbs	2	1.93	1.73	1.20∼2.51
	Negative expression	1	3.59	3.34	1.97∼5.67
	Bizarre person	1	3.18	2.75	1.32∼5.72
	Complete or partial loss of limbs	0	1.82	1.82	1.26∼2.63
	Fist	0	3.66	3.66	1.70∼7.85

of shadows or shading was a significant predictive feature of a mental disorder, which is consistent with many previous studies (15, 36). Thus, the inclusion of excessive details in the whole drawing is an important indicator of inner anxiety. (3) Small or simplified: in this whole drawing characteristic, the picture is drab and meaningless, and specific drawing characteristics include *no additional decoration*, *simplified drawing*, *no motion*, *no theme*, *small drawing size*, *weak or intermittent lines*, *emphasis on straight lines*, and *scribbled drawing*. Previous researchers have paid more attention to the decorations of the drawing. *No additional decoration* other than the house, tree, and person usually represents low psychological energy and a lack of enthusiasm and motivation in life. A study of schizophrenia supported this view and found a significant enrichment of drawings after the patients were treated (37). Both the *no motion* and *no theme* characteristics predict mental disorders. Previous studies have also found that the drawings of depressed patients are more likely to lack emotion and theme (38). In addition, the size of the picture is usually related to the self-awareness and psychological state of the painter. A *small drawing size* indicates that the subject may have a low self-evaluation or be insecure (19). Moreover, *weak or intermittent lines* often suggest indecision, as well as unclear self-awareness and emotional tendencies (39), and are more likely to be reflected in patients with mental disorders. Thus, a small or simplified whole drawing reflects low mental energy, avoidance, and withdrawal.

House drawing characteristics

By analyzing the size of the house, windows and doors, the floor, etc., the family atmosphere, self-image, and interpersonal status of the illustrator can be revealed (40). In this study, we found that the significant house drawing characteristics are as follows: (1) Item absence: this characteristic means that the house is missing necessary features, such as *no door* and *no window*. Doors and windows are channels of contact with the outside world, and *no door* suggests strong defensiveness (38), corresponding to the self-closure and refusal to communicate in patients with mental disorders. Both *no door* and *no window* were found to be significant predictive characteristics of mental disorders in this meta-analysis. Chen (37) also noted that there is a significant increase in doors and windows after schizophrenic patients receive treatment. Thus, the absence of items in the house drawing indicates an individual's autism and defensiveness. (2) Bizarre or distortion: this characteristic means that the shape or features of the house deviate from reality, such as a *leaning house* or a *bizarre house*. *A leaning*

TABLE 5 Characteristics and implications of effective predictive drawing for mental disorders.

Type	Drawing characteristics	Indicates meaning
Item absence	Excessive separation among items, omitted house, tree or person, no door, no window, loss of facial features, poker face, complete or partial loss of limbs, and incomplete person	Loss of self-awareness and psychological defenses
Bizarre or distortion	Leaning house, bizarre house, truncated tree, dead tree, bizarre tree, sharp branch, flattened crown, bizarre person, inappropriate body proportions, and fist	Psychological conflict and sense of unreality
Excessive details	Shaded or blackened drawing, shadow, decorated roof, smoking chimney, shaded or blackened wall, roots, shaded or blackened person, and negative expression	Nervousness, sensitivity, and irritability
Small or simplified	No additional decoration, simplified drawing, no motion, no theme, small drawing size, weak or intermittent lines, emphasis on straight lines, scribbled drawing, very small house, two-dimensional house, very small tree, very small person, and single line limbs	Low mental motivation, avoidance and retreat

house suggests unbearable stress and can significantly predict a mental disorder (23). Some researchers have found that *bizarre houses* (e.g., churches, temples, and pavilions) are also more likely to appear in the paintings of schizophrenic patients (41). Thus, bizarre houses or distortion of house drawing reflects inner repression and escape from reality. (3) Excessive details: this characteristic represents an excessive house depiction, including the following specific characteristics: *decorated roof*, *smoking chimney*, *shaded, or blackened walls*. Some researchers have argued that individuals with high activity levels usually create more meticulous drawings, and the opposite is true for individuals with low activity levels, such as those suffering from depression (42); however, others have suggested that detailed delineation represents neuroticism, sensitivity, and irritability (36, 38). The results of this meta-analysis supported the latter, showing that a *decorated roof* and *walls that are shaded or painted black* were both significant predictors of mental disorders. In addition, a *smoking chimney* indicates that the subject is experiencing family conflict, anxiety, and tension (37) and has a positive predictive effect regarding mental disorders. Therefore, excessive details of a house drawing reflect an individual's concern for family and the apparent anxiety. (4) Small or simplified: this house drawing characteristic indicates that the house drawing is too simple or flat, including *very small houses* and *two-dimensional houses*. The house size usually represents the family relationship and status of the artist, and v*ery small houses* are mostly seen in families with low intimacy and prominent conflicts (37). Deng (39) found that 84.4% of the schizophrenia group painted houses that were too small, which was significantly higher than that of the normal group (34.4%). *A two-dimensional house* appears monotonous and lacks dimensionality, which usually reflects introverted and withdrawn personalities and is more likely to appear in the drawings of depressed individuals (43). Thus, a small or simplified house drawing reflects low security and poor intimacy.

Tree drawing characteristics

Many projective tests have used tree imagery as a theme; in addition to the HTP test, a common test using this theme is the tree test (44). Tree imagery often reflects emotional experiences related to growth and can reflect the relationship between an individual's subjective feelings and the external environment (21). The results showed that the significant characteristics of a tree drawing include the following: (1) Bizarre or distortion: this tree imagery has characteristics that are different from usual, including *truncated trees*, *sharp branches*, *bizarre trees*, *dead* trees, and *flattened crowns*. *Truncated trees* or *dead trees* often symbolize emotional indifference, lack of vitality, and loss of willingness to live (45, 46) and can significantly predict mental disorders. Hui (38) also found that *dead trees* emerged only in the depressed group. In addition, a *flattened canopy* indicates that external stress overwhelms subjects (26), which is supported by the results of this meta-analysis. *Sharp branches* are often thought to be associated with aggression and destructiveness. Chen (37) found that the percentage of *sharp branches* drawn by schizophrenic patients decreased from 37.7 to 6.7% once they received treatment. Therefore, bizarre or distorted tree drawings mainly reflect the unrealistic and aggressive traits of individuals. (2) Excessive details: this characteristic implies complex depictions of tree characteristics, such as *roots*. *Roots* indicate an immature mind and internal conflict (39), and the results of the meta-analysis demonstrate that the trait is one of the indicators of mental disorders. (3) Small or simplified: this tree drawing characteristic means that the tree imagery is too simple, and the significant characteristic is a *very small tree*. Tree imagery symbolizes lives and energy. Large trees represent vitality, while *very small trees* imply loneliness and a lack of self-confidence, which are more likely to appear in the paintings of patients with mental disorders (46).

Person drawing characteristics

The person's imagery often directly reflects the participant's self-concept (40). In addition to the HTP test, the human drawing test is also widely used in clinical assessment (47). We found that multiple drawing characteristics of a person could predict mental disorders, including the following: (1) Item absence: this characteristic means that the figure is drawn with incomplete characteristics such as facial features or limbs, including *an incomplete person*, *loss of facial features*, *poker face*, *complete or partial loss of limbs*. Machover (48) indicated that an *incomplete person* represents an incomplete self-image. If a part of the figure is omitted from the painting, this signals the loss of function of that part. *Complete or partial loss of limbs* also reflects the loss of self-awareness and even the lack of will to live in patients with mental disorders (27). Therefore, the absence of items in the drawing of a person means that the individual's self-awareness is weak or even lost. (2) Bizarre or distortion: this characteristic represents that the body is disproportionate or has uncommon features, such as *inappropriate body proportions*, *a bizarre person,* and the drawing of a *fist*. *A bizarre person* or *inappropriate body proportions* imply conflict between individuals and the external environment and are more likely to appear in the drawings of patients with mental disorders, consistent with many previous studies (49, 50). The drawing of a *fist* has a similar meaning to that of a sharp branch, indicating strong aggression and rebelliousness (13, 48), and is also a significant predictor of mental disorders. Thus, bizarre or distorted person's drawings reflect the individual's conflict or aggressiveness toward the external environment. (3) Excessive details: this characteristic represents that the figure drawing is depicted in unreasonable detail, such as a *shaded or blackened person* and *negative expression*. Researchers have argued that *shaded or blackened persons* imply the melancholy and depressed state of the painter (51), and the results of the meta-analysis support this view. In

addition, *negative expressions* (e.g., sadness and anger) tend to reflect negative emotions and are more likely to be expressed in persons with mental disorders (52). Therefore, excessive details in the person's drawing usually reflect an individual's negative emotions, such as depression and anxiety. (4) Small or simplified: this person drawing characteristic indicates a person drawing that is too small or oversimplified, and includes the following two significant characteristics: *very small person* and *single line limbs*. Figure size is important for explaining individual self-awareness, and a *very small person* symbolizes weak self-awareness and low mental energy in subjects (53) and appears in a much higher proportion of patients with mental disorders than in normal groups (34). *Single line limbs* mean that the figure drawing is overly simple and abstract; this characteristic almost exclusively occurs in patients with psychiatric disorders and is a significant predictive feature of disorders such as schizophrenia (41, 54). Thus, a small or simplified person drawing reflects weak self-awareness and low self-esteem.

Subgroup analysis

Furthermore, we know that there are differences in clinical symptoms between affective-type disorders and thought-type disorders. According to projective theory, it can be speculated that the differences would be reflected in the drawing characteristics. Therefore, we further explored the independent predictive characteristics of these two mental disorders through heterogeneity analysis. The results support the hypothesis, showing that some characteristics can only predict a specific type of mental disorder, while some characteristics have the same predictive effect for both types of mental disorders. We present the affective-specific indicators, thought-specific indicators, and common indicators separately below. These findings could provide a more practical reference for the screening and diagnosis of different types of mental disorders.

Affective-specific indicators included *no motion, leaning house,* and *decorated roof*. *No motion* is an important reflection of emptiness, reflecting a depressed mood and lack of mental motivation, which coincides with the clinical manifestations of depression. The results of previous comparative studies showed that the proportion of *no motion* was significantly higher in depressed patients than in the normal group (34, 55), but no significant difference was found in the comparison of individuals with schizophrenia and the normal group (27). The distorted characteristics represent a state of stress, and a *leaning house* suggested great stress in subjects. It was significantly reflected in individuals with both depression and anxiety disorders, appearing much more frequently in these groups than in the normal group (38, 56). Furthermore, meticulous drawings have been shown to represent sensitivity and irritability, coinciding with the clinical manifestations of anxiety disorders (15). Thus, a *decorated roof* was more frequently observed in the drawings of patients with affective-type disorders (57). Based on these findings, attention should be focused on distortion and excessive details in drawings when screening for affective-type disorders (e.g., depression and anxiety).

In addition, thought-specific indicators included *excessive separation among items*, *no window*, *loss of facial features,* and *inappropriate body proportions*. *Excessive separation among items* means that the house, tree, and person are separate and independent, which is more consistent with the broken and detached thinking of patients with thought-type disorders (e.g., schizophrenia). The results of the comparison study showed that this characteristic was only present in the schizophrenia group and not in the normal group (41, 58). However, no significant difference was found in the anxiety disorder group compared to the normal group (15). In addition, a comparison study found that 32.7% of patients with schizophrenia did not draw windows, and another 8.2% drew cutoff or odd windows, while 91.8% of the normal group drew regular windows (27). Relative to the normal group, *no window* is more likely to appear in the drawings of patients with schizophrenia, and there is a significant increase after treatment (37). However, there was no significant difference between the depressed and normal groups (50, 59). In addition, *loss of facial features* and *inappropriate body proportions* are more common due to the wild imaginations of thought-category disorder patients (60). Some patients may experience physical discomfort that is projected into their drawings. Many previous studies support this result, and found that schizophrenic patients were more likely to draw people with disproportionate head-to-body ratios, but no significant differences were found in another comparative study of depressed and normal individuals (27, 56, 61). Based on these findings, when screening for and diagnosing thought-type disorders (e.g., schizophrenia), focus should be placed on the obvious absence or excessive separation of drawing characteristics.

Common indicators of mental disorders included *no additional decoration*, *simplified drawings*, *very small houses*, *two-dimensional houses,* and *very small trees*. *Simplified drawings* without additional decoration have been proven to be significant predictors of mental disorders, implying that subjects are unresponsive and lack enthusiasm and motivation for life, which are typical symptoms of mental disorders. Many previous comparative studies on mental disorders such as depression and schizophrenia and normal individuals have found significant differences (19, 62). As mentioned previously, *very small or two-dimensional houses* and *very small trees* reflect the low psychological energy and insecurity of the subjects, and all appeared much more frequently in patients with depression, anxiety, and schizophrenia than in the normal group (34, 39, 41). Therefore, the common characteristics of mental disorders all reflect the lack of mental motivation.

Based on the above findings, oversimplified painting, small drawing size, and small imagery should be of concern regardless of which mental disorders are being screened for and diagnosed.

Strengths and limitations

This study is innovative in some ways. First, this paper innovatively integrates the characteristics of drawing in related studies since the development and application of HTP measurement through meta-analysis. This provides a reference standard for the selection of indicators in future HTP studies and offers the possibility for the development of objectification of the test (10, 12). In future studies, objectified indicators should be selected, and feature coding criteria should be formed to continuously promote the formation of an objectified HTP system. Second, this study found indicators specific to thought-type disorders and indicators specific to affective-type disorders and explored the theoretical implications, thus forming a theoretical guide for HTP testing. In the future, we should explore the predictive indicators of drawing for different psychological traits or mental disorders and continuously improve the theoretical guidance and application value of the HTP test. Finally, the predictive characteristics derived from this study can provide a basis for the screening and diagnosis of mental disorders, and their use in combination with the scale can improve the accuracy of mental disorder diagnosis (21). Meanwhile, the validity of the drawing characteristics of the HTP test needs to be continuously verified in clinical practice, which will in turn form an objective, complete, and valid predictor of mental disorders.

This study also has some shortcomings. First, it is still unclear whether subjects from other regions would have similar results because the included study population mainly originated from Asia, and only Chinese and English databases were searched. Second, it is difficult to classify drawing characteristics completely independent of each other when we encode them, resulting in overlapping meanings of certain characteristics. For example, *shaded or blackened drawings* contained *shaded or blackened persons*, while *incomplete persons* contained *complete or partial loss of limbs*. Attention should be given to the selection and interpretation of such drawing characteristics. Third, some of the drawing characteristics have been studied less often, which may have some influence on the accuracy of the results. More caution should be exercised, and more verification should be performed in interpreting these characteristics. Fourth, limited by the lack of basic information reported in the current literature, this study only explored the classification of mental disorders and could not explore the differences in gender and age. The analysis of the causes of drawing characteristics needs further depth. Finally, subgroup analysis can only explore two categories of mental disorders, and it is difficult to achieve a more refined classification, such as depression and anxiety, in affective-type disorders. It can be speculated that drawings with *missing or very small person* is more likely associated with depression, while *decorated roofs* are more consistent with anxiety. These results need to be further tested in future studies.

Conclusion

In this study, we found that the greatest predictor of mental disorders was the whole drawing characteristic, followed in order by house, tree, and person characteristics. The drawing characteristics that significantly predicted mental disorders can be grouped into the following four categories: item absence, bizarre or distortion, excessive details, and small or simplified. Moreover, subgroup analysis distinguished between affective-specific indicators, thought-specific indicators, and common indicators of mental disorders. The above findings can provide reference standards for the selection of drawing characteristics and provide theoretical guidance for the screening and clinical diagnosis of mental disorders.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Author contributions

HG and TC contributed to the conception and design of the study. BF organized the database and performed the statistical analysis. HG, TC, and BF wrote the draft of the manuscript. TC, YM, XZ, HF, and ZD reviewed and edited the manuscript. TC and QG supervised the study and acquired funding. All authors contributed to the manuscript revision and read and approved the submitted version.

Funding

This study was supported by the National Key R&D Program of China (Grant no. 2022YFC2009900), the National Natural Science Foundation of China (Grant nos. 81820108018 and 81401398), the Key R&D Program of Sichuan Province (Grant no. 2023YFS0076), the Sichuan Science and Technology Program (Grant no. 2019YJ0049), and the Sichuan Provincial Health and Family Planning Commission (Grant no. 19PJ080).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2022.1041770/full#supplementary-material

References

1. Stein DJ, Phillips KA, Bolton D, Fulford KW, Sadler JZ, Kendler KS. What is a mental/psychiatric disorder? From DSM-IV to DSM-V. *Psychol Med.* (2010) 40:1759–65. doi: 10.1017/s0033291709992261

2. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet.* (2013) 382:1575–86. doi: 10.1016/s0140-6736(13)61611-6

3. World Health Organization [WHO]. *Mental Disorders Fact Sheets*. Geneva: WHO (2019).

4. De Vaus J, Hornsey MJ, Kuppens P, Bastian B. Exploring the east-west divide in prevalence of affective disorder: a case for cultural differences in coping with negative emotion. *Pers Soc Psychol Rev.* (2018) 22:285–304. doi: 10.1177/1088868317736222

5. Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *J Abnorm Psychol.* (2017) 126:454–77. doi: 10.1037/abn0000258

6. Bech P, Bille J, Møller SB, Hellström LC, Østergaard SD. Psychometric validation of the Hopkins Symptom Checklist (SCL-90) subscales for depression, anxiety, and interpersonal sensitivity. *J Affect Disord.* (2014) 160:98–103. doi: 10.1016/j.jad.2013.12.005

7. Uher R, Payne JL, Pavlova B, Perlis RH. Major depressive disorder in DSM-5: implications for clinical practice and research of changes from DSM-IV. *Depress Anxiety.* (2014) 31:459–71. doi: 10.1002/da.22217

8. Lamiell J, Foss M, Cavenee P. On the relationship between conceptual schemes and behavior reports A closer look. *J Pers.* (2006) 48:54–73. doi: 10.1111/j.1467-6494.1980.tb00965.x

9. Shweder RA, D'Andrade RG. Accurate reflection or systematic distortion? A reply to Block, Weiss, and Thorne. *J Pers Soc Psychol.* (1979) 37:1075–84. doi: 10.1037/0022-3514.37.6.1075

10. Xiong M, Ye YT. The house-tree-person technique and its application in counseling. *J Jimei Univ.* (2012) 13:28–32. doi: 10.3969/j.issn.1671-6493.2012.01.006

11. Smeijsters H, Cleven G. The treatment of aggression using arts therapies in forensic psychiatry: results of a qualitative inquiry. *Arts Psychother.* (2006) 33:37–58. doi: 10.1016/j.aip.2005.07.001

12. Tong HJ. Surveying and expecting:three majar test technologies in psychology. *J Nanjing Normal Univ.* (2002) 3:81–8. doi: 10.3969/j.issn.1001-4608-B.2002.03.012

13. Buck JN. The H-T-P technique, a qualitative and quantitative scoring manual. *J Clin Psychol.* (1948) 4:317.

14. Camara W, Nathan J, Puente A. Psychological test usage: implications in professional psychology. *Prof Psychol.* (2000) 31:141–54. doi: 10.1037/0735-7028.31.2.141

15. Sheng L, Yang G, Pan Q, Xia C, Zhao L. Synthetic house-tree-person drawing test: a new method for screening anxiety in cancer patients. *J Oncol.* (2019) 2019:5062394. doi: 10.1155/2019/5062394

16. Michal-Smith H. The identification of pathological cerebral function through the H-T-P technique. *J Clin Psychol.* (1953) 9:293–5.

17. Li N, Zuo QS, Yang H, Xu RL. Application of HTP projective drawing test in psychological general survey for freshmen. *J Minzu Normal Univ Xingyi.* (2018) 6:26–31. doi: 10.3969/j.issn.1009-0673.2018.06.006

18. Dewaraja R, Sato H, Ogawa T. Anxiety in tsunami-affected children in Sri Lanka measured by Revised Children's Manifest Anxiety Scale and Synthetic House–Tree–Person Test. *Int Congr Series.* (2006) 1287:74–8. doi: 10.1016/j.ics.2005.12.035

19. Yang G, Zhao L, Sheng L. Association of synthetic house-tree-person drawing test and depression in cancer patients. *Biomed Res Int.* (2019) 2019:1478634. doi: 10.1155/2019/1478634

20. Chen K, Xu GX. A research on the diagnosis of depression through the projective drawing test. *J Psychol Sci.* (2008) 31:722–4. doi: 10.16719/j.cnki.1671-6981.2008.03.002

21. Cai W, Tang YL, Wu S, Chen ZZ. The tree in the projective tests. *Adv Psychol Sci.* (2012) 20:782–90.

22. Chen G, Yan WS. Utility of the Rorschach inkblot test in clinical psychological diagnosis. *China J Health Psychol.* (2022) 30:475–80. doi: 10.13342/j.cnki.cjhp.2022.03.031

23. Zhou HQ. *Research on the Relationship between Rumination Thinking of Junior Middle School Students and H-T-P Drawing Characteristics*. Jinzhou: Bohai University (2021).

24. Xiang JJ, Liao MS, Zhu MJ. Assessment of junior elementary pupils'depression tendency via House-Tree-Person test. *China J Health Psychol.* (2020) 28:1057–61. doi: 10.13342/j.cnki.cjhp.2020.07.023

25. Kato D, Suzuki M. Developing a scale to measure total impression of synthetic house-tree-person drawings. *Soc Behav Pers.* (2016) 44:19–28. doi: 10.2224/sbp.2016.44.1.19

26. Li JD, Fu HY. A study on characteristics of HTP in depression. *Psychol China.* (2021) 3:656–63. doi: 10.35534/pc.0306079r

27. Xie LY, Ye XH. Primary application of synthetic House-Tree-Person technique in china: a comparison of schizophrenics and normal controls. *Chin Mental Health J.* (1994) 8:250–2.

28. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* (2021) 372:n71.

29. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* (2003) 327:557–60. doi: 10.1136/bmj.327.7414.557

30. Rothstein HR, Sutton AJ, Borenstein M. Publication bias in meta-analysis. In: Borenstein M, Rothstein HR, Sutton AJ editors. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. England: John Wiley & Sons, Inc (2005). p. 1–7. doi: 10.1002/0470870168

31. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics.* (2000) 56:455–63. doi: 10.1111/j.0006-341x.2000.00455.x

32. Li X, Cao BD, Yang W, Qi JH, Liu J, Wang YF. Characteristic of the synthetic House-Tree-Person test in children with high-functioning autism. *Chin Mental Health J.* (2014) 28:260–6. doi: 10.3969/j.issn.1000-6729.2014.04.005

33. Gabbard GO. *Psychodynamic Psychiatry in Clinical Practice*. Washington, DC: American Psychiatric Pub (2014).

34. Ning SY, Zheng L, Li X, Hui WJ. A study on the application of the House Tree Person Test to assess depression in adolescents. *Chin J Clin Res.* (2015) 28:305–7. doi: 10.13429/j.cnki.cjcr.2015.03.011

35. Johnson DL, Johnson CA. Comparison of four intelligence tests used with culturally disadvantaged children. *Psychol Rep.* (1971) 28:209–10. doi: 10.2466/pr0.1971.28.1.209

36. Wang XY. The application of the Room Shuren Test in the psychological screening of junior high school freshmen. *Popular Psychol.* (2017) 1:24–5.

37. Chen LY. *The Study of Art Psychotherapy Effects on Psychiatric Rehabilitation*. Guangzhou: Guangzhou University of Chinese Medicine (2015).

38. Hui WJ. *Using Drawing Test to Assess the Thndency of Depression in Adolescence*. Harbin: Heilongjiang University of Chinese Medicine (2014).

39. Deng CY. *The Study of Correlation between Schizophrenics SHTP Test and BPRS*. Guangzhou: Guangzhou University of Chinese Medicine (2014).

40. Zhang Y. The value of HTP projective test in the psychological screening of new students. *Ideol Theoret Educ.* (2010) 5:70–3. doi: 10.16075/j.cnki.cn31-1220/g4.2010.05.007

41. Kwak YH, Lee KM. A comparative study on reactional characteristics of S-HTP between normal and schizophrenia patients. *Medicine.* (2010) 17:297–318. doi: 10.35594/kata.2010.17.2.007

42. Lange-Kuettner C, Kerzmann A, Heckhausen J. The emergence of visually realistic contour in the drawing of the human figure. *Br J Dev Psychol.* (2002) 20:439–63. doi: 10.1348/026151002320620415

43. Yan H, Yu HH, Chen JD. Application of the House-tree-person test in the depressive state investigation. *Chin J Clin Psychol.* (2014) 22:842–4. doi: 10.16128/j.cnki.1005-3611.2014.05.065

44. Koch C. *The Tree Test; the Tree-Drawing Test as an Aid in Psychodiagnosis*. Oxford: Grune and Stratton (1952).

45. Fukunishi I, Sugawara Y, Takayama T, Makuuchi M, Kawarasaki H, Surman OS. Association between pretransplant psychological assessments and posttransplant psychiatric disorders in living-related transplantation. *Psychosomatics.* (2002) 43:49–54. doi: 10.1176/appi.psy.43.1.49

46. Wang HL, Liu LL, Gao M, Ma HX. Effectiveness of drawing art tests in university students mental health test. *J HeBei United Univ.* (2019) 21:236–41. doi: 10.19539/j.cnki.2095-2694.2019.03.015

47. Laak JT, De Goede M, Aleva A, Rijswijk PV. The draw-a-person test: an indicator of children's cognitive and socioemotional adaptation? *J Genet Psychol.* (2005) 166:77–93. doi: 10.3200/GNTP.166.1.77-93

48. Machover K. *Personality Projection in the Drawing of the Human Figure: A Method of Personality Investigation*. Springfield, IL: Charles C Thomas Publisher (1949). p. 3–32. doi: 10.1037/11147-001

49. Lee EJ. Factors associated with nicotine addiction and coping skills in the synthetic house-tree-person drawing test. *J Korean Acad Psychiatr Mental Health Nurs.* (2020) 29:185–93. doi: 10.12934/jkpmhn.2020.29.2.185

50. Xiang JJ, Liao MS, Zhu MJ. Assessment of attention deficit hyperactivity tendencies in primary school students by the House-Tree-Person Drawing Test. *Children' Study.* (2020) 7:33–7.

51. Koide R, Fujihara K. A study on HTP organic signs. *Shinrigaku Kenkyu.* (1992) 63:277–80. doi: 10.4992/jjpsy.63.277

52. Zhu MJ, Chen T, Pei HC, Wang PC, Xing YL, Luo J, et al. Assessment of teenagers' narcissistic personality disorder inclination-based on the projective drawing test. *China J Health Psychol.* (2020) 28:676–80. doi: 10.13342/j.cnki.cjhp.2020.05.010

53. Meehan MC. Psychological evaluation of children's human figure drawings. *JAMA.* (1968) 205:190. doi: 10.1001/jama.1968.03140290082037

54. Zhu HL, Xiang JJ, Chen WJ, Shen HY, Gao L. Characteristics of HTP paintings of post-traumatic stress disorder adolescents in Sichuan earthquake area. *J Educ Dev.* (2011) 6:39–42. doi: 10.16215/j.cnki.cn44-1371/g4.2011.06.002

55. Li XM. *A Study on the Effect of H-T-P Test on the Evaluation of Junior High School Students' Anxiety*. Wuhan: Jianghan University (2020).

56. Lee EJ. Correlations among depressive symptoms, personality, and Synthetic House-Tree-Person Drawings in South Korean adults. *Psychologia.* (2019) 61:211–20. doi: 10.2117/psysoc.2019-A104.61-4

57. Zhao Y, Wang QY, Xiang JJ, Wang Q. Drawing characteristics of somatization tendency children in house-tree-person test. *Chin Mental Health J.* (2015) 29:115–20. doi: 10.3969/j.issn.1000-6729.2015.02.007

58. Zhou AB, Xie P, Pan CC, Tian Z, Xie JW. Performance of patients with different schizophrenia subtypes on the Synthetic House–Tree–Person Test. *Soc Behav Pers.* (2019) 47:1–8. doi: 10.2224/sbp.8408

59. Deng Y, Zhou CP, Wang YF. An exploration of the correlation between drawing characteristics and depressive tendencies. *Ability Wisdom.* (2017) 31:197. doi: 10.3969/j.issn.1673-0208.2017.31.168

60. Wang QS, Xiang JJ, Liu JX. Childhood trauma and self-concept of those with history of suicide attempt. *Chin Mental Health J.* (2007) 6:407–10. doi: 10.3321/j.issn:1000-6729.2007.06.016

61. Eisel HE. *A Comparative Study of the House-Tree-Person Drawings of Schizoid Personalities and Individuals with Below-Average Intelligence in a Prison Setting*. Ph. D. thesis. Ohio: The Ohio State University (1978).

62. Kirchner JH, Marzolf SS. Personality of alcoholics as measured by sixteen personality factor questionnaire and house-tree-person color-choice characteristics. *Psychol Rep.* (1974) 35(1 Pt 2):627–42. doi: 10.2466/pr0.1974.35.1.627

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 11 January 2023
DOI 10.3389/fpsyt.2022.1028342

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University, Germany

REVIEWED BY
Malahat Akbarfahimi,
Iran University of Medical Sciences, Iran
Marjorie Nicholas,
MGH Institute of Health Professions, United States

*CORRESPONDENCE
Filiberto Toledano-Toledano
✉ filiberto.toledano.phd@gmail.com

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 25 August 2022
ACCEPTED 19 December 2022
PUBLISHED 11 January 2023

CITATION
Merino-Soto C, Núñez Benítez MÁ, Domínguez-Guedea MT, Toledano-Toledano F, Moral de la Rubia J, Astudillo-García CI, Rivera-Rivera L, Leyva-López A, Angulo-Ramos M, Flores Laguna OA, Hernández-Salinas G, Rodríguez Castro JH, González Peña OI and Garduño Espinosa J (2023) Medical outcomes study social support survey (MOS-SSS) in patients with chronic disease: A psychometric assessment. *Front. Psychiatry* 13:1028342. doi: 10.3389/fpsyt.2022.1028342

Medical outcomes study social support survey (MOS-SSS) in patients with chronic disease: A psychometric assessment

Cesar Merino-Soto[1], Miguel Ángel Núñez Benítez[2], Miriam Teresa Domínguez-Guedea[3], Filiberto Toledano-Toledano[4,5,6]*, José Moral de la Rubia[7], Claudia I. Astudillo-García[8], Leonor Rivera-Rivera[9], Ahidée Leyva-López[9], Marisol Angulo-Ramos[6], Omar Arodi Flores Laguna[10], Gregorio Hernández-Salinas[11], Jorge Homero Rodríguez Castro[12], Omar Israel González Peña[4] and Juan Garduño Espinosa[13]

[1]Instituto de Investigación de Psicología, Universidad de San Martín de Porres, Surquillo, Peru, [2]Instituto Mexicano del Seguro Social, IMSS, Unidad de Medicina Familiar 31, Mexico City, Mexico, [3]Departamento de Psicología y Ciencias de la Comunicación, Universidad de Sonora, Hermosillo, Sonora, Mexico, [4]Hospital Infantil de México Federico Gómez, Unidad de Investigación en Medicina Basada en Evidencias, Mexico City, Mexico, [5]Instituto Nacional de Rehabilitación Luis Guillermo Ibarra Ibarra, Unidad de Investigación Sociomédica, Mexico City, Mexico, [6]Instituto Nacional de Ciencias e Innovación para la Formación de Comunidad Científica, INDEHUS, Dirección de Investigación y Diseminación del Conocimiento, Mexico City, Mexico, [7]Facultad de Psicología, Universidad Autónoma de Nuevo León, Monterrey, Mexico, [8]Servicios de Atención Psiquiátrica (SAP). Secretaría de Salud, Mexico City, Mexico, [9]Centro de Investigación en Salud Poblacional, Instituto Nacional de Salud Pública (INSP), Cuernavaca, Morelos, Mexico, [10]Facultad de Ciencias Empresariales y Jurídicas, Universidad de Montemorelos, Montemorelos, Mexico, [11]Tecnológico Nacional de México/Instituto Tecnológico Superior de Zongolica-Extensión Tezonapa, Heroica Veracruz, Mexico, [12]División de Estudios de Posgrado e Investigación, Tecnológico Nacional de Mexico/Instituto Tecnologico de Ciudad Victoria, Ciudad Victoria, Tamaulipas, Mexico, [13]Dirección de Investigación, Hospital Infantil de México Federico Gómez, Instituto Nacional de Salud, Mexico City, Mexico

Purpose: Currently, information on the psychometric properties of the Medical outcomes study-social support survey (MOS-SSS) for patients with chronic disease in primary health care, suggests problems in the dimensionality, specifically predominant unidimensionality in a multidimensional measure. The aim of this study was to determine the internal structure (dimensionality, measurement invariance and reliability) and association with other variables.

Methods: A total of 470 patients with chronic disease from a Family Medicine Unit at the Instituto Mexicano del Seguro Social, IMSS, with a mean age of 51.51 years were included. Participants responded to the Questionnaire of Sociodemographic Variables (Q-SV), SF-36 Health-Related Quality of Life Scale–version 1.1, and MOS-SSS.

Results: Non-parametric (Mokken scaling analysis) and parametric (confirmatory factor analysis) analyses indicated unidimensionality, and three-factor model was not representative. A new 8-item version (MOS-S) was developed, where measurement invariance, equivalence with the long version, reliability, and relationship with the SF-36 were satisfactory.

Conclusion: The MOS-SSS scale is unidimensional, and the shortened version yields valid and reliable scores for measuring social support in patients with chronic disease at the primary health care.

KEYWORDS

chronic disease, MOS-SSS, primary health care, psychometrics, Mexico, psychometric assessment

1. Introduction

Globally, an increase in the prevalence of non-communicable diseases (NCDs) (1, 2) has been reported, with the most common being diabetes mellitus (DM), systemic arterial hypertension (HAS), osteoarticular diseases and heart disease. In Mexico, the National Health and Nutrition Survey 2018–2019 (3) reported a prevalence of DM of 10.3% and HAS of 18.4%. NCDs are among the leading causes of death worldwide (1, 2) and in Mexico (4, 5); in addition, they are among the main reasons for consultation at the primary health care (6).

At the international level, the Global Health Metrics (7) reported an increase in the burden of disease associated with NCDs due to the years lost due to premature death (YLLD). Such an increase implies a process of gradual and continuous loss of health, affecting performance, independence, functionality and quality of life (8). It has been shown that the number of comorbid medical conditions is closely related to health-related quality of life (HRQoL) (9, 10) and to limitations in people's ability to perform activities of daily living (ADLs) (11, 12).

All these health-disease processes involve processes of adversity, risk and vulnerability for patients, their families and the health care system (13, 14). However, despite the warning of national and international agencies about the impact of chronic diseases (CDs) in young adults and older adults, no research processes have been developed to characterize the impact of chronic disease on individual, family and sociocultural indicators associated with social support processes (8).

Abbreviations: AFF, affective support; DM, diabetes mellitus; EMI, emotional/informational support; HAS, systemic arterial hypertension; HRQoL, health-related quality of life; MOS, medical outcomes study; MOS-S, short version of MOS-SSS; MOS-SSS, medical outcomes study-social support survey; NCD, non-communicable diseases; PSI, positive social interaction; Q-SV, Questionnaire of sociodemographic variables; TAN, tangible support; YLLD, years lost due to premature death.

The first studies on social support focused on psychosocial processes and stress (15) and social support as a buffer for stressful life processes (16). It is a construct that encompasses three components: support schema, support relationships, and support transactions (17). It has been defined as the social resources that people perceive as available or actually provided to them by non-professionals in the context of formal support groups and informal helping relationships that serve as an aid in coping with adverse life events and conditions (18–21). Social support is a determinant of health, and it fulfills different emotional, instrumental, informational, and companionship functions (22).

Thus, individuals' connections with their social environment occur at the community, social network, and intimate relationship levels (23). The social support has been classified into the following categories (24): (1) material help; (2) behavioral assistance; (3) intimate interaction; (4) guidance; (5) feedback; and (6) positive social interaction (25).

Empirical evidence indicates that social relationships can moderate the effects of stress on people's health and well-being, which impacts their family, social and work environments (16, 26–30), and in fact, associations between social support and mortality risk have been demonstrated (31–34). Moreover, the sources of social support differ cross-culturally (35, 36). Thus, different mechanisms have been identified through which social networks can influence chronic disease management: sharing knowledge, facilitating access to resources, engaging and maintaining productive relationships with network members (37). For example, a follow-up study conducted in women with school-aged children in the context of the COVID-19 pandemic showed how stress was associated with a higher probability of depression, while social support acted as a buffer against the effects of psychosocial stress and protected physical and mental health (38). Another study conducted in older adults suggested that having few social support networks could be a risk factor for reduced physical functioning, which was linked

to dependence in at least one of the ADLs and instrumental ADLs (39). However, other studies have reported that supportive behaviors do not have a positive effect on well-being (40) or may even be detrimental to the recipient (41) or the provider (42).

Given the considerable health implications of social support in patients with chronic disease, the need for a psychometrically sound instrument to measure social support in this population at the primary health care level is indicated. Such findings would provide validation of the Medical outcomes study (MOS) scale of social support in patients with chronic diseases at the first level of health care, obtain useful information to generate empirically based interventions aimed at developing and promoting social support resources, and may provide a novel and complementary approach to improve social support outcomes in this population (8). Therefore, the aim of this study was to evaluate the psychometric properties of the MOS in patients with chronic illness, determine the factor structure of the MOS, estimate its internal consistency reliability, and describe the distribution of MOS scores and the level of social support in the sample. It was hypothesized that the MOS social support survey (MOS-SSS) would show adequate psychometric properties.

The main theoretical formulations and recent empirical research results have focused on social integration, perceived social support, received social support and enacted social support (43–45). Regarding the assessment of social support, various generic and specialized measurement instruments have been developed in the international literature for adults and children and have been classified into measures of social integration, perceived social support, received social support and enacted social support (46, 47). These measures include the Family Relationship Index (FRI) (48), Inventory of Social Support Behaviors (ISSB) (49), Social Provisions Scale (SPS) (50), Social Support Network Inventory (SSNI) (51), among others (52–63).

In Mexico, the evaluation and measurement of social support has been carried out through the following measurement instruments: the Adult Social Support Scale (EAS) (64), the Social Support Scale in Family Caregivers of Older Adults (65), the Social Support Scale in Mexican Adults (66), the Social Support Network Scale (SSNS) (67) that has been validated in family caregivers of children with cancer (8), and the Perceived Social Support Scale (MSPSS) that has been validated in informal primary caregivers of cancer patients and presented satisfactory psychometric properties (68).

Although there are several available measurement instruments, Sherbourne and Stewart (69) developed and evaluated a multidimensional, self-administered 19-item Likert scale of social support (Medical outcomes study-social support survey; MOS-SSS) for patients with chronic illness that assesses five dimensions: emotional support (the expression of positive affect, empathetic understanding and the encouragement of expressions of feelings), informational support (the provision of advice, information, guidance or feedback), tangible support (the provision of material or behavioral assistance), positive social interaction (the availability of other people to do fun things with you), and affective support (including expressions of love and affection). The content of the MOS-SSS was constructed to focus on the sources of social support involved in patient well-being (69), and therefore its content validity is supported by the selection process of the literature and conceptually relevant items. The internal consistency for the 5 dimensions was >0.91, and the overall internal consistency was 0.97. This scale has been validated and adapted to multiple countries and languages, specifically for Chinese (70, 71), Taiwanese (72), Australian (73), Canadian and French (74), and Portuguese (75) populations. The Spanish version (76), used in the present study, was validated in primary health care for patients and consists of 20 items. The first question collects information on the size of the social network. The subsequent 19 items collect values referring to four dimensions of functional social support: emotional/informational support (items 3, 4, 8, 9, 13, 16, 17 and 19), tangible support (items 2, 5, 12 and 15), positive social interaction (items 7, 11, 14 and 18) and affective support (items 6, 10 and 20). In the study by Ahumada et al. (76), the factor analysis revealed the existence of 3 factors, which explained 68.72% of the overall variance. Cronbach's alpha coefficients for the three factors were >0.85. Factor 1 (items 3, 4, 8, 9, 11, 13, 14, 16, 17 and 19) corresponds to emotional/informational support; factor 2 (items 6, 7, 10, 18 and 20) corresponds to affective support; and factor 3 (items 2, 5, 12 and 15) measures instrumental support.

The Spanish version of the MOS-SSS has been adapted and validated in Mexico in two studies. First, in HIV + patients (77), exploratory factor analysis revealed two factors, namely, emotional/informative support and tangible support; these two factors explained 72.22% of the variance, with Cronbach's alpha values of 0.97 and 0.89, respectively. Three changes were made to the scale: (1) item 2, "someone to help you when you have to be in bed," was changed to "someone to help you when you have to be sick in bed"; (2) in item 9, "someone to confide in or talk to about yourself and your concerns," "yourself" was replaced; and (3) item 1, "approximately how many close friends or close family members do you have?" was changed to item 20 (77). Second, in Mexican patients with cardiovascular disease (78), the results showed four factors: emotional/informational support, positive social interaction, instrumental support and affective support. The internal consistency with Cronbach's alpha was 0.97, and Cronbach's alpha for the four factors were >0.95; the four factors explained 87.48% of the variance (78).

To explore the influential methodological points in the previous MOS-SSS validation studies, a systematic scoping review was conducted that focused on the properties of the internal structure. The search was made in a generic engine (Google) and specialized engines (PubMed, Google Scholar) with the keywords in Spanish and English: "validity," "medical

outcomes study (MOS)," and "social support." The inclusion criteria were articles with validity results in any language and year; studies whose complete content could not be retrieved and manuscripts that were not peer reviewed were excluded. Eligible manuscripts were reviewed by one of the coauthors (MAR), with 100% agreement reached. Some of these articles served as sources to search for additional validation articles [i.e., (79–81)]. The results are presented in **Table 1**, which shows the essential characteristics of the methodology applied and its influence on internal structure decision making.

The results of the scoping review indicated that the most tested model was the correlated factors model, and although this model accommodates the generalized tendency of the most used model in psychometric research (104), there are other reasonable models that can be solved in the assessment of the dimensionality of the MOS-SSS, given the evidence of factors with high or very high correlations with each other (**Table 1**, under the Fact R heading). Along similar lines, model comparisons were almost absent with the exception of a few studies [e.g., (79, 85, 94)], given that they directly tested the correlated factors model, and the confirmatory methodology was not exploited to verify other reasonably competitive models with support in antecedent research.

On the other hand, in this review, it was also found that inter-factor or inter-observed score correlations were rarely reported, even though these psychometric parameters are important for assessing the discriminative validity of the dimensions, and it is usual for confirmatory factor analysis (CFA) to report it (unless explicitly not estimated). In the studies where interfactor correlations were reported [including the study by Sherbourne and Stewart (69)], these tended to show high values, to a degree that raises suspicions about the conceptual discrimination of the dimensions; furthermore, with the exception of a few studies [e.g., (87)], the discriminative validity of the dimensions in the remaining studies is a matter of reasonable doubt. There were 4 abbreviated versions that were essentially motivated by the unidimensional representativeness of the items and the similarity of the factor loadings. On the other hand, the predominant analysis strategies did not consider the items as categorical variables and therefore used estimators for normally distributed continuous variables [i.e., maximum likelihood (ML)]. This may lead to a degree of non-ignorable underestimation of loadings and interfactor correlations, as is usual with CFA and ML estimators (105, 106), even with robust modifications for ML (107). This potential problem was also noted by Higgins et al. (96). It is apparent that, without adjustments or the use of polychoric correlations, factor loadings and correlations may have non-ignorable biases (106). Finally, total scores were obtained in several studies, even though the multidimensional model was advocated and established, which suggested that the MOS construct is represented with several obtainable scores, but not a single global score. A contrast with the rest of the studies was established with Margolis et al. (79), one of the few studies that, as an argument for their study, explicitly acknowledged the highly inconsistent internal structure of the MOS found in preceding studies. This study represented a methodological advance in the evaluation of the structure of the MOS-SSS because it used a recommended methodology for categorical variables [similar to Higgins et al. (96)] and included the comparison of models, including the bifactor model. Their bifactor model did not converge properly (negative variance), and it was concluded that the MOS-SSS model can be represented by a single dimension with numerous correlated errors. These latter findings on the dimensionality of the MOS-SSS, specifically the probable predominant unidimensionality, require careful examination for proper interpretation of its scores.

Given the background set of MOS-SSS validation studies in different cultural groups, the trends in the reporting of the results, and the results obtained, the present study aligns with what was expressed by Stewart and Napoli-Springer (108) and emphasized by Margolis et al. (79), which is the need to reevaluate a measure when the inconsistency in its dimensionality is a verifiable feature in the preceding literature. This need is critical to ensure the interpretation of the MOS-SSS measure and to define usable observed scores for theory and practice. In this sense, the objective was to obtain evidence of the internal structure of the MOS-SSS, incorporating a sequence of methodological decisions to define the number of dimensions, the internal validity of its items, and the parsimony of its interpretation by means of a proposed abbreviated version. This objective was also accompanied by other analyses that provided the remainder of validity evidence: measurement invariance (not performed in almost all previous studies), comparison of measurement models, and equivalence between versions of the MOS-SSS (full version vs. new abbreviated version).

2. Materials and methods

The type of study was non-experimental and cross-sectional, and the participants were chosen using non-probabilistic, convenience-based sampling method.

2.1. Participants

A total of 470 patients with chronic diseases participated (women: 297, 63.2%; men: 173, 36.8%) with an average age of 51.51 years (SD = 15.45). The participants were interviewed in a family medicine unit in Mexico City. The inclusion criteria were (a) affiliated with and receiving regular treatment in the family medicine service for the control of chronic diseases (DM, HAS, chronic renal disease, chronic obstructive pulmonary disease, obstructive sleep

TABLE 1 Review of MOS-SSS validation studies for evidence of internal structure.

References	Country	Items	Factors	Dimensionality	Fact R	*P* total	Equiv/inva	Equiv long/short
Yu et al. (71)	China 110	19	EMI TAN AFF PSI	CFA: ML	Min = 0.88 Max = 0.99	Yes	N.R.	N.Rel.
Westaway et al. (82)	South Afrika 263	19	EMI TAN	PCA: varimax	N.R.	Yes	N.R.	N.Rel.
Shyu et al. (72)	Taiwan 265	19	EMI TAN	EFA: varimax	R = 0.71	No	N.R.	N.Rel.
Alonso et al. (83)	Portugal 101	19	EMI TAN AFF PSI	PCA: varimax CFA: ML	Min = 0.97 Max = 0.99	Yes	N.R.	N.Rel.
Requena et al. (84)	Spain 400	19	EMI TAN AFF	PCA: varimax	N.R.	No	N.R.	N.Rel.
Gjesfjeld et al. (85)	USA 330	12 4	EMI TAN AFF PSI	CFA: ML	N.R.	Yes	N.R.	Yes $R > 0.90$
Espínola and Enrique (86)	Argentina 375	19	EMI TAN AFF	PCA: varimax	N.R.	Yes	N.R.	N.Rel.
Pais-Ribeiro and Ponte (87)	Portugal 225	19	EMI TAN AFF PSI	PCA: varimax	Min = 0.15 Max = 0.60	Yes	N.R.	N.Rel.
Zanini et al. (81)	Brazil 129	19	EMI TAN AFF PSI	EFA: varimax	N.R.	No	N.R.	N.Rel.
Robitaille et al. (88)	Canada 3,131	19	EMI TAN AFF PSI	CFA: N.R.	N.R.	No	Métrica	N.Rel.
Ashing-Giwa and Rosales (89)	320 Multinational	19	EMI TAN AFF PSI	N.R.	N.R.	No	N.R.	N.Rel.
Londoño et al. (90)	Colombia 179	19	EMI TAN AFF PSI	EFA: varimax/oblicua CFA: N.R.	N.R.	No	N.R.	N.Rel.
Moser et al. (91)	USA 3,241	8	EMI TAN	PCA: varimax CFA N.R.	N.R.	No	N.R.	N.R.
Soares et al. (92)	Brazil	6	One dimension	PCA: varimax	N.Rel.	Yes	N.R.	N.Rel.
Wang et al. (70)	China 200	19	EMI TAN AFF PSI		Min = 0.68 Max = 0.89	Yes	N.R.	N.Rel.
Gomez-Campelo et al. (93)	Spain 1,594	8	One dimension	CFA: ULS	N.Rel.	Yes	N.R.	N.R.
Holden et al. (73)	Australia 20,493	6	One dimension	CFA: ADF	N.Rel.	Yes	N.R.	Yes $R > 0.90$

(Continued)

TABLE 1 (Continued)

References	Country	Items	Factors	Dimensionality	Fact R	*P* total	Equiv/inva	Equiv long/short
Basurto et al. (77)	Mexico	19	Emoc (14) Tang (5)	PCA: varimax CFA: MLR	N.R.	No	N.R.	N.Rel.
Giangrasso and Casale (94)	Italia 485	19	EMI TAN AFF PSI	CFA: N.R.	Min = 0.46 Max = 0.75	Yes	N.R.	N.Rel.
Conte et al. (95)	USA 505	19	EMI TAN AFF PSI	PCA: N.R. CFA: N.R.	N.R.	No	N.R.	N.Rel.
Higgins et al. (96)	USA 406	8 4	One dimension	CFA: WLSMV	N.Rel.	No	N.R.	N.R.
Norhayati et al. (97)	Malasya 144	16	EMI Tan Pos	CFA: N.R.	Min = 0.39 Max = 0.86	No	N.R.	N.Rel.
Yu et al. (98)	China 200	19	Emoc (14) Tang (5)	EFA: oblicua		Yes	N.R.	N.R.
Togari and Yokoyama (99)	Japan 2,052	8	Instrum (4) Emoc (4)	PCA: promax	N.R.	Yes	N.R.	N.Rel.
Zanini and Peixoto (80)	Brazil 998	19	EMI TAN AFF PSI	CFA: ML	Min = 0.41 Max = 0.73	No	N.R.	N.Rel.
Priede et al. (100)	Spain 128	19	E-I-SI Instru Afec	PCA: varimax	N.R.	No	N.R.	N.Rel.
Margolis et al. (79)	USA 199	19	One dimension	CFA: WLSMV	Min = 0.88 Max = 0.96	No	N.R.	N.Rel.
Yilmaz and Bozo (101)	Turkey	19	EMI TAN AFF PSI	EFA: varimax	N.R.	No	N.R.	N.Rel.
Martin-Carbonell et al. (102)	Colombia 463	19	EMI TAN AFF PSI	CFA: ULS	Min = 0.77 Max = 0.95	Yes	No	N.Rel.
Navarrete et al. (78)	Mexico 229	19	EMI TAN AFF PSI	PCA: N.R. CFA: ML	Min = 0.59 Max = 0.75	No	N.R.	N.Rel.
Bavarsad et al. (103)	Iran 420	5	Inst (2) Emoc (3)	PCA: varimax CFA: ML	0.55	Yes	No	No

EFA, exploratory factor analysis; PCA, principal components analysis; CFA, confirmatory factor analysis; varimax and oblique, types of rotations; ML WLSMV, ULS, ADF, MLR, estimators; Fact R, interfactor correlations; P total, total score computed; Equiv/inva, measurement equivalence/invariance; Equiv long/short, equivalence between long and short forms; EMI, emotional/informational support; TAN, tangible support; AFF, affective support; PSI, positive social interaction; N.R., not reported; N.Rel., not relevant.

apnea syndrome, degenerative osteoarthrosis, cerebral vascular disease, rheumatoid arthritis, cancer, hypothyroidism and epilepsy), (b) at least 20 years of age, (c) male or female, and (d) signed an informed consent form. The exclusion criteria were (a) inability to read and write and (b) refusal to participate in the study. The elimination criteria included (a) partial or incomplete responses to the measurement instruments and (b) having been detected as a potential generator of biased responses. In this patient sample, chronic mental health diseases, as well as, some chronic autoimmune diseases such as Multiple Sclerosis (MS), Systemic Lupus Erythematosus (SLE), Myasthenia Gravis, Inflammatory Bowel Disease (IBD), among others, have a low prevalence; patients with these diagnoses are generally seen at a third level of care; similarly, patients with HIV/AIDS

are seen at a second level of care. Consequently, care of these patients at the *primary health care* is infrequent. No patients with these diagnoses were found in the family medicine office during the study period, so they were not included. Finally, patients with Cerebrovascular Disease (CVD) and/or Cerebrovascular Accident (CVA) were included in the study, six patients participated (1.3%).

2.2. Ethical considerations

This study is part of the research project HIM/2015/017/SSA.1207 "Efectos del entrenamiento en mindfulness sobre el distrés psicológico y la calidad de vida del cuidador familiar," which was approved by the Research, Ethics, and Biosafety Committees of the Hospital Infantil de México Federico Gómez, Instituto Nacional de Salud, in Mexico City. To conduct this study, we followed the rules and ethical considerations for human research currently applicable in Mexico (109, 110) and those described in Sociedad Mexicana de Psicología American Psychological Association (111). All patients were informed about the objectives and scope of the research and their rights in accordance with the Declaration of Helsinki (112). Patients who agreed to participate in the study signed a letter of informed consent. Participation in this study was voluntary and did not involve payment.

2.3. Procedure

Once the research protocol was approved, the battery of measurement instruments was integrated. The patients were identified by the research team in the waiting rooms and in the consultation room of the family medicine unit. Then, the team members asked the patients for their voluntary participation in the study, and they were presented with the informed consent letter, which they signed. Likewise, they were guaranteed their right to withdraw from the study at any time they wished without an impact on or risk to their care in the institution. The participants were informed about the objective of the research, the instruments they would complete and the time they should have available for this activity. At all times, the interviewer verified that there were no unanswered questions to prevent having missing values. At the end of the interview, the patients were verbally thanked and were given the opportunity to express any doubts or concerns about their participation.

2.4. Measures

2.4.1. Medical outcomes study-social support survey (MOS-SSS)

This self-report questionnaire consisted of 20 items rated on a five-point Likert-type scale that ranged from 1 "never" to 5 "always"; the first item reported on the size of the social network, and the subsequent 19 items measured four dimensions of functional social support: emotional/informational support (the expression of positive affect and the provision of advice, information, guidance or feedback) (eight items: 3, 4, 8, 9, 13, 16, 17, and 19), instrumental support (the provision of material or behavioral assistance) (four items: 2, 5, 12, and 15), positive social interaction (the availability of other people to do fun things with you) (four items: 7, 11, 14, and 18) and affective support (including expressions of love and affection) [three items: 6, 10, and 20 (69)]. The present study used the Spanish version from Ahumada et al. (76).

2.4.2. Questionnaire of sociodemographic variables for research on family caregivers of children with chronic diseases (Q-SV)

This questionnaire contained 20 items that collect information on sociodemographic, medical, sociocultural and family variables from families of children with chronic diseases. The content of this questionnaire maximized the amount of demographic information, with content relevant to these families (113).

2.4.3. SF-36 scale of health-related quality of life

This is a Likert-type scale (36 items) that evaluated positive and negative states of physical and mental health; item 2 is a transition item that asks about the change in the general state of health with respect to the previous year and was not used for the calculation of any of the 8 dimensions of health status: physical function (ten items), physical role (four items), bodily pain (two items), general health (five items), vitality (four items), social function (two items), emotional role (three items) and mental health (five items). The reported Cronbach's alpha reliability coefficient was reported to range from 0.56 to 0.84 for the different dimensions (114).

2.5. Data analysis

First, data cleaning focused on the detection of excessively inconsistent and consistent responses, i.e., possible response biases. These response patterns were examined by means of the multivariate distance D^2 (115) and the longest sequence of consecutive responses [longstring; Johnson (116)]. For D^2, the cutoff point for detection was $D^2 > 36.19$ (at $p = 0.01$); for the longstring method, the cutoff point for detection was half the number of items in the total instrument (117, 118), i.e., 19/2 = 9.5 (set to 10). To reduce false negatives, Tukey's fences were also used, with parameter k set to 1. The database consisted of retaining participants not detected by the two independent methods. Both methods are recommended for the identification of suspected cases of insufficient effort when answering long

questionnaires (119). The analysis was performed with the R program *careless* (120).

With the database without the participants showing potentially biased responses, descriptive and association statistics were obtained for items treated as ordinal categorical variables (121), specifically to identify associations with sex (Glass rank biserial correlation coefficient), chronological age (Spearman's correlation coefficient), marital status (ordinal eta-squared), and education (ordinal eta-squared). These associations may be potential indicators of the differential functioning of items in the compared groups and of the sensitivity of the content for score interpretation purposes (122). The analysis was performed with the R programs *rcompanion* (123) and *MVN* (124).

To test the internal structure of the instrument, we first applied a non-parametric approach, Mokken scaling analysis (MSA) (125), that is a method focused on the psychometric properties of the observed score by analyzing the number of dimensions, the scaling of items and scores, local independence, and the monotonic item-score relationship (125, 126), as these are characteristics that build the monotonic homogeneity model (MHM) (125). MSA does not require the assumptions of parametric analyses [e.g., structural equation modeling or item response theory; Crişan et al. (127)] and is a preliminary procedure for subsequent latent construct analysis (127, 128). Additionally, this method was considered appropriate given the moderate sample size in each randomly drawn subsample and the small number of items in some of the MOS-SSS subscales. Within the MSA, to determine the number of instrument scales, the automated item selection procedure (AISP) (125, 126) was used with the normal search based on the increasing scalability of items grouped by the scalability coefficient H (127). The analysis was performed with the R program *mokken* (129).

To obtain parametric estimates of the internal structure of the MOS and based on the results of the MSA, parallel analysis (PA) (130) was used to identify the number of latent dimensions, and confirmatory factor analysis of structural equation modeling (CFA-SEM) was used to contrast measurement models. Used on categorical variables, such as MOS-SSS items, PA is still an optimal method for estimating the number of latent dimensions (131). PA was used on the interitem polychoric correlations of the simulated data in PA using the psych program (132). The total sample was divided into two halves to assess the replicability of the number of dimensions.

With CFA-SEM, we evaluated (a) the 4-factor multidimensional model of Ahumada et al. (76), which was the source of the MOS-SSS version used in this study, and (b) the unidimensional model, whose result was obtained from the MSA. The weighted least square mean and variance adjusted (WLSMV) estimator was used on the interitem polychoric correlations, given that the items were treated as categorical variables (133).

TABLE 2 Sociodemographic and clinical variables characteristics of patients with CDs (n = 470).

	N	%	*M*	SD
Sex				
Female	297	63.2		
Male	173	36.8		
Age				
20–29	47	10		
30–39	71	15.1		
40–49	100	21.3		
50–59	88	18.7		
60–69	108	23		
70–79	45	9.6		
80–89	10	2.1		
90–99	1	0.2		
Total	–	–	51.51	15.45
Marital status				
Married	261	55.5		
Single	67	14.3		
Widowed	38	8.1		
Divorced	22	4.7		
Free-union	82	17.4		
Instruction				
Primary incomplete	35	7.4		
Primary complete	83	17.7		
Secondary incomplete	15	3.2		
Secondary complete	122	26		
High school incomplete	6	1.3		
High school complete	86	18.3		
Technical	49	10.4		
Bachelor	52	11.1		
Graduate	1	0.2		
No studies	21	4.5		
Monthly income (Mexican currency)				
0–2,699	56	11.9		
2,700–6,799	250	53.2		
6,800–11,599	120	25.5		
11,600–34,999	44	9.4		
Disease				
HAS	296	63		
DM	265	56.4		
CKD	14	3		

(Continued)

TABLE 2 (Continued)

	N	%	M	SD
COPD	22	4.7		
OSA	10	2.1		
AMI	14	3		
CVD	6	1.3		
Osteomuscular diseases	90	19.1		
RA	5	1.1		
Cancer	19	4		
Hypothyroidism	20	4.3		
Epilepsy	6	1.3		

CDs, chronic diseases; HAS, systemic arterial hypertension; DM, diabetes mellitus; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; OSA, obstructive sleep apnea syndrome; AMI, acute myocardial infarction; CVD, cerebral vascular disease; RA, rheumatoid arthritis.

Fitting of these models was performed with approximate fit indices, such as CFI (>0.95), RMSEA (<0.05), and SRMR (<0.05). The 4-factor model was further evaluated on its discriminative validity among its factors with the heterotrait-monotrait ratio (HTMT) criterion (134, 135), a measure sensitive to the degree of statistical differentiation in SEMs (136–138). HTMT compares interitem correlations of different constructs (heteroattribute – heteromethod correlations) with interitem correlations of the same construct (monoattribute – heteromethod correlations). Two cutoff points were chosen, namely, HTMT > 0.90 (139) and HTMT > 0.85 (137, 138, 140), to identify factors with poor discriminative validity and moderate discriminative validity, respectively.

The dimensionality results of the MOS-SSS were assessed for replicability by randomly partitioning the sample into two halves, n = 159 and n = 158 (this was done on the total clean sample, n = 317).

Based on the results of the internal structure, an abbreviated version was created (MOS-S) from the internal strength of the scale, which retained more construct variance (141, 142). Thus, 2 items with the highest factor loadings were chosen in the unidimensional factor solution, but each item also corresponded with each theoretical dimension. For items with equal loadings, the content was considered to maximize content heterogeneity and was chosen for the short version. Equivalence between the short and long versions was assessed by linear correlation (142) but with correction for overlap (143), which is especially used to remove correlated error variance when both versions come from the same administered group (141). Equivalence was further assessed with the classificatory agreement generated by both scores at levels of 3 (tertiles), 4 (quartiles) and 5 groups (quintiles); the coefficient of agreement AC (144) was used.

Measurement invariance of the MOS-S was evaluated with respect to the sex group of the patients. Taking into account the sample size of the study [>300; Chen (145)], the suggested invariance criteria for CFI, SRMR, and RMSEA were <0.010, <0.030, and <0.015, respectively (145). Participant sex was chosen as a possible source of measurement invariance because this phenotypic characteristic is usually included in studies of invariance of psychosocial measures (122).

The reliability of the scores of the final version of the MOS-SSS was estimated with the omega coefficients for categorical variables (146) and with the Molenaar-Sijtsma coefficient (MS-rho) (147), both from SEM and MSA modeling, respectively. The alpha coefficients were also estimated. These estimates were made with the R programs *mokken* (129) and *MBESS* (148).

Finally, as evidence of the relationship between the construct measured by both versions of the MOS and the dimensions of the SF-36, correlation analyses were performed using Pearson's linear association coefficient. The difference in the correlations obtained between each of the versions of the MOS and the SF-36 was evaluated with Hittner et al.'s (149) z test and the confidence interval for the difference between dependent correlations (150). The procedure was performed with the R program cocor (151).

3. Results

3.1. Sociodemographic and clinical characteristics of patients with chronic disease

The results indicated that most of participants were female (63.2%) and the average age was 51.51 years (SD = 15.45); the three most prevalent diseases were HAS (63%), DM (56.4%) and musculoskeletal diseases (19.1%); and the total number of diseases experienced by a patient ranged from a minimum of 1 to a maximum of 5 diseases (M = 1.63, SE = 0.76). In most cases, high school was the highest level of education attained by the participants (26%). A high percentage of the patients were married (55.5%), and the majority reported a monthly income of between $2,700 and $6,799 (53.2%). This information can be seen in **Table 2**.

3.2. Possible response biases

Seventy-nine cases were detected with the D^2 method, with an inconsistent response pattern in the set of 19 items (16.8%); with the longstring method, consecutive identical responses had a median of 5 consecutive identical responses (M = 7.6, Q1 = 4, Q3 = 10) and a range between 2 and 19 identical responses. Applying Tukey's fences criterion (longstring > 16), 74 (15.7%) cases were detected between 17 and 19 consecutive identical responses. The linear correlation between the two estimators

was high ($r = -0.58$, $p < 0.01$, 95% CI = −0.64, −0.52; for the classification of detected cases: Cramer $V = 1.0$, 95% CI = 0.98, 1.00, $\chi^2 = 7{,}520.0$, $p < 0.01$, gl = 4,672). Because of this high divergence, no subjects were detected with both methods; the detected cases were excluded (79 and 74, 32.5%), and the sample for analysis consisted of 317 participants (78.9%).

3.3. Descriptive, normality and association statistics for the MOS items

The mean responses on the items tended to be similar since all were >3.0 (Table 3). The highest mean response (4.00) was only 0.29 times higher than the lowest mean response (3.1). The above pattern of similarity was also observed in the dispersion of the items estimated by the standard deviation of each item, where the maximum (1.49) and minimum (1.24) mean dispersion was only 0.20 times. Regarding the distribution of the items in the range of response options, skewness and kurtosis were in the same direction (i.e., negative), suggesting a highly similar distributional behavior. In the same line of results, univariate normality did not hold for all items.

3.4. Associations between the MOS items and demographic variables

Regarding the association of the MOS items with demographic variables (Table 3), the association with sex (min = −0.01, max = 0.15, Md = 0.07), chronological age (min = −0.08, max = 0.07, Md = 0.01), marital status (min = 0.00, max = 0.00, Md = 0.00), and education (min = −0.01, max = 0.00, Md = 0.00) were all maintained around zero, and there was an absence of statistical significance.

3.5. Evidence of the internal structure of the MOS social support survey

3.5.1. Non-parametric modeling (Mokken scaling analysis)

Table 4 shows the results of the MSA modeling. The AISP algorithm for Mokken scale selection yielded a likely different MOS structure. At the 0.30, 0.40, and 0.50 levels of the scaling coefficient H, the number of scale differences remained constant, where a single dimension was the apparent best definition of the internal structure of the MOS. This result was replicated in the two samples randomly drawn from the total sample, indicating that the unidimensionality of the MOS was replicable.

3.5.2. Parametric modeling (−CFA-SEM)

Given the results of the non-parametric modeling, where the apparent unidimensionality can be accepted, the dimensionality was again examined using linear parametric modeling. Table 5 shows the results of the PA on the total sample and on the two randomly drawn samples. The calculated eigenvalues clearly differentiated between a model possibly represented by a single factor (eigenvalues > 11.00) compared to the dimensionality of two or more factors (eigenvalues < 1.00). The corresponding graphs in each analysis also show the representativeness of a single dominant factor and its replicability.

The final decision regarding dimensionality was evaluated in the total sample, with the comparative fit of two models, one representing the 4-factor multidimensional model [from Ahumada et al. (76)] and the unidimensional model (suggested in the previous sections of the present study). The fit of the 4-factor multidimensional model (MOS-4F) revealed WLSMV-$\chi^2 = 332.745$ (df = 146), CFI = 0.999, RMSEA = 0.064 (90% CI = 0.055, 0.073), and SRMR = 0.039. The unidimensional model (MOS-1F) also showed an acceptable fit with WLSMV-$\chi^2 = 509.44$ (df = 171), CFI = 0.998, RMSEA = 0.086 (90% CI = 0.078, 0.095), and SRMR = 0.048. The factor loadings obtained in both models were high (>0.60); although the RMSEA indicated that the degree of misfit was lower in the multidimensional model compared to the unidimensional one ($RMSEA_{MOSS-4F} < RMSEA_{MOSS-1F}$), it was observed that there was no substantial difference in the fit indices between the two models.

A comparative inspection in detail of the obtained parameters (i.e., factor loadings and interfactor correlations; Table 6) revealed that the size of the factor loadings was highly similar ($r = 96$, $p < 0.01$; congruence coefficient = 0.99). Additionally, the correlations between factors ranged between 0.99 and0.86, a size range that can be considered high (137). Assessment of the discriminative validity between factors (under the heading "Correlations/HTMT" in Table 6) yielded HTMT indices that essentially bordered on or exceeded both criteria for poor discrimination (HTMT ≥ 0.94). Given the results of both MSA and CFA-SEM analyses, the unidimensional model appears to represent social support well, without loss of internal validity in the present sample.

3.5.3. Short version (MOS-S)

The items with the highest factor loadings in their previous content dimensions were as follows (Table 6): 9, 16, 17, 12, 15, 7, 11, 14, 6 and 10. Based on the content analysis, the content of item 17 can be subsumed in item 9, where sharing and expressing concerns can be oriented to several purposes, among them, problem solving. Item 14 seemed more directly linked to the content of the rest of its theoretical dimension because of the reference to the condition of health or illness. The final short version consisted of eight items: 9, 16, 12, 15, 7, 11, 6 and 10. Table 6, under the heading "short version," shows the recalculated parameters for the items of this abbreviated version, with CFA-SEM and MSA. Strong factor loadings are observed

(>0.81) and are similar to their corresponding factor loadings in the long version (congruency coefficient = 0.99).

3.5.4. Measurement invariance

The configurational, metric and scalar invariances were satisfactory (Table 7). Additionally, the differences between these models indicated that the invariance in the psychometric parameters of the MOS-S was maintained up to the invariance in the residuals. Based on these results, the parameters obtained in the total sample are equally representative for both sex groups of patients.

3.6. Reliability

The reliability of the score from the new abbreviated version of the MOS was α = 0.95 (95% CI = 0.94, 0.96) and ω = 0.97 (95% CI = 0.97, 0.99); based on the MSA framework, the reliability was rho-MS = 0.96. The standard error of measurement (using SD in the total sample = 9.75, and the alpha coefficient) corresponding to this score was 2.18. The reliability of the long version of the MOS, with a single score, was α = 0.99 (95% CI = 0.99, 1.00) and ω = 0.99 (95% CI = 0.99, 1.00). The difference between the internal consistency of the shortened version and the unidimensional long version can be considered trivial.

3.7. Equivalence between versions (MOS-SSS and MOS-S)

The linear association between the scores of both unidimensional short and long versions was $r = 0.98$ ($t = 95.3$, df = 315, $p < 0.01$); with correction for overlapping, the correlation was 0.95. The degree of agreement (Gwet's AC1 coefficient) between the classification of scores into tertiles, quartiles, and quintiles produced by both scores (short and long-unidimensional version) was, respectively, AC1 = 0.90 ($p < 0.01$; 95% CI = 0.86, 0.94), AC1 = 0.86 ($p < 0.01$; 95% CI = 0.82, 0.90), and AC1 = 0.80 ($p < 0.01$; 95% CI = 0.75, 0.85).

3.8. Association with other variables

The linear association of both versions of the MOS (19-item and 8-item versions) is shown in Table 8. Except for physical role, the rest of the correlations were statistically non-significant and practically zero. Statistical comparison between

TABLE 3 Descriptive and association statistics for MOS-SSS/MOS-S items (n = 317).

	Descriptive					Association			
	M	SD	Sk	*K*	AD	Sex	Age	Marital	Instruct.
MOS3	3.61	1.29	−0.52	−0.88	15.28	0.06	−0.02	−0.00	−0.01
MOS4	3.48	1.32	−0.40	−1.00	13.12	0.12	−0.01	−0.00	−0.00
MOS8	3.57	1.34	−0.62	−0.83	16.33	0.03	0.06	−0.00	0.00
MOS9	3.62	1.34	−0.58	−0.95	17.50	0.13	0.00	−0.00	0.00
MOS13	3.41	1.42	−0.39	−1.20	15.09	0.10	−0.03	−0.00	0.00
MOS16	3.33	1.43	−0.23	−1.33	14.67	0.06	−0.00	0.00	−0.00
MOS17	3.34	1.41	−0.25	−1.29	14.03	0.11	−0.03	−0.00	0.00
MOS19	3.56	1.32	−0.46	−0.98	14.74	0.07	0.01	−0.00	−0.01
MOS2	3.10	1.49	−0.15	−1.42	14.37	0.08	0.04	−0.00	−0.01
MOS5	3.48	1.48	−0.46	−1.24	18.93	0.07	0.07	−0.00	−0.00
MOS12	3.58	1.46	−0.54	−1.12	2.85	−0.01	0.03	0.00	−0.01
MOS15	3.47	1.40	−0.44	−1.14	15.76	0.02	0.02	0.00	−0.01
MOS7	3.88	1.24	−0.87	−0.32	21.70	0.07	0.00	0.00	−0.00
MOS11	3.57	1.38	−0.46	−1.11	17.49	0.06	0.02	−0.00	−0.00
MOS14	3.53	1.32	−0.35	−1.20	15.99	0.08	0.03	−0.00	0.00
MOS18	3.60	1.27	−0.45	−0.97	14.90	0.04	−0.08	−0.00	−0.01
MOS6	4.00	1.25	−1.09	0.01	28.71	0.10	0.04	−0.00	−0.00
MOS10	3.77	1.42	−0.75	−0.84	25.94	0.15	0.01	−0.00	0.00
MOS20	3.81	1.32	−0.71	−0.79	23.26	0.10	0.01	−0.00	−0.00

Sk, skew coefficient; K, kurtosis coefficient; AD, Anderson–Darling normality test; Marital, marital status; Instruc., instruction level.

TABLE 4 MSA: Number of dimensions (AISP) and monotonic homogeneity model (MHM).

	Total sample (n = 317)					Random sample 1 (n = 159)					Random sample 2 (n = 158)				
	AISP			MHM		AISP			MHM		AISP			MHM	
	0.3	0.4	0.5	H	Crit	0.3	0.4	0.5	H	Crit	0.3	0.4	0.5	H	Crit
MOS3	1	1	1	0.72	0	1	1	1	0.68	0	1	1	1	0.75	0
MOS4	1	1	1	0.69	0	1	1	1	0.67	0	1	1	1	0.70	0
MOS8	1	1	1	0.76	0	1	1	1	0.73	0	1	1	1	0.79	0
MOS9	1	1	1	0.77	0	1	1	1	0.74	0	1	1	1	0.80	0
MOS13	1	1	1	0.74	0	1	1	1	0.72	0	1	1	1	0.77	0
MOS16	1	1	1	0.78	0	1	1	1	0.73	0	1	1	1	0.82	0
MOS17	1	1	1	0.76	0	1	1	1	0.71	0	1	1	1	0.80	0
MOS19	1	1	1	0.73	0	1	1	1	0.68	0	1	1	1	0.79	0
MOS2	1	1	0	0.46	38	1	0	0	0.39	0	1	1	1	0.52	9
MOS5	1	1	1	0.66	0	1	1	1	0.64	0	1	1	1	0.68	0
MOS12	1	1	1	0.69	0	1	1	1	0.66	0	1	1	1	0.71	0
MOS15	1	1	1	0.70	0	1	1	1	0.66	0	1	1	1	0.73	0
MOS7	1	1	1	0.75	0	1	1	1	0.76	0	1	1	1	0.75	0
MOS11	1	1	1	0.75	0	1	1	1	0.72	0	1	1	1	0.79	0
MOS14	1	1	1	0.75	0	1	1	1	0.73	0	1	1	1	0.77	0
MOS18	1	1	1	0.70	0	1	1	1	0.66	0	1	1	1	0.75	0
MOS6	1	1	1	0.76	0	1	1	1	0.77	0	1	1	1	0.76	0
MOS10	1	1	1	0.76	0	1	1	1	0.75	0	1	1	1	0.78	0
MOS20	1	1	1	0.62	0	1	1	1	0.68	0	1	1	1	0.64	0

MSA, Mokken scaling analysis; AISP, automated item selection procedure; MHM, monotonic homogeneity model.

10.3389/fpsyt.2022.1028342

the two versions of the MOS indicated an absence of substantial differences, and differences that were rather trivial in size. Although a statistically significant difference was found (in social function), the size of this difference can be considered trivial (see the range of the difference in these correlations, Δ_r).

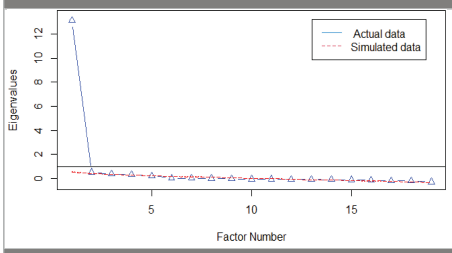
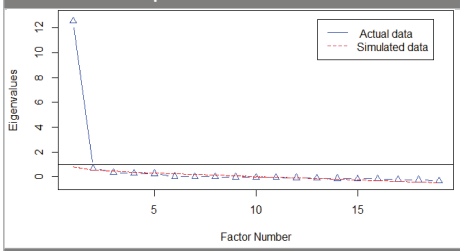
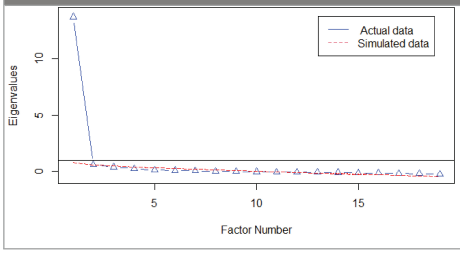
4. Discussion

The objectives of this research were to obtain evidence of the validity of the MOS scale with respect to its factorial structure, its internal consistency reliability, and its relationship with other variables. The complementary objectives were to describe the distribution of its scores. This study was implemented in patients with chronic disease at the primary health care, where the measurement of social support is relevant for knowing the resources that can impact the patient's quality of life.

According to the results, the unidimensional model adequately represents the construct of social support measured by the MOS. The validity of the items with respect to their latent constructs was not affected by the shift from multidimensional modeling to the unidimensional model. One implication of this is that social support represented by a single score does not alter the significance of the items in defining an overall construct. However, another implication is that the items do not represent content that previously appeared to be differentiated, i.e., the items do not represent specific dimensions such as the four dimensions obtained in Ahumada et al. (76). This unidimensional representation of the construct measured by the MOS leads to rethinking the theoretical definition of social support coming from the MOS framework, as well as testing a definition for the interpretation of the total score of the instrument. This definition is more parsimonious since it is focused on a general domain and not divided into separate dimensions.

The results are not congruent with the conclusions of the Hispanic studies (see **Table 1**), including those reported by Sherbourne and Stewart (69), because these studies reported the apparent multidimensionality of the MOS. As described in the Introduction, this discrepancy is fueled by the methodological characteristics of these studies that influenced decision-making about internal structure, as well as by the incomplete reporting of their factorial results. Specifically, few of these studies reported interfactor correlations [e.g., (78, 83)], and when reported, the size of the interfactor correlations showed a range between 0.59 and 0.75 (78) or 0.97 and 0.99 (83). These magnitudes are clearly high or very high and show that the discriminative validity of the MOS scales is not defensible

TABLE 5 Parallel analysis (number of factors).

Total sample (n = 317)	*N* factors	Eigenvalues	
		Real	Simulated
	1	13.12	0.58
	2	0.52	0.39
	3	0.35	0.33
	4	0.31	0.27
	5	0.19	0.22
Random sample 1 (n = 159)	***N* factors**	**Real**	**Simulated**
	1	12.54	0.81
	2	0.68	0.54
	3	0.34	0.46
	4	0.27	0.38
	5	0.24	0.30
Random sample 2 (n = 158)	***N* factors**	**Real**	**Simulated**
	1	13.70	0.77
	2	0.59	0.56
	3	0.36	0.47
	4	0.23	0.38
	5	0.11	0.30

10.3389/fpsyt.2022.1028342

TABLE 6 CFA-SEM of the MOS-SSS and MOS-S (n = 317).

	MOS–SSS				MOS-1F	MOS-S		
						CFA	MSA	
	EMI	TAN	PSI	AFF			H	Crit
MOS3	0.86				0.85	–	–	–
MOS4	0.84				0.83	–	–	–
MOS8	0.93				0.92	–	–	–
MOS9	0.94				0.93	0.83	0.80	0
MOS13	0.92				0.92	–	–	–
MOS16	0.95				0.94	0.92	0.81	0
MOS17	0.94				0.93	–	–	–
MOS19	0.91				0.90	–	–	–
MOS2		0.60			0.55	–	–	–
MOS5		0.87			0.79	–	–	–
MOS12		0.92			0.85	0.88	0.74	0
MOS15		0.92			0.85	0.87	0.75	0
MOS7			0.91		0.91	–	–	–
MOS11			0.92		0.92	0.93	0.81	0
MOS14			0.91		0.91	0.90	0.79	0
MOS18			0.86		0.86	–	–	–
MOS6				0.94	0.92	0.89	0.78	0
MOS10				0.95	0.92	0.94	0.81	0
MOS20				0.78	0.76	–	–	–
Correlations/HTMT								
	EMI	TAN	PSI	AFF				
EMI	1	0.87	0.98	0.91				
TAN	0.88	1	0.86	0.84				
PSI	0.99	0.88	1	0.97				
AFF	0.94	0.86	0.99	1				

MOS-SSS, 19 items MOS-SSS; EMI, emotional/informational support; TAN, tangible support; AFF, affective support; PSI positive social interaction; HTMT, heterotrait-monotrait correlation. MOS-1F, unidimensional model; MOS-S, final model for MOS-S (8 items); CFA, CFA-SEM; MSA, Mokken scaling analysis; H, *scalability coefficients*; Crit weighted criterion for the monotonic homogeneity model.

and that a unidimensional factorial solution may be the best representation of the construct. This problem in discriminative validity was also reported in the MOS creation study, in which the relationship between emotional and informational support was 0.99 (69), and the unhypothesized item-scale correlations studied were approximately 0.50. On the other hand, in other studies, the degree of discriminative validity could be assessed because the analytic strategy forced us to estimate the interfactor correlation (varimax rotation), or it was not reported. Along with this type of orthogonal rotation, in which the factors are assumed to be completely independent, the analysis of principal components and the number of dimensions using the Kaiser criterion (eigenvalue > 1) were also frequent. This package of methodological choices is known as the little jiffy (152).

Another reason for divergence was that several studies reported 2 and 3 factors (77, 86, 88, 90, 98). However, these studies did not report interfactor correlations and/or did not compare measurement models (e.g., unidimensional or bifactor models), and it is difficult to be sure whether a single dimension competed with the multidimensionality found in these studies. However, regarding the study by Margolis et al. (79), our study found partial convergence, given that they concluded unidimensionality but with the addition of correlated errors and high factor loadings on the global factor. In the present study, the high psychometric similarity of the 19 items was considered a strong justification to produce a shortened version and to avoid the occurrence of correlated errors and maximize the parsimonious measurement of the MOS.

TABLE 7 MOS-S measurement invariance (group = sex).

	Configurational	Metric	Intercepts	Residuals
Fit measures				
WLSMV-x2 (df)	83.83** (40)	106.07** (63)	118.68** (70)	118.68** (78)
CFI	0.999	0.999	0.999	0.999
SRMR	0.038	0.039	0.039	0.039
Differences				
Δ_{CFI}	–	0.001	0.00	
Δ_{SRMR}	–	0.000	0.00	0.00

MOS-S, short version of MOS-SSS. **$p < 0.01$.

TABLE 8 Association with other variables: comparison MOS-SSS vs. MOS-S.

	MOS-SSS (19 items)	MOS–S	Z_{HMS}	95% CI Δ_r
SF-36				
Physical functioning	0.154	0.150	0.38	−0.01, 0.02
Role limitations due to physical health	0.19*	0.19*	−0.35	−0.02, 0.01
Pain	−0.15	−0.15	0.78	−0.01, 0.02
General health	0.01	0.02	0.96	−0.01, 0.03
Energy/fatigue	−0.03	−0.03	−0.11	−0.02, 0.01
Social functioning	−0.03	−0.06	2.86*	0.00, 0.04
Role limitations due to emotional problems	0.11	0.16	−0.44	−0.02, 0.01
Emotional well-being	−0.00	0.00	−0.46	−0.02, 01

Z_{HMS}: Hittner et al.'s (149) z-test. Δ_r: 95% confidence interval for difference. *$p < 0.006$ (nominal alpha with Bonferroni's correction: 0.05/8 scores = 0.006).

The present study also made progress in generating an abbreviated measure, given that a) factor loadings were highly similar in the unidimensional solution, and therefore the construct validity of the items did not differentiate between items that may have been more valid than others; b) an abbreviated measure is parsimonious to interpret, and c) this may be an important opportunity for choosing between screening measures or lengthy community surveys. This result adds to the existing abbreviated versions and may provide an equivalent measure of social support as these measures, given that the items are psychometrically similar with respect to their overall construct, social support. However, a comparative evaluation of these short versions with respect to subject classification and association with external variables is needed. Because previous brief versions were generated from models with different numbers of factors and an emphasis on tangible support [e.g., (73, 91)] or different samples of participants [e.g., mothers of children in clinical treatment; Gjesfjeld et al. (85)], the version obtained here may be more appropriate for the study sample. Given the strength of the validity of the items in their single dimension, it is likely that this version is generalizable to other groups of participants, but this assertion is conditional on future studies.

In the analysis of the equivalence between the unidimensional score with the 19 items and the abbreviated version, the high linear correlation between the two versions of the instrument indicates that the scaling of people based on the scores would be practically equal and that both scores can be used equivalently to differentiate the magnitude of perceived social support. When people are classified ordinally into groups of 3, 4 or 5 clusters, the agreement was also somewhat high, although it was higher in the tercile classification (i.e., low, medium, high), which suggested that the classification will be more equivalent between both MOS-SSS and MOS-S scores with fewer clusters. In summary, the analysis of the equivalence of the two versions of the MOS for differentiating subjects using direct scores or rankings (i.e., based on tertiles, quartiles, or quintiles) is highly similar. This high similarity is associated with the high coefficient of consistency obtained with both scores because it indicated that the error variance is very small, and the variability around the direct score will not produce severe changes in the description of the person assessed.

This level of reliability may indicate that the MOS score is useful in clinical practice, where individual decisions require highly accurate measures, i.e., with as little error variability as possible. Given that there appears to be no substantial loss of accuracy, according to the results of the equivalence between scores and internal consistency, the use of the abbreviated version is recommended for screening and clinical assessment purposes; specifically, for individual descriptions related to

the diagnosis and psychosocial variables derived from social support, for individual reports on the patient's social support status, and for making individual decisions on personalized interventions. Another implication of the obtained reliability results is that there was a strong replicability of the scores in a hypothetical situation where the MOS-S measurement is applied repeatedly. This indicates that the degree of error is small and advisable for clinical purposes because a reliability coefficient > 0.90 implies little probability of measurement error when applied for decision making on individual examinees. This is especially useful in individual interventions.

The association of the MOS with the SF-36 yielded low linear dependence, indicating divergence between the constructs assessed by these measures but also the possible specificity of these scores in this participant sample. In this sense, the physical role score was comparatively more strongly associated with the MOS, and it is very consistent with this research, given the basic characteristics of the sample. The study sample comprised patients with chronic diseases, and given the specific condition and severity of the disease, these patients will require support for roles that require moderate or intense physical exertion. In this sense, the new version has potential usefulness in the context of the importance of measuring social support for patients, since it has been shown that social support is an important determinant of physical and mental health because it moderates the effects of stress, improves the well-being of people, and has effects that extend to their family, social and work environment (16, 26–30, 38).

Among the limitations of this study, we can identify the use of non-probabilistic sampling so that population representativeness is not guaranteed. A second limitation is the cross-sectional design, which does not allow us to estimate the temporal reliability or to test the temporal stability of the factor model. A possible limitation is that participants with valid responses (i.e., false positives) may have been included in the removed group because of possible response bias. As a balance to this problem, we used two accepted methods (116–118) that detected two distinct patterns usually associated with possible response insufficiency/bias: extreme consistency (longstring) and inconsistency (outliers). A qualitative examination of this selection, and a sensitivity analysis, can verify whether the detection was correct and its impact large. But surely, some detection is preferable to none. Finally, the relationship with convergent measures of social support was not included, so this source of validity should be included in future studies. As a final note, replication of this work in future studies will allow more precise conclusions to be drawn regarding the factor structure of the MOS scale in patients with chronic disease at the primary health care. In addition, it will be possible to establish the relationship between social support and the degree of severity of chronic diseases and to carry out predictive studies between social support and the severity of chronic diseases in patients being attended in primary health care.

5. Conclusion

Due to the multiple clinical implications of social support in patients with chronic disease, the high global and national prevalence of these diseases, most of which are treated at the primary health care, and the instability of the internal structure of the MOS-SSS, the validity of this scale in patients with chronic disease was studied. Based on the results obtained in this study, a unidimensional representation of all MOS items was obtained. Since the items were psychometrically similar, a new 8-item, unidimensional, highly reliable, abbreviated version with invariant structure in the sex group of the patients was developed. This version showed adequate psychometric properties in patients with chronic disease at the primary health care.

Data availability statement

The original contributions presented in this study are included in this article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

This study involving human participants was reviewed and approved by the Research, Ethics, and Biosafety Committees of the Hospital Infantil de México Federico Gómez, Instituto Nacional de Salud, in Mexico City. Patients who agreed to participate in the study signed a letter of informed consent.

Author contributions

MÁNB and JGE: conceptualization. CM-S, MTD-G, MA-R, and FT-T: methodology. MA-R, OAFL, and MÁNB: validation. CM-S and JMR: formal analysis. MÁNB, CM-S, FT-T, and MA-R: investigation. LR-R and FT-T: resources. MTD-G, JMR, and GH-S: data curation. JHRC, FT-T, and OIGP: writing—original draft preparation. FT-T, MD-G, CM-S, GH-S, MA-R, and JHRC: writing—review and editing. CIA-G and LR-R: visualization. AL-L, CIA-G, and JM-R: supervision. JGE and FT-T: project administration and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding

This work presents some results of the research project HIM/2015/017/SSA.1207 "Effects of mindfulness training on psychological distress and quality of life of the family caregiver,"

Main researcher: FT-T, Ph.D. The present research was funded with federal funds for health research and was approved by the Commissions of Research, Ethics and Biosafety [Comisiones de Investigación, Ética y Bioseguridad], Hospital Infantil de México Federico Gómez, National Institute of Health. The funding agency had no control over the design of the study; the collection, analysis and interpretation of the data; or the writing of the manuscript.

Acknowledgments

We would like to thank MÁNB, Instituto Mexicano del Seguro Social, IMSS, Unidad de Medicina Familiar 31, Ciudad de México, Mexico.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. World Health Organization [WHO]. *Noncommunicable diseases*. Geneva, Switzerland: World Health Organization (2021).

2. The World Bank. *Panorama de la salud: latinoamérica y el caribe*. Paris: OECD (2020).

3. Shamah-Levy T, Vielma-Orozco E, Heredia-Hernández O, Romero-Martínez M, Mojica-Cuevas J, Cuevas-Nasu L, et al. *Encuesta nacional de salud y nutrición 2018–19: resultados nacionales*. Cuernavaca, México: Instituto Nacional de Salud Pública (2020).

4. INEGI. Comunicado de prensa núm. 538/19, 2019. Características de las defunciones registradas en méxico durante 2018: editorial universidad de sevilla. Ciudad de México: INEGI (2019).

5. Instituto Mexicano del Seguro Social. *Dirección de prestaciones médicas. Unidad de atención médica. Diagnóstico situacional de atención a la salud 2019. Programa presupuestario E-011: universidad nacional autonoma de mexico*. Ciudad de México: IMSS (2019). p. 31–7.

6. Instituto Mexicano del Seguro Social. *Informe de labores 2018-2019 y programa de actividades: instituto mexicano de tecnología del agua*. Ciudad de México: IMSS (2019).

7. G 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the global burden of disease study 2019. *Lancet*. (2020) 396:1204–22. doi: 10.1016/S0140-673630925-9

8. Toledano-Toledano F, Moral de la Rubia J, Frometa R, Betanzos F, Guzmán L, García M. The social support networks scale (SSNS) for family caregivers of children with cancer: a psychometric evaluation. *Int J Environ Res Public Health*. (2020) 17:7820. doi: 10.3390/ijerph17217820

9. Miller D, Rogers W, Kazis L, Spiro A, Ren X, Haffer S. Patients' self-report of diseases in the medicare health outcomes survey based on comparisons with linked survey and medical data from the veterans health administration. *J Ambul Care Manag*. (2008) 31:161–77. doi: 10.1097/01.jac.0000314707.88160.9c

10. Rothrock N, Hays R, Spritzer K, Yount S, Riley W, Cella D. Relative to the general US population, chronic diseases are associated with poorer health-related quality of life as measured by the patient-reported outcomes measurement information system (PROMIS). *J Clin Epidemiol*. (2010) 63:1195–204. doi: 10.1016/j.jclinepi.2010.04.012

11. Verbrugge L, Lepkowski J, Imanaka Y. Comorbidity and its impact on disability. *Milbank*. (1989) 67:450–85. doi: 10.2307/3350223

12. Cieza A, Stucki G. The international classification of functioning disability and health: its development process and content validity. *Eur J Phys Rehabil Med*. (2008) 44:303–13.

13. Toledano-Toledano F, Moral de la Rubia J. Factors associated with anxiety in family caregivers of children with chronic diseases. *Biopsychosoc Med*. (2018) 12:20. doi: 10.1186/s13030-018-0139-7

14. Toledano-Toledano F, Luna D, Moral de la Rubia J, Martínez Valverde S, Bermúdez Morón C, Salazar García M, et al. Psychosocial factors predicting resilience in family caregivers of children with cancer: a cross-sectional study. *Int J Environ Res Public Health*. (2021) 18:748. doi: 10.3390/ijerph18020748

15. Cassel J. Psychosocial processes and "stress": theoretical formulation. *Int J Health Serv*. (1974) 4:471–82. doi: 10.2190/wf7x-y1l0-bfkh-9qu2

16. Cobb S. Social support as a moderator of life stress. *Psychosom Med*. (1976) 38:300–14. doi: 10.1097/00006842-197609000-00003

17. Pierce G, Sarason B, Sarason I, Joseph H, Henderson C. Conceptualizing and assessing social support in the context of the family. In: Pierce G, Sarason B, Sarason I editors. *Handbook of social support and the family*. Boston, MA: Springer (1996). p. 3–23.

18. Aranda B, Pando M. Conceptualización del apoyo social y las redes de apoyo social [conceptualization of the social support and the social support network]. *Rev Investig Psicol*. (2013) 16:233–45. doi: 10.15381/rinvp.v16i1.3929

19. Cohen S, Gottlieb B, Underwood L. Social relationships and health: challenges for measurement and intervention. *Adv Mind Body Med*. (2001) 17:129–41.

20. González-Quiñones J, Restrepo-Chavarriaga G. Prevalencia de felicidad en ciclos vitales y relación con redes de apoyo en población colombiana. *Rev Salud Pública*. (2010) 12:228–38. doi: 10.1590/s0124-00642010000200006

21. World Health Organization[WHO]. *Division of health promotion, education, and communication*. Geneva: World Health Organization (1998).

22. Wills T, Shinar O. Measuring perceived and received social support. In: Cohen S, Underwood L, Gottlieb B editors. *Social support measurement and intervention: a guide for health and social scientists*. New York: Oxford University Press (2000). p. 86–135.

23. Lin N. Coceptualizing social support. In: Lin N, Dean A, Ensel W editors. *Social support, life events, and depression*. Orlando, FL: Academic Press, INC (1986). p. 17–48.

24. Pattison E. A theoretical-empirical base for social system therapy. In: Foulks E, Wintrob R, Westermeyer J, Favazza A editors. *Current perspectives in cultural psychiatry*. New York: Spectrum (1977). p. 217–53.

25. Barrera M, Ainlay S. The structure of social support: a conceptual and empirical analysis. *J Community Psychol*. (1983) 11:133–43. doi: 10.1002/1520-6629(198304)11:23.0.co;2-l

26. House J, Umberson D, Landis K. Structures and processes of social support. *Annu Rev Sociol*. (1988) 14:293–318. doi: 10.1146/annurev.so.14.080188.001453

27. Reblin M, Uchino B. Social and emotional support and its implication for health. *Curr Opin Psychiatry*. (2008) 21:201–5. doi: 10.1097/YCO.0b013e3282f3ad89

28. Harandi T, Taghinasab M, Nayeri T. The correlation of social support with mental health: a meta-analysis. *Electron Physician.* (2017) 9:5212–22. doi: 10.19082/5212

29. McCoy M, Theeke LA. A systematic review of the relationships among psychosocial factors and coping in adults with type 2 diabetes mellitus. *Int J Nurs Sci.* (2019) 6:468–77. doi: 10.1016/j.ijnss.2019.09.003

30. Scott H, Pitman A, Kozhuharova P, Lloyd-Evans B. A systematic review of studies describing the influence of informal social support on psychological wellbeing in people bereaved by sudden or violent causes of death. *BMC Psychiatry.* (2020) 20:265. doi: 10.1186/s12888-020-02639-4

31. Berkman L, Syme S. Social networks, host resistance, and mortality: a nine-year follow-up study of alameda county residents. *Am J Epidemiol.* (1979) 109:186–204. doi: 10.1093/oxfordjournals.aje.a112674

32. House J, Robbins C, Metzner H. The association of social relationships and activities with mortality: prospective evidence from the tecumseh community health study. *Am J Epidemiol.* (1982) 116:123–40. doi: 10.1093/oxfordjournals.aje.a113387

33. Stringhini S, Berkman L, Dugravot A, Ferrie J, Marmot M, Kivimaki M, et al. Socioeconomic status, structural and functional measures of social support, and mortality: the british whitehall II cohort study, 1985-2009. *Am J Epidemiol.* (2012) 175:1275–83. doi: 10.1093/aje/kwr461

34. Hill T, Uchino B, Eckhardt J, Angel J. Perceived social support trajectories and the all-cause mortality risk of older Mexican American women and men. *Res Aging.* (2016) 38:374–98. doi: 10.1177/0164027515620239

35. Litwin H. Social networks and well-being: a comparison of older people in mediterranean and non-mediterranean countries. *J Gerontol B Psychol Sci Soc Sci.* (2010) 65:599–608. doi: 10.1093/geronb/gbp104

36. Bélanger E, Ahmed T, Vafaei A, Curcio C, Phillips S, Zunzunegui M. Sources of social support associated with health and quality of life: a cross-sectional study among Canadian and Latin American older adults. *BMJ Open.* (2016) 6:e011503. doi: 10.1136/bmjopen-2016-011503

37. Vassilev I, Rogers A, Kennedy A, Koetsenruijter J. The influence of social networks on self-management support: a metasynthesis. *BMC Public Health.* (2014) 14:719. doi: 10.1186/1471-2458-14-719

38. Rivera N, McGuinn L, Osorio-Valencia E, Martinez-Medina S, Schnaas L, Wright R, et al. Changes in depressive symptoms, stress and social support in Mexican women during the COVID-19 pandemic. *Int J Environ Res Public Health.* (2021) 18:8775. doi: 10.3390/ijerph18168775

39. Mendoza-Núñez V, González-Mantilla F, Correa-Muñoz E, Retana-Ugalde R. Relationship between social support networks and physical functioning in older community-dwelling Mexicans. *Int J Environ Res Public Health.* (2017) 14:993. doi: 10.3390/ijerph14090993

40. Iglesias S, Arias A. Structural and functional social support in elderly objective and subjective health ratings. *Eur J Investig Health Psychol Educ.* (2015) 5:243–52. doi: 10.1989/ejihpe.v5i2.116

41. Bolger N, Amarel D. Effects of social support visibility on adjustment to stress: experimental evidence. *J Pers Soc Psychol.* (2007) 92:458–75. doi: 10.1037/0022-3514.92.3.458

42. Schulz R, Sherwood P. Physical and mental health effects of family caregiving. *Am J Nurs.* (2008) 108:23–7. doi: 10.1097/01.NAJ.0000336406.45248.4c

43. Barrera M. Distinctions between social support concepts, measures, and models. *Am J Community Psychol.* (1986) 14:413–45. doi: 10.1007/bf00922627

44. Vangelisti A. Challenges in conceptualizing social support. *J Soc Pers Relationsh.* (2009) 26:39–51. doi: 10.1177/0265407509105520

45. Eagle D, Hybels C, Proeschold-Bell R. Perceived social support, received social support, and depression among clergy. *J Soc Pers Relationsh.* (2019) 36:2055–73. doi: 10.1177/0265407518776134

46. Streeter C, Franklin C. Defining and measuring social support: guidelines for social work practitioners. *Res Social Work Pract.* (1992) 2:81–98. doi: 10.1177/104973159200200107

47. Gottlieb B, Bergen A. Social support concepts and measures. *J Psychosom Res.* (2010) 69:511–20. doi: 10.1016/j.jpsychores.2009.10.001

48. Holahan C, Moos R. Social support and psychological distress: a longitudinal analysis. *J Abnorm Psychol.* (1981) 90:365–70. doi: 10.1037/0021-843x.90.4.365

49. Barrera M, Sandler I, Ramsay T. Preliminary development of a scale of social support: studies on college students. *Am J Community Psychol.* (1981) 9:435–47. doi: 10.1007/bf00918174

50. Cutrona C, Russell D. The provisions of social relationships and adaptation to stress. In: Jones W, Perlman D editors. *Advances in personal relationships: a research annual*. Greenwich, Connecticut: JAI Press (1981). p. 37–67.

51. Flaherty J, Gaviria F, Pathak D. The measurement of social support: the social support network inventory. *Compr Psychiatry.* (1983) 24:521–9. doi: 10.1016/0010-440x90019-6

52. Procidano M, Heller K. Measures of perceived social support from friends and from family: three validation studies. *Am J Community Psychol.* (1983) 11:1–24. doi: 10.1007/bf00898416

53. Sarason I, Levine H, Basham R, Sarason B. Assessing social support: the social support questionnaire. *J Pers Soc Psychol.* (1983) 44:127–39. doi: 10.1037/0022-3514.44.1.127

54. Vaux A, Harrison D. Support network characteristics associated with support satisfaction and perceived support. *Am J Community Psychol.* (1985) 13:245–65. doi: 10.1007/bf00914932

55. Vaux A, Phillips J, Holly L, Thomson B, Williams D, Stewart D. The social support appraisals (SS-A) scale: studies of reliability and validity. *Am J Community Psychol.* (1986) 14:195–219. doi: 10.1007/bf00911821

56. Broadhead W, Gehlbach S, De Gruy F, Kaplan B. The duke???UNC functional social support questionnaire. *Med Care.* (1988) 26:709–23. doi: 10.1097/00005650-198807000-00006

57. Tracy E, Whittaker J. The social network map: assessing social support in clinical practice. *Fam Soc J Contemp Hum Serv.* (1990) 71:461–70. doi: 10.1177/104438949007100802

58. Mitchell P, Powell L, Blumenthal J, Norten J, Ironson G, Pitula C, et al. A short social support measure for patients recovering from myocardial infarction: the ENRICHD social support inventory. *J Cardiopulm Rehabil.* (2003) 23:398–403. doi: 10.1097/00008483-200311000-00001

59. Ong A, Ward C. The construction and validation of a social support measure for sojourners: the index of sojourner social support (ISSS) scale. *J Cross Cult Psychol.* (2005) 36:637–61. doi: 10.1177/0022022105280508

60. Littlewood K, Swanke J, Strozier A, Kondrat D. Measuring social support among kinship caregivers: validity and reliability of the family support scale. *Child Welfare.* (2012) 91:59–78.

61. Malecki C, Demary M. Measuring perceived social support: development of the child and adolescent social support scale (CASSS). *Psychol Schools.* (2002) 39:1–18. doi: 10.1002/pits.10004

62. Gordon-Hollingsworth A, Thompson J, Geary M, Schexnaildre M, Lai B, Kelley M. Social support questionnaire for children: development and initial validation. *Meas Eval Couns Dev.* (2016) 49:122–44. doi: 10.1177/0748175615596780

63. Dambi J, Corten L, Chiwaridzo M, Jack H, Mlambo T, Jelsma J. A systematic review of the psychometric properties of the cross-cultural translations and adaptations of the multidimensional perceived social support scale (MSPSS). *Health Qualt Life Outcomes.* (2018) 16:80. doi: 10.1186/s12955-018-0912-0

64. Lever J, García G, Estrada A. Elaboración de una escala de apoyo social (EAS) para adultos. *Univ Psychol.* (2013) 12:129–37. doi: 10.11144/javeriana.upsy12-1.eeas

65. Domínguez Guedea M, Mandujano Jaquez M, Georgina Quintero M, Sotelo Quiñónez T, Gaxiola Romero J, Valencia Maldonado J. Escala de apoyo social para cuidadores familiares de adultos mayores mexicanos. *Univ Psychol.* (2013) 12:391–402. doi: 10.11144/Javeriana.UPSY12-2.easc

66. García-Torres M, García-Méndez M, Rivera-Aragón S. Apoyo social en adultos mexicanos: validación de una escala. *Acta Invest Psicol.* (2017) 7:2561–7. doi: 10.1016/j.aipprr.2017.02.004

67. Fontes M, Heredia M, Peñaloza J, Cedeño M, Rodríguez-Orozco A. Funcionamiento familiar y su relación con las redes de apoyo social en una muestra de morelia. *México Salud Mental.* (2012) 35:147–54.

68. Santos-Vega M, Ortega-Andeane P, Toledano-Toledano F. Validez y confiablidad de la escala de apoyo social percibido (MSPSS). *Psicooncología.* (2021) 18:333–45. doi: 10.5209/psic.77756

69. Sherbourne C, Stewart A. The MOS social support survey. *Soc Sci Med.* (1991) 32:705–14. doi: 10.1016/0277-953690150-b

70. Wang W, Zheng X, He H, Thompson D. Psychometric testing of the Chinese mandarin version of the medical outcomes study social support survey in patients with coronary heart disease in mainland China. *Quali Life Res.* (2013) 22:1965–71. doi: 10.1007/s11136-012-0345-x

71. Yu D, Lee D, Woo J. Psychometric testing of the Chinese version of the medical outcomes study social support survey (MOS-SSS-C). *Res Nurs Health.* (2004) 27:135–43. doi: 10.1002/nur.20008

72. Shyu Y, Tang W, Liang J, Weng L. Psychometric testing of the social support survey on a Taiwanese sample. *Nurs Res.* (2006) 55:411–7. doi: 10.1097/00006199-200611000-00005

73. Holden L, Lee C, Hockey R, Ware R, Dobson A. Validation of the MOS social support survey 6-item (MOS-SSS-6) measure with two large population-based samples of Australian women. *Qual Life Res.* (2014) 23:2849–53. doi: 10.1007/s11136-014-0741-5

74. Anderson D, Bilodeau B, Deshaies G, Gilbert M, Jobin J. Validation canadienne-francaise du "MOS social support survey" [French-Canadian validation of the MOS social support survey]. *Can J Cardiol.* (2005) 21:867–73.

75. Fachado A, Martinez A, Villalva C, Pereira M. Adaptação cultural e validação da versão Portuguesa questionário medical outcomes study social support survey (MOS-SSS) [cultural adaptation and validation of the medical outcomes study social support survey questionnaire (MOS-SSS)]. *Acta Méd Port.* (2007) 20:525–34.

76. Ahumada L, Castillo J, Muñoz E, Moruno I. Validación del cuestionario MOS de apoyo social en atención primaria. *Med Fam.* (2005) 6:10–8.

77. Basurto A, Román S, Villalobos E, Pérez V, Rosas A. Adaptación y validación del cuestionario MOS de apoyo socialen pacientes mexicanos con VIH+. *Rev Latinoam Med Conduct.* (2014) 4:93–101.

78. Navarrete B, Vázquez Ó, Alcaraz R, Penedo F, Lerma A. Propiedades psicométricas del cuestionario MOS de apoyo social en una muestra de pacientes con enfermedades cardiovasculares en población mexicana. *Psicol Salud.* (2021) 31:225–35. doi: 10.25009/pys.v31i2.2691

79. Margolis R, Bellin M, Sacco P, Harrington D, Butz A. Evaluation of MOS social support in low-income caregivers of African American children with poorly controlled asthma. *J Asthma.* (2019) 56:951–8. doi: 10.1080/02770903.2018.1510504

80. Zanini D, Peixoto E. Social support scale (MOS-SSS): analysis of the psychometric properties via item response theory. *Paidéia.* (2016) 26:359–68. doi: 10.1590/1982-43272665201612

81. Zanini D, Verolla-Moura A, Queiroz I. Apoio social: aspectos da validade de constructo em estudantes universitários. [Social support: validity aspects of the construct in undergraduation student]. *Psicol Em Estudo.* (2009) 14:195–202. doi: 10.1590/s1413-73722009000100023

82. Westaway M, Seager J, Rheeder P, Van Zyl D. The effects of social support on health, well-being and management of diabetes mellitus: a black South African perspective. *Ethn Health.* (2005) 10:73–89. doi: 10.1080/1355785052000323047

83. Alonso A, Montes A, Menéndez V, Graça M. Adaptação cultural e validação da versão portuguesa questionário medical outcomes study social support survey (MOS-SSS). *Acta Méd Port.* (2007) 20:525–33.

84. Requena G, Salamero M, Gil F. Validación del cuestionario MOS-SSS de apoyo social en pacientes con cáncer. *Med Clín.* (2007) 128:687–91. doi: 10.1157/13102357

85. Gjesfjeld C, Greeno C, Kim K. A confirmatory factor analysis of an abbreviated social support instrument: the MOS-SSS. *Res Soc Work Pract.* (2007) 18:231–7. doi: 10.1177/1049731507309830

86. Espínola S, Enrique H. Validacioìn Argentina del cuestionario MOS de apoyo social percibido. *Psicodebate.* (2007) 7:155–68. v7i0.433 doi: 10.18682/pd

87. Pais-Ribeiro J, Ponte A. Propriedades métricas da versão portuguesa da escala de suporte social do MOS (MOS social support survey) com idosos [metric properties of the portuguese version of the MOS social support survey with a sample of aged people]. *Psicol Saúde Doenças.* (2009) 10:163–74.

88. Robitaille A, Orpana H, McIntosh C. Psychometric properties, factorial structure, and measurement invariance of the english and French versions of the medical outcomes study social support scale. *Health Rep.* (2011) 22:33–40.

89. Ashing-Giwa K, Rosales M. A cross-cultural validation of patient-reported outcomes measures: a study of breast cancers survivors. *Qual Life Res.* (2012) 22:295–308. doi: 10.1007/s11136-012-0140-8

90. Londoño N, Rogers H, Castilla J, Posada S, Jaramillo M, Oliveros M, et al. Validación en colombia del cuestionario MOS de apoyo social. *Int J Psychol Res.* (2012) 5:142–50.

91. Moser A, Stuck A, Silliman R, Ganz P, Clough-Gorr K. The eight-item modified medical outcomes study social support survey: psychometric evaluation showed excellent performance. *J Clin Epidemiol.* (2012) 65:1107–16. doi: 10.1016/j.jclinepi.2012.04.007

92. Soares A, Biasoli I, Scheliga A, Baptista R, Brabo E, Morais J, et al. Validation of the Brazilian portuguese version of the medical outcomes study-social support survey in Hodgkin's lymphoma survivors. *Support Care Cancer.* (2012) 20:1895–900. doi: 10.1007/s00520-011-1292-8

93. Gomez-Campelo P, Pérez-Moreno E, de Burgos-Lunar C, Bragado-Álvarez C, Jiménez-García R, Salinero-Fort M. Psychometric properties of the eight-item modified medical outcomes study social support survey based on Spanish outpatients. *Qual Life Res.* (2014) 23:2073–8. doi: 10.1007/s11136-014-0651-6

94. Giangrasso B, Casale S. Psychometric properties of the medical outcome study social support survey with a general population sample of undergraduate students. *Soc Indic Res.* (2014) 116:185–97. doi: 10.1007/s11205-013-0277-z

95. Conte K, Schure M, Goins R. Correlates of social support in older American Indians: the native elder care study. *Aging Ment Health.* (2014) 19:835–43. doi: 10.1080/13607863.2014.967171

96. Higgins G, Marcum C, Golder S, Hall M, Logan T. Confirmatory factor analysis of the medical outcomes study – social support survey: examining the factor structure among victimized women on probation and parole. *Am J Crim Justice.* (2015) 40:811–22. doi: 10.1007/s12103-015-9290-x

97. Norhayati M, Aniza A, Nik Hazlina N, Azman M. Psychometric properties of the revised malay version medical outcome study social support survey using confirmatory factor analysis among postpartum mothers. *Asia Pac Psychiatry.* (2015) 7:398–405. doi: 10.1111/appy.12184

98. Yu Y, Yang J, Shiu C, Simoni J, Xiao S, Chen W, et al. Psychometric testing of the Chinese version of the medical outcomes study social support survey among people living with HIV/AIDS in China. *Appl Nurs Res.* (2015) 28:328–33. doi: 10.1016/j.apnr.2015.03.006

99. Togari T, Yokoyama Y. Application of the eight-item modified medical outcomes study social support survey in Japan: a national representative cross-sectional study. *Qual Life Res.* (2016) 25:1151–8. doi: 10.1007/s11136-015-1155-8

100. Priede A, Andreu-Vaillo Y, Martínez-López P, Ruiz-Torres M, Hoyuela F, González-Blanch C. Validación de la escala mos-sss de apoyo social en una muestra de pacientes oncológicos recién diagnosticados. *Proceedings of the 9th international and 14th national congress of clinical psychology*. Spain: Santander (2016). p. 45–53.

101. Yilmaz T, Bozo Ö. Psychometric qualities of medical outcomes study social support survey (MOS-SSS) in Turkish culture. *Int J Eurasia Soc Sci.* (2019) 10:899–915. doi: 10.35826/ijoess.2516

102. Martín-Carbonell M, Cerquera-Córdoba A, Fernández-Daza M, Higuita J, Patrignani G, Martel M, et al. Estructura factorial del cuestionario de apoyo social MOS en ancianos colombianos con dolor crónico. *Ter Psicol.* (2019) 37:211–24. doi: 10.4067/s0718-48082019000300211

103. Bavarsad M, Foroughan M, Zanjari N, Shushtari Z, Harouni G. Psychometric properties of modified MOS social support survey 5-item (MSSS-5-item) among Iranian older adults. *BMC Geriatr.* (2021) 21:409. doi: 10.1186/s12877-021-02353-0

104. Jackson D, Gillaspy J, Purc-Stephenson R. Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychol Methods.* (2009) 14:6–23. doi: 10.1037/a0014694

105. Beauducel A, Herzberg P. On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Struct Equ Model Multidiscip J.* (2006) 13:186–203. doi: 10.1207/s15328007sem1302_2

106. Flora D, Curran P. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychol Methods.* (2004) 9:466–91. doi: 10.1037/1082-989X.9.4.466

107. Li C. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behav Res Methods.* (2016) 48:936–49. doi: 10.3758/s13428-015-0619-7

108. Stewart A, Nápoles-Springer A. Advancing health disparities research. *Med Care.* (2003) 41:1207–20. doi: 10.1097/01.mlr.0000093420.27745.48

109. Sociedad Mexicana de Psicología. *Código ético del psicólogo*. Trillas: Ciudad de México (2010).

110. Ley General de Salud. *Titulo quinto, investigación para la salud, capítulo único: dykinson*. Ciudad de México: Diario Oficial de la Federación (2022). p. 45–6.

111. American Psychological Association. *Ethical principles of psychologists and code of conduct*. Washington, DC: Author (2021).

112. World Medical Association. World medical association declaration of helsinki: ethical principles for medical research involving human subjects. *JAMA.* (2013) 310:2191–4. doi: 10.1001/jama.2013.281053

113. Toledano-Toledano F, Rodríguez-Rey R, Moral de la Rubia J, Luna D. A Sociodemographic variables questionnaire (Q-SV) for research on family caregivers of children with chronic disease. *BMC Psychol.* (2019) 7:85. doi: 10.1186/s40359-019-0350-8

114. Zúniga M, Carrillo-Jiménez G, Fos P, Gandek B, Medina-Moreno M. Evaluación del estado de salud con la encuesta SF-36: resultados preliminares en méxico. *Salud Pública Méx.* (1999) 41:110–8. doi: 10.1590/s0036-36341999000200005

115. Mahalanobis P. On the generalized distance in statistics. *Proc Natl Inst Sci India.* (1936) 2:49–55.

116. Johnson J. Ascertaining the validity of individual protocols from Web-based personality inventories. *J Res Pers.* (2005) 39:103–29. doi: 10.1016/j.jrp.2004.09.009

117. Curran P. Methods for the detection of carelessly invalid responses in survey data. *J Exp Soc Psychol.* (2016) 66:4–19. doi: 10.1016/j.jesp.2015.07.006

118. Huang J, Curran P, Keeney J, Poposki E, DeShon R. Detecting and deterring insufficient effort responding to surveys. *J Bus Psychol.* (2012) 27:99–114. doi: 10.1007/s10869-011-9231-8

119. Meade A, Craig S. Identifying careless responses in survey data. *Psychol Methods.* (2012) 17:437–55. doi: 10.1037/a0028085

120. Yentes R, Wilhelm F. *Careless: procedures for computing indices of careless responding (R packages version 1.2.0)*. (2018).

121. Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends Sport Sci.* (2014) 1:1–25. doi: 10.3724/sp.j.1041.2016.00435

122. Salas-Blas E, Merino-Soto C, Pérez-Amezcua B, Toledano-Toledano F. Social networks addiction (SNA-6) – short: validity of measurement in mexican youths. *Front Psychol.* (2022) 12:774847. doi: 10.3389/fpsyg.2021.774847

123. Mangiafico S. *Rcompanion: functions to support extension education program evaluation (R package version 2.3.26.)*. (2021).

124. Korkmaz S, Goksuluk D, Zararsiz G. MVN: an R package for assessing multivariate normality. *Res J.* (2014) 6:151–62. doi: 10.32614/rj-2014-031

125. Mokken R. *A theory and procedure of scale analysis*. Netherlands: Mouton (1971).

126. Molenaar I. Nonparametric models for polytomous responses. In: van der Linden W, Hambleton R editors. *Handbook of modern item response theory*. New York, NY: Springer (1997). p. 369–80.

127. Crişan D, Tendeiro J, Meijer R. On the practical consequences of misfit in mokken scaling. *Appl Psychol Meas.* (2020) 44:482–96. doi: 10.1177/0146621620920925

128. Palmgren P, Brodin U, Nilsson G, Watson R, Stenfors T. Investigating psychometric properties and dimensional structure of an educational environment measure (DREEM) using mokken scale analysis – a pragmatic approach. *BMC Med Educ.* (2018) 18:235. doi: 10.1186/s12909-018-1334-8

129. Van der Ark L. New developments in mokken scale analysis in R. *J Stat Softw.* (2012) 48:1–27. doi: 10.18637/jss.v048.i05

130. Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika.* (1965) 30:179–85. doi: 10.1007/bf02289447

131. Garrido L, Abad F, Ponsoda V. Are fit indices really fit to estimate the number of factors with categorical variables? Some cautionary findings via monte carlo simulation. *Psychol Methods.* (2016) 21:93–111. doi: 10.1037/met0000064

132. Revelle W. *Psych: procedures for personality and psychological research*. Evanston, Illinois, USA: Northwestern University (2021).

133. Finney S, DiStefano C. Nonnormal and categorical data in structural equation modeling. In: Hancock G, Mueller R editors. *Structural equation modeling: a second course*. North Carolina: IAP Information Age Publishing (2013). p. 439–92.

134. Henseler J, Ringle C, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J Acad Mark Sci.* (2015) 43:115–35. doi: 10.1007/s11747-014-0403-8

135. Roemer E, Schuberth F, Henseler J. HTMT2–an improved criterion for assessing discriminant validity in structural equation modeling. *Ind Manag Data Syst.* (2021) 121:2637–50. doi: 10.1108/imds-02-2021-0082

136. Cheung G, Wang C. Current approaches for assessing convergent and discriminant validity with SEM: issues and solutions. *Acad Manag Proc.* (2017) 2017:12706. doi: 10.5465/ambpp.2017.12706abstract

137. Franke G, Sarstedt M. Heuristics versus statistics in discriminant validity testing: a comparison of four procedures. *Internet Res.* (2019) 29:430–47. doi: 10.1108/intr-12-2017-0515

138. Voorhees C, Brady M, Calantone R, Ramirez E. Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies. *J Acad Mark Sci.* (2016) 44:119–34. doi: 10.1007/s11747-015-0455-4

139. Bagozzi R, Yi Y, Phillips L. Assessing construct validity in organizational research. *Adm Sci Q.* (1991) 36:421–58. doi: 10.2307/2393203

140. Kline R. *Principles and practice of structural equation modeling*. New York, NY: Guilford Press (2011).

141. Girard T, Christensen B. Clarifying problems and offering solutions for correlated error when assessing the validity of selected-subtest short forms. *Psychol Assess.* (2008) 20:76–80. doi: 10.1037/1040-3590.20.1.76

142. Stanton J, Sinar E, Balzer W, Smith P. Issues and strategies for reducing the length of self-report scales. *Pers Psychol.* (2002) 55:167–94. doi: 10.1111/j.1744-6570.2002.tb00108.x

143. Levy P. The correction for spurious correlation in the evaluation of short-form tests. *J Clin Psychol.* (1967) 23:84–6. doi: 10.1002/1097-4679(196701)23:13.0.co;2-2

144. Gwet K. *Handbook of inter-rater reliability*. Gaithersburg, MD: Advanced Analytics, LLC (2014).

145. Chen F. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Model Multidiscip J.* (2007) 14:464–504. doi: 10.1080/10705510701301834

146. Green S, Yang Y. Reliability of summed item scores using structural equation modeling: an alternative to coefficient alpha. *Psychometrika.* (2009) 74:155–67. doi: 10.1007/s11336-008-9099-3

147. Molenaar I, Sijtsma K. Mokken's approach to reliability estimation extended to multicategory items. *Kwant Methoden.* (1988) 9:115–26.

148. Kelley K. Methods for the behavioral, educational, and social sciences: an R package. *Behav Res Methods.* (2007) 39:979–84. doi: 10.3758/bf03192993

149. Hittner J, May K, Silver NC. A monte carlo evaluation of tests for comparing dependent correlations. *J Gen Psychol.* (2003) 130:149–68. doi: 10.1080/00221300309601282

150. Zou G. Toward using confidence intervals to compare correlations. *Psychol Methods.* (2007) 12:399–413. doi: 10.1037/1082-989x.12.4.399

151. Diedenhofen B, Musch J. Cocor: a comprehensive solution for the statistical comparison of correlations. *PLoS One.* (2015) 10:e0121945. doi: 10.1371/journal.pone.0121945

152. Dominguez-Lara S, Fernández-Arata M, Merino-Soto C, Navarro-Loli J, De la Cruz G. Inventario de violencia y acoso psicológico en el trabajo (IVAPT) en colombia: el peligroso little jiffy. *Salud Uninorte.* (2018) 34:536–7. doi: 10.14482/sun.34.2.658.47

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 16 January 2023
DOI 10.3389/fpsyt.2022.985456

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University, Germany

REVIEWED BY
Maryam Rassouli,
Shahid Beheshti University of Medical Sciences, Iran
Hadis Ashrafizadeh,
Dezful University of Medical Sciences, Iran
Azam Shirinabadi Farahani,
Shahid Beheshti University of Medical Sciences, Iran

*CORRESPONDENCE
Filiberto Toledano-Toledano
✉ filiberto.toledano.phd@gmail.com

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 03 July 2022
ACCEPTED 23 December 2022
PUBLISHED 16 January 2023

CITATION
Jiménez S, Moral de la Rubia J, Varela-Garay RM, Merino-Soto C and Toledano-Toledano F (2023) Resilience measurement scale in family caregivers of children with cancer: Multidimensional item response theory modeling.
Front. Psychiatry 13:985456.
doi: 10.3389/fpsyt.2022.985456

Resilience measurement scale in family caregivers of children with cancer: Multidimensional item response theory modeling

Said Jiménez[1], José Moral de la Rubia[2], Rosa María Varela-Garay[3], Cesar Merino-Soto[4] and Filiberto Toledano-Toledano[1,5,6]*

[1]Unidad de Investigación en Medicina Basada en Evidencias, Hospital Infantil de México Federico Gómez, National Institute of Health, Mexico City, Mexico, [2]Facultad de Psicología, Universidad Autónoma de Nuevo León, Monterrey, Mexico, [3]Departamento de Trabajo Social y Servicios Sociales, Facultad de Ciencias Sociales, Universidad Pablo de Olavide, Seville, Spain, [4]Instituto de Investigación en Psicología, Universidad de San Martin de Porres, Lima, Peru, [5]Unidad de Investigación Sociomédica, Instituto Nacional de Rehabilitación Luis Guillermo Ibarra Ibarra, Mexico City, Mexico, [6]Dirección de Investigación y Diseminación del Conocimiento, Instituto Nacional de Ciencias e Innovación para la Formación de Comunidad Científica, INDEHUS, Mexico City, Mexico

Background: Currently, information about the psychometric properties of the Resilience Measurement Scale (RESI-M) in family caregivers of children with cancer according to item response theory (IRT) is not available; this information could complement and confirm the findings available from classical test theory (CTT). The objective of this study was to test the five-factor structure of the RESI-M using a full information confirmatory multidimensional IRT graded response model and to estimate the multidimensional item-level parameters of discrimination (MDISC) and difficulty (MDIFF) from the RESI-M scale to investigate its construct validity and level of measurement error.

Methods: An observational study was carried out, which included a sample of 633 primary caregivers of children with cancer, who were recruited through nonprobabilistic sampling. The caregivers responded to a battery of tests that included a sociodemographic variables questionnaire, the RESI-M, and measures of depression, quality of life, anxiety, and caregiver burden to explore convergent and divergent validity.

Results: The main findings confirmed a five-factor structure of the RESI-M scale, with RMSEA = 0.078 (95% CI: 0.075, 0.080), TLI = 0.90, and CFI = 0.91. The estimation of the MDISC and MDIFF parameters indicated different values for each item, showing that all the items contribute differentially to the measurement of the dimensions of resilience.

Conclusion: That regardless of the measurement approach (IRT or CTT), the five-factor model of the RESI-M is valid at the theoretical, empirical, and methodological levels.

KEYWORDS

resilience, psychometric properties, family caregivers, cancer, item response theory

1. Introduction

Childhood cancer has serious repercussions on the physical and psychological health of pediatric patients, their families and their caregivers; caregiving can be experienced as a stressful process that can cause psychological and physical effects and consequences (1–4). Childhood cancer patients and their families often experience anxiety, depression, and parental stress (5–7); poor health (8); and social and economic overload and caregiver burnout (9). Therefore, caregiving has effects on the quality of life, caregiver profile and resilience of families caring for children with cancer (10, 11). The research literature has identified a number of contextual factors and sociodemographic characteristics in family caregivers of pediatric patients that increase the risk for physical and psychological health impacts (12). The main demographic variables include gender (13), unemployment (5), low income (14), low levels of education (15), social support networks (16), caregiver marital status (17), number of children in the family (18), child age (19), and the psychosocial profile of family caregivers (11). Contextual factors include the time elapsed since diagnosis (20), subjective perceptions of disease severity of both patients and caregivers (21), the duration of the disease (22), the personality type of the parents (23), and the duration and impact of care (24, 25). In this regard, evidence indicates that the term "caregiver", which was first used in 1966 referring to those "helping those who suffer", is a multidimensional construct, and its use in research lacks a coherent conceptualization and an operational definition (26). However, in chronic illness contexts, the family caregiver has been defined as the person who has a significant emotional bond with the patient; who may be a family member who is part of the patient's family life cycle; who offers emotional, expressive, instrumental and tangible support; and who provides assistance and comprehensive care during the chronic illness, acute illness or disability of a child, adult, or elderly person (11). In this sense, resilience to chronic illness is a process of positive adaptation despite the loss of health, which implies the development of vitality and skills to overcome the negative effects of adversity, risk, and vulnerability caused by the disease (27).

The measurement and assessment of resilience depends on how it is defined, and the factors associated with it (28). One of the measurements developed for Mexican population was the Resilience Measurement Scale (RESI-M) (29). It is an instrument of 43 items with a Likert-type scale with four response options, ranging from 1 "totally disagree" to 4 "totally agree." The items of the RESI-M were derived from two instruments that measure resilience that are widely used in the international literature, namely, the Connor-Davidson Resilience Scale (CD-RISC) (30) and the Resilience Scale for Adults (RSA) (31). Both scales measure resilience in adults. According to Palomar and Gómez (29), the factors of the RESI-M have been defined as follows: (1). Strength and Self-Confidence refers to the clarity that individuals have about their objectives, the effort they make to achieve their goals, the confidence they have that they will succeed and the optimism, strength and tenacity with which they face their challenges. (2). Social Competence indicates the competence of individuals to relate to others and the ease with which they make new friends, make people laugh and enjoy a conversation. (3). Family Support addresses family relationships and family support, loyalty among family members, and family members sharing similar views of life and spending time together. (4). Social Support, mainly from friends, points to the individual having people who can help, give encouragement and care about him or her in difficult times. (5). Structure refers to the ability of people to organize themselves, to plan activities and time, and to have rules and systemic activities, even in difficult times.

The research literature on the psychometric properties of the RESI-M in different contexts and Mexican subpopulations has shown empirical evidence that it is a valid and reliable scale. In this regard, in a sample of 348 Mexican adults (235 women and 113 men), the psychometric properties of the RESI-M were evaluated, the structure was reproduced by means of principal component analysis, and 58.71% of the variance explained by the five factors was reported, the overall internal consistency was high ($\alpha = 0.92$) (32). In another study conducted by Sanjuan-Meza et al. (33) with indigenous women in Mexico, the results of the psychometric analysis of the RESI-M showed a final version of the instrument with 34 questions (out of the original 43), acceptable reliability ($\alpha = 0.942$), and six factors that explained 56.34% of the total variance (33). In another validation study of the RESI-M in patients with chronic renal failure treated with hemodialysis, after exploratory factor analysis, two of the 43 items were eliminated. The five factors explained 63.6% of the total variance, with an overall $\alpha = 0.96$, and the five factors were negatively correlated with symptoms of anxiety, depression, and distorted thoughts (34). Another study aimed to obtain the psychometric properties of the RESI-M in family caregivers of children with chronic conditions (35) and showed an adequate fit with the data based on a maximum likelihood estimator. The overall internal consistency was 0.95, and the variance explained was 63%. Likewise, in a validation study of the RESI-M in family caregivers of children with cancer, the RESI-M showed reliability and construct validity and overall internal consistency ($\alpha = 0.976$), and the explained variance was 47%. Confirmatory factor analysis showed that the five-factor model fit the data well: NFI = 0.970, CFI = 0.997, SRMR = 0.055, and RMSEA = 0.019. The RESI-M scale total score was positively correlated with psychological well-being and negatively correlated with depression, parental stress, and anxiety (27).

The findings obtained in these studies suggest that (a) the RESI-M is a multidimensional measure representing psychosocial and individual aspects of resilience; (b) the dimensions of the RESI-M remain stable; (c) the dimensions of the RESI-M are correlated, such that they would covary in the resilient behavior exhibited by the individual in situations in general; (d) the covariation of these attributes in behavior is not, however, equal among the dimensions, to the degree that some would covary more strongly than others; (e) the content of the construct of resilience appears to be unstable across studies because the number of items does not remain the same across studies (i.e., a small number, and different items need to be eliminated); and (f) the methods for studying internal structure have used an approach based on linear models.

Research regarding the impact of resilience on family caregivers is promising, but one of its limitations is having reliable measurement instruments that have been validated in this specific population. The RESI-M can be useful for this purpose and has the advantage of having been developed in the international cultural context. However,

Abbreviations: RESI-M, resilience measurement scale; INDEHUS, Instituto Nacional de Ciencias e Innovación para la Formación de Comunidad Científica.

it should be considered that this test was originally validated for use in the general population; therefore, its use in other specific populations, such as family caregivers of children with cancer, would compromise the validity and reliability of its results due to the lack of psychometric data. Although there is scientific evidence in the literature about the validity and reliability of the RESI-M in various Mexican contexts, no research results have been found that show empirical findings of the psychometric properties of RESI-M having been analyzed, evaluated and studied based on the item response theory (IRT) in a population of family caregivers of children with cancer. The IRT framework takes into account the non-linear relationship of the items with the latent attribute and the categorical expression of the items to represent the participants' responses to the measurement instrument (36). One of the main advantages of item calibration in the IRT framework is the psychometric properties provided by graded response modeling (37). In this model, the importance of each item in the measurement of the construct it is intended to measure is weighted, as opposed to classical test theory (CTT), which assumes that all items contribute equally to the measurement of the construct. Another advantage of the analysis in the IRT framework is in terms of the reliability of the instrument since the information functions allow the exploration of the accuracy of the measurements of the RESI-M factors depending on a range of values in the constructs. In contrast, CTT assumes that measurement reliability is the same at all levels of measured traits (36). Within this IRT framework, as one of the models applied to polytomous items (i.e., ordinal or Likert responses), the graded response model (GRM) has gained much acceptance because it models the variability of item discrimination and threshold spacing (36, 37), which is more realistic for most psychosocial measures.

In response to this need for reliable measurement instruments of resilience for the family caregiver population as well as to the existing knowledge gaps and to bridge the gap in this field of knowledge, the aim of the present study was to analyze the psychometric properties of the RESI-M. To this end, we formulated six objectives: (1). To evaluate the five-factor structure of the RESI-M using a full-information confirmatory and multidimensional IRT GRM; (2). To estimate the multidimensional item-level parameters of discrimination (MDISC) and difficulty (MDIFF) from the RESI-M scale; (3). To plot the item characteristic curves (ICCs) of the RESI-M; (4). To calculate the estimated precision of latent traits using the information functions of the five factors of the RESI-M; (5). To obtain measurements of the five latent factors of the RESI-M for cancer patients' caregivers; and (6). to investigate test score validity by correlating the measurements of the five latent factors with the total scores of the Beck Depression Inventory (BDI) (38), Beck Anxiety Inventory (BAI) (39), WHOQoL-BREF (40), and Zarit Burden Interview (ZBI) (41). Taking the antecedent validation studies of the RESI-M as a framework, hypothesis regarding the psychometric content were formulated. In relation to dimensionality, the hypothesis was that the number of dimensions of the RESI-M would remain at five dimensions; the second hypothesis was that the dimensions of the RESI-M would be correlated. The third hypothesis was that the items would show high levels of discrimination. Regarding relationships with external variables, a negative linear association was expected with maladaptive responses, such as anxiety symptoms, depression symptoms, and subjective burden symptoms, and a positive linear association was expected with adaptive responses, such as quality of life.

2. Materials and methods

2.1. Participants

A non-experimental, transversal, *ex post facto* study was conducted using a convenience and non-probabilistic sampling technique. A total of 633 family caregivers of hospitalized children with cancer were interviewed at the Hospital Infantil de Meìxico Federico Goìmez National Institute of Health in Mexico City. The sample included women (81.4%) and men (18.6%) aged between 18 and 52 years, with an average of 31.7 years (SD = 7.6). The inclusion criteria for the study were (1) being a family caregiver of a child who was receiving cancer treatment, (2) being at least 18 years old, and (3) having signed an informed consent form. The exclusion criteria were (1) inability to read and write and (2) refusal to participate in the study. The deletion criteria included partial or incomplete responses to the psychosocial measurement instruments. The pediatric patients included both girls (47.7%) and boys (52.3%) aged between 1 and 17 years, with an average age of 5.8 (SD = 4.9). In most cases, the time elapsed since cancer diagnosis ranged from one week to one year (68.4%), and the hospitalization period was one week to one month (85.3%).

2.2. Instruments

A battery of test instruments, including a sociodemographic variables questionnaire for research with families of children with chronic diseases and four self-report instruments measuring psychosocial variables (resilience, depression, anxiety, quality of life, and caregiver burden), were used. To guarantee the accuracy of the data obtained, the instruments were validated in the Mexican population and with families of children with chronic diseases.

2.2.1. Sociodemographic variables questionnaire (Q-SV) for research with family caregivers of children with chronic diseases

This questionnaire contains 20 items that evaluate information on sociodemographic, medical, sociocultural and family variables in families of children with chronic diseases. For this study, the diagnosis, the age and sex of the patient and caregiver, the relationship between the patient and caregiver (mother, father, or another family member), the educational level (no schooling, primary education, secondary education, undergraduate education, postgraduate education), occupation (homemaker, worker, trader, employee, student, pensioner, unemployed), marital status (married, living together, separated, divorced, single parent, widowed), years of partnership, number of children, type of family (nuclear, seminuclear, extended, single-parent), family life cycle (with young children, with school-age children, with adult children), social support networks (family, friends, religion, institutions, government), religion (Catholic, Christian, none), and monthly income were determined (12).

2.2.2. Resilience measurement scale in Mexicans (RESI-M)

This scale has been validated in family caregivers of children with cancer (35). This scale contains 43 four-point Likert-type items, ranging from 1 "strongly disagree" to 4 "strongly agree,"

and measures the level of overall resilience and five dimensions: Strength and Self-Confidence (19 items), Social Competence (eight items), Family Support (six items), Social Support (five items), and Structure (five items) (29).

2.2.3. Beck depression inventory II (BDI-II)

This inventory has been validated in a population of family caregivers of children with chronic diseases (42). This inventory includes 21 items, each with four statements that assess depressive symptomatology and episodes. It uses a rating scale from 0 to 3, where the higher the score is, the higher the level of depression. The level of depression is interpreted as follows: minimum from 1 to 4, mild from 5 to 13, moderate from 14 to 27, and severe from 28 to 63 points. Among the 330 family caregivers in the present study, the overall internal consistency of the 21 items was excellent ($\alpha = 0.90$; 95% CI = 0.89, 0.91; $\omega = 0.92$; 95% CI = 0.91, 0.94) (38).

2.2.4. Beck anxiety inventory (BAI)

This instrument has been validated in family caregivers of children with cancer by Toledano-Toledano et al. (43). With 16 items, this inventory assesses anxious symptomatology using a four-point scale, ranging from 0 "Little or nothing" to 3 "Severely." The level of anxiety obtained is minimum (1 to 5 points), mild (6 to 15), moderate (16 to 30), or severe (31 to 63). In the present sample, the overall internal consistency of the 21 items was excellent ($\alpha = 0.94$; 95% CI = 0.94, 0.95; $\omega = 0.97$; 95% CI = 0.96, 0.98) (39).

2.2.5. WHOQOL-BREF inventory of quality of life

This inventory has been validated in a Mexican population (40). It includes 26 five-point Likert-type items ranging from 1 to 5. Two items constitute general questions about quality of life, and the remaining items are grouped into the following dimensions: physical health (seven items), psychological health (six items), social relations (three items), and environment (eight items). Among the 330 family caregivers in the present study, the overall internal consistency of the 26 items was excellent ($\alpha = 0.92$) (40).

2.2.6. Zarit burden interview (ZBI)

This instrument has been validated in a Mexican population (44). It assesses the subjective burden, attitudes and emotional reactions of the caregiver when faced with the responsibility of care and the perception of the situation. It contains 22 items distributed across three factors: impact of caregiver (13 items), interpersonal relationship (six items), and self-efficacy expectations (three items). The scores of the items range from 0 "Never" to 4 "Always." In the present study, only the ZBI total score was used, and its overall internal consistency was excellent among the 330 family caregivers ($\alpha = 0.85$; 95% CI = 0.82, 0.87; $\omega = 0.98$, 95% CI = 0.93, 1.00) (41).

2.3. Procedure

The family caregivers were interviewed by the corresponding author of this study in the wards of the Hematology-Oncology Service of the Hospital Infantil de México Federico Gómez, National Institute of Health. All the family caregivers interviewed were invited to participate voluntarily; the objectives of the research were explained to them, and all of their concerns regarding the study were addressed. The family caregivers who agreed to participate signed informed consent forms and answered the instruments individually during a single session. Participants did not face any consequences for withdrawing their consent, as specified on the informed consent sheet. Before collecting the completed instruments, the interviewer checked that there were no questions without answers. If there were questions without answers, the participant was asked to respond to them, and in this way, we managed to avoid missing values.

2.4. Ethical considerations

This study is a part of the research project HIM/2015/017/SSA.1207 "Effects of mindfulness training on psychological distress and quality of life of the family caregiver," which was approved on December 16, 2014, by the Research, Ethics, and Biosafety Commissions of the Hospital Infantil de México Federico Gómez, National Institute of Health, in Mexico City. While conducting this study, the ethical rules and considerations for research with humans currently enforced in Mexico (45) and those outlined by the American Psychological Association (46) were followed. All family caregivers were informed of the objectives and scope of the research and their rights according to the Declaration of Helsinki (47). The caregivers who agreed to participate in the study signed an informed consent letter. Participation in this study was voluntary and did not involve payment.

2.5. Statistical analyses

2.5.1. Item response theory modeling

A confirmatory multidimensional IRT model was used in which five correlated factors were a priori specified to evaluate the structure and psychometric properties of the RESI-M. To evaluate their robustness in comparison with alternative measurement models, competing models were also specified: unidimensional (representing the absence of differentiated content and scores), multidimensional orthogonal (including the specific factors but restricting the correlations between them) and bifactor (representing the coexistence of a general factor and specific factors). As the scale is composed of polytomous items with ordered response categories, the GRM (37) was used, and its parameters were estimated with the Metropolis-Hastings Robbins-Monroe (MHRM) method using the "mirt" package in R (48). To facilitate model interpretation, the GRM's slopes and thresholds were re-parametrized according to Reckase (36) to obtain the multidimensional discrimination (MDISC) and difficulty (MDIFF) parameters. The goodness of fit of the models was evaluated using the M_2 * statistic and its associated RMSEA value; other fit indices were also obtained (e.g., CFI > 0.95, SRMR < 0.05). In the evaluation of the bifactor model, the extracted common variance [ECV; (49)], which indicates the degree of common variability derived from the general factor, was additionally estimated. ECV > 0.70 suggests essential unidimensionality (50). Likewise, ICCs were calculated, and the information functions of the five factors in the RESI-M scale were calculated. The ICCs allowed the investigation of the response probabilities to each category across the range in the latent trait θ, while the information functions indicated the change in the precision of the estimates in a range of $-4 \leq \theta \leq 4$. Finally, the measurements in the 5 factors of the 633 caregivers were obtained, and for the sake of validity, their linear relation with total

scores of the BDI, the BAI, the WHOQoL-BREF, and the ZBI was computed using simple linear regression controlling by sex and age of the caregiver.

2.5.2. Linear model

For comparability with previous RESI-M studies, the linear model was used to estimate the internal consistency coefficients α and ω, with confidence intervals (95%) generated by bootstrap sampling (n = 1,000 samples). This procedure was implemented by the *omega* command (51).

3. Results

3.1. Characteristics of the family caregivers

The sample included 515 women (81.4%) and 116 men (18.6%) aged between 18 and 49 years, with an average age of 31.6 (SD = 7.5). Regarding education, 2.7% of the participants had no education, 19.7% had primary school education, 44.6% had secondary school education, 25.5% had upper secondary (high school) education, and 7.4% had university or college education. The median and mode of the number of children was two, ranging from 0 to 10. More details are provided in Table 1. The pediatric patients included both girls (47.7%) and boys (52.3%), aged between 1 and 17 years, with an average age of 5.8 (SD = 4.9). In most cases, the time elapsed since cancer diagnosis ranged from one week to one year (68.4%), and the hospitalization period was one week to one month (85.4%).

TABLE 1 Summary statistics of sociodemographic variables.

Sociodemographic variable		*N*	%
Sex			
	Men	118	18.6
	Women	515	81.4
Schooling			
	No schooling	18	2.8
	Primary	124	19.6
	Secondary	282	44.5
	Higher secondary (high school)	163	25.8
	University or college	46	7.3
Occupation			
	Homemaker	413	65.2
	White-collar worker	87	13.7
	Merchant	58	9.2
	Blue-collar worker	26	4.1
	Unemployed	49	7.7
Marital status			
	Married	257	40.6
	Living together	244	38.5
	Separated	53	8.4
	Single mother	53	8.4
	Divorced	18	2.9
	Widowed	6	0.9
	Other	2	0.3
Income per month			
	<141 US dollars	390	61.6
	Between 141 and 281 US dollars	140	22.1
	Between 282 and 563 US dollars	85	13.4
	>563 US dollars	18	2.8
Religious adscription			
	Catholic Christian	512	80.9
	Non-Catholic Christian	75	11.8
	No religion	46	7.3
		M	SD
Age (years)		31.7	7.58
Number of children		2.32	1.17

n, frequency; %, percentage; mean, arithmetic mean; SD, standard deviation.

3.2. Model results

3.2.1. Internal structure and model fit

From all competing models, the multidimensional IRT model with correlated factors and the bifactor model obtained the best goodness of fit indices (Table 2). The bifactor model yielded lower RMSEA and higher TLI and CFI values than the multidimensional IRT model with correlated factors; however, the ECV derived from the primary factor was 0.62, which weakens the conclusion that a bifactor structure underlies the RESI-M (50) and was the reason why we decided to report on the functioning of the multidimensional IRT model with correlated factors. Even though this confirmatory model had a statistically significant value of $M_2 * (774) = 3714.12$, $p < .001$, the RMSEA suggested an acceptable fit, with $RMSEA = 0.078$ (95% CI: 0.075, 0.080), as did the TLI and CFI statistics, which were 0.90 and 0.91, respectively.

3.2.2. Multidimensional item parameters

The multidimensional parameters (MDISC and MDIFF) of the RESI-M obtained in the sample of caregivers are included in Table 3. In IRT, the a parameter corresponds to the slope of the function, in this case, MDISC (36), which allows individuals with low or high levels of the latent trait to be distinguished. Likewise, the parameters b_1 , b_2, and b_3 that correspond to the thresholds are presented as measurements of MDIFF (36), which indicates how much of the latent trait is required for a respondent to endorse a particular category. Items with a greater $_a$ value have better discrimination (i.e., they have a stronger relationship with the latent construct), and response categories with a larger b value indicate that the caregiver must have a high level of *resilience* to select that category. The range of a values was from 1.40 for item 2 of the Strength and Self-Confidence factor to 4.88 for item 35 of the Social Support factor; therefore, according to the classification proposed by Baker (52), 19% of the items had "high" discrimination, while the majority (81%) had "very high" discrimination.

TABLE 2 Goodness-of fit-indices and information criteria from all competing models.

Model	ML	df	*p*	RMSEA	TLI	CFI	AIC	BIC	loglik
Unidimensional	7199.27	774	<0.001	0.12	0.78	0.80	43,388	44,154	−21,522
MD (orthogonal)	3718.36	774	<0.001	0.08	0.90	0.91	41,084	41,849	−20,370
MD (correlated factors)	3714.12	774	<0.001	0.08	0.90	0.91	39,816	40,626	−19,726
Bifactor	2587.33	731	<0.001	0.06	0.93	0.94	39,529	40,486	−19,549

MD, multidimensional; ML, likelihood-ratio-chi-2 test statistics; df, degree of freedom; *p*, probability value; AIC, Akaike information criterion; RMSEA, root mean square error of approximation; TLI, Tucker–Lewis index; BIC, Bayesian information criterion; loglik, log likelihood.

3.2.3. Information functions

Figure 1 shows the ICCs of the items with the highest discrimination of each of the five factors. Each panel includes the probability of selecting the response categories depending on a range of $-4 \leq \theta \leq 4$ in the latent trait. The ICCs reveal the GRM response predictions across different levels of the Strength and Self-Confidence (SSC), Social Competence (SC), Family Support (FS), Social Support (SS), and Structure (Str) factors. As the scores in the latent trait are standardized, the average of the scale occurs when $\theta = 0$; at this level of the traits, it is possible to observe that the most likely response to these items is the category "Agree." Levels above the average of the latent traits are required to select the highest response category, and levels below $\theta < -1$ are required to select the lowest categories.

3.2.4. Score reliability

Additionally, the test information functions (TIFs) for the five RESI-M factors are shown in Figure 2. The TIFs allow the test precision to be explored to measure different levels of the traits. At the levels of θ where the function increases, we found the most precise measures; this is also where the test can collect the most information from the latent traits. For example, the Strength and Self-Confidence factor TIF provided information in a wide range of θ values; however, the information was substantially higher for values lower than the average when $\theta \approx -2$. Additionally, it could be observed that the function had another maximum at levels above the average ($\theta \approx 1$), and that pattern was present in all factor functions.

In the linear modeling, α coefficients for the RESI-M scores were as follows: *Strength and Self-Confidence* (α = 0.93, 95% CI = 0.92,.94, se = 0.004), Social Competence (α = 0.86, 95% CI = 0.83, 0.88, se = 0.011), Family Support (α = 0.88, 95% CI = 0.87,.90, se = 0.009), Social Support (α = 0.91, 95% CI = 0.89, 0.93, se = 0.008), and Structure (α = 0.75, 95% CI = 0.71, 0.79, se = 0.019). Additionally, the ω coefficients for the factors were as follows: Strength and Self Confidence (ω = 0.95, 95% CI = 0.94, 0.96), Social Competence (ω = 0.88, 95% CI = 0.86, 0.90), Family Support (ω = 0.94, 95% CI = 0.91, 0.96), Social Support (ω = 0.92, 95% CI = 0.90, 0.94), and Structure (ω = 0.78, 95% CI = 0.74, 0.81).

3.2.5. Factors' individual scores

The estimation of the factors' individual scores, their dispersion, and Pearson's correlation coefficient are shown in Figure 3. The distributions of the standardized scores of the factors are included in the figure diagonal. In the lower part of the matrix, the figures depict the position of each caregiver in two dimensions as points of the scatter plots, while in the upper part of the matrix, the Pearson correlation coefficients are reported, which were positive and statistically significant ($p < .001$). A slight positive bias can be observed in the distributions of the Strength and Self-Confidence (SSC), Social Competence (SC), and Structure (Str) factors, as well as a slight negative bias for the Family Support (FS), and Social Support (SS) factors. Therefore, it can be inferred that the majority of the caregivers had low scores on the Strength and Self-Confidence, Social Competence, and Structure factors, while the Family Support and Social Support scores of the majority of the caregivers were high. Regarding the correlations of the factors, positive and strong associations ($r > .7$) were detected for the relationships between the Strength and Self-Confidence and Social Competence factors, Strength and Self-Confidence and Family Support factors, and Structure and Social Competence factors.

3.2.6. Validity

Finally, regarding the validity of the measures of latent traits in primary caregivers, the linear relation between the five factors of the RESI-M and the total scores of the BDI, BAI, WHOQoL-BREF, and ZBI provided evidence of convergent and discriminant validity. Table 4 includes the standardized slope parameters matrix (simple linear regression coefficients) between all the RESI-M factors and the total scores of the aforementioned scales controlling by sex and age of the caregiver. In general, it can be noted that the scores of the resilience factors predict negative relations with depression (BDI), anxiety (BAI), and caregiver burden (ZBI) and a positive association with the quality-of-life scale (WHOQoL-BREF). Although the strength of the relationships varied from weak to moderate estimates, all of them were statistically significant at $p < .05$.

4. Discussion

The aim of the study was to evaluate the construct validity of the RESI-M, focusing on the internal structure, the reliability of the scores, and the relationship with external constructs. In contrast to previous studies of the RESI-M (27, 35), the present study used a full-information confirmatory multidimensional IRT GRM, a model that allows parameters of the metric structure of the instruments to be obtained in a non-linear framework and that is more detailed at the item level and at the score level. The results of our evaluation support the multidimensional structure of the RESI-M. We confirmed the five dimensions of the scale proposed in previous evaluations conducted with a linear analysis framework. In terms of item functioning, all 43 items of the RESI-M were informative (i.e., the degree to which they contain information about the construct measured) and contributed specifically to assessing different aspects of resilience. Adequate item functioning comprised five latent dimensions that accurately measure the factors Strength and Self-Confidence (SSC), Social Competence (SC), Family Support (FS), Social Support (SS), and Structure (Str), ranging from minus three to two standard deviations below the mean to 1 to 2 standard deviations above the mean.

TABLE 3 Multidimensional parameters of discrimination and difficulty from the full information confirmatory graded response model (GRM).

Factor and item		a	b_1	b_2	b_3
Strength and Self Confidence (SSC)					
1	What has happened to me in the past makes me feel confident. . .	1.50	−2.77	−1.62	0.69
2	I know where to look for help.	1.41	−3.00	−1.64	0.98
3	I am a strong person.	1.94	−3.31	−1.67	0.64
4	I know very well what I want.	2.25	−3.14	−1.48	0.62
5	I have control over my life.	1.57	−3.01	−1.37	1.14
6	I like challenges.	1.51	−2.84	−1.26	1.05
7	I strive to reach my goals.	3.29	−2.73	−1.89	0.34
8	I am proud of my achievements.	2.96	−2.67	−1.56	0.39
9	I know I have skills.	3.24	−2.61	−1.97	0.25
10	Believing in myself helps me overcome difficult moments.	2.27	−2.99	−1.89	0.20
11	I think I will succeed.	2.37	−2.76	−1.68	0.38
12	I know how to achieve my goals.	3.03	−2.98	−1.28	0.73
13	Whatever happens, I will always find a solution.	2.08	−3.14	−2.08	0.39
14	My future looks good.	2.53	−2.51	−1.01	0.90
15	I know that I can solve my personal problems.	2.78	−3.35	−1.96	0.53
16	I am satisfied with myself.	2.86	−2.59	−1.46	0.62
17	I have realistic plans for the future.	1.96	−2.86	−1.62	0.71
18	I trust my decisions.	2.70	−3.36	−1.53	0.62
19	When I am not well, I know that better times will come.	1.62	−3.32	−2.39	0.44
Social Competence (SC)					
20	I feel comfortable with other people.	1.60	−2.79	−1.19	1.32
21	It is easy for me to establish contact with new people.	1.96	−2.32	−0.93	1.28
22	It is easy for me to make new friends.	2.20	−2.15	−0.77	1.18
23	It is easy for me to think of good topics of conversation.	2.96	−2.32	−0.80	0.99
24	I adapt easily to new situations.	2.20	−2.33	−0.93	1.06
25	It is easy for me to make other people laugh.	1.67	−2.80	−0.57	1.66
26	I enjoy being with other people.	1.84	−3.10	−1.29	1.35
27	I know how to start a conversation.	2.52	−2.50	−0.92	1.20
Family Support (FS)					
28	I have a good relationship with my family.	3.20	−2.27	−1.62	0.14
29	I enjoy being with my family.	3.56	−2.87	−1.86	−0.19
30	In our family, we are loyal to each other.	4.48	−2.13	−1.48	0.18
31	In our family, we enjoy doing activities together.	4.81	−2.09	−1.35	0.07
32	Even in difficult times, our family has an optimistic attitude. . .	1.94	−2.45	−1.86	0.46
33	In our family we agree in relation to what we consider. . .	1.90	−3.34	−1.99	0.53
Social Support (SS)					
34	I have some friends/relatives who truly care about me.	3.99	−1.95	−1.43	0.30
35	I have some friends/relatives who support me.	4.88	−1.86	−1.27	0.34
36	I always have someone who can help me when I need it.	3.15	−1.97	−1.37	0.28
37	I have some friends/relatives who encourage me.	4.35	−2.02	−1.36	0.30
38	I have some friends/relatives who value my skills.	3.14	−2.12	−1.37	0.64

(Continued)

TABLE 3 (Continued)

Factor and item		a	b_1	b_2	b_3
Structure (Str)					
39	Rules and routine make my life easier.	1.83	−2.42	−0.89	1.45
40	I keep my routine even in difficult times.	1.59	−2.51	−0.62	1.74
41	I prefer to plan my activities.	1.94	−2.39	−1.11	1.41
42	I work better when I have goals.	1.81	−3.07	−1.68	0.94
43	I am good at organizing my time.	2.30	−2.11	−0.81	1.16

a: slope parameter (discrimination). b_i: thresholds parameter.

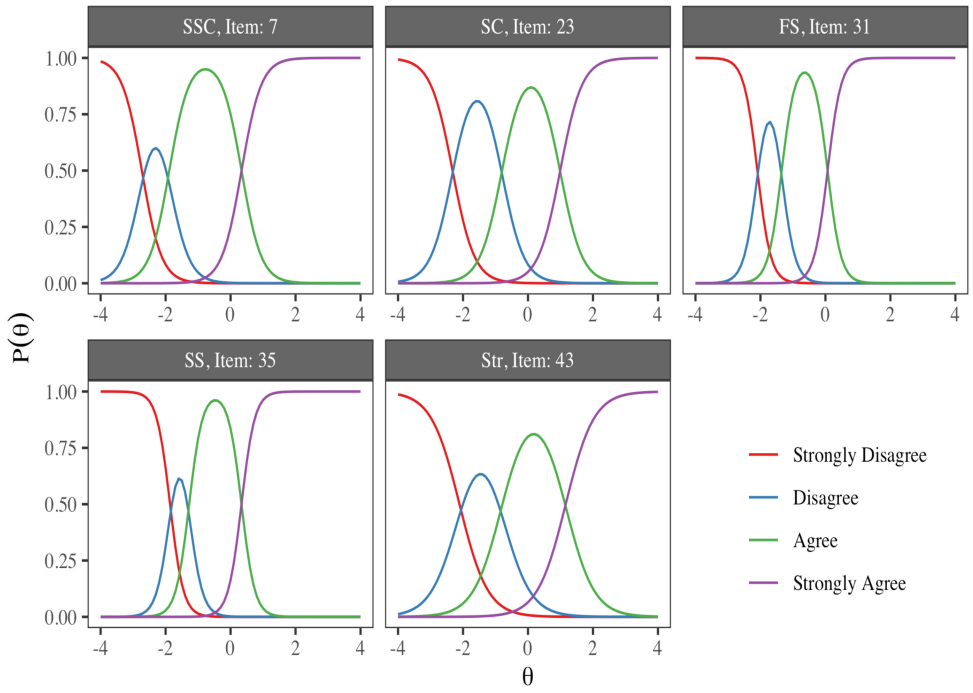


FIGURE 1
Item characteristic curves with the greatest discrimination from each factor of the resilience measurement scale (RESI-M) scale. The ICCs reveal the GRM response predictions at different levels of the Strength and Self-Confidence (SSC), Social Competence (SC), Family Support (FS), Social Support (SS), and Structure (Str) factors.

The MDISC and MDIFF parameters of the RESI-M items were different for each item, which supports the idea that all items contribute differentially to the measurement of the dimensions of resilience. This is not a problem for the measurement of the attribute because it is a realistic expression of the differential content of the items and the conceptual structure of the construct being measured. All items had high or very high discrimination, which indicates that they have the ability to distinguish with high accuracy between individuals who have low or high levels of the dimensions assessed; in the context of caregivers of cancer patients, this could be very useful to detect who would need specific psychological intervention and describe them with high accuracy in the RESI-M framework.

Regarding the precision of the estimates of the latent traits by means of the information functions of the five dimensions of the RESI-M, IRT modeling made it possible to detect within the RESI-M which of the factors and at which levels of the traits there was more measurement precision and therefore more reliability. The findings obtained indicate that the measurement precision had a bimodal form, in that the further away from the mean the subject's position is, the higher the precision will be. This bimodal form of the information function suggests that the construct is sensitive to individual differences at the extremes of the construct but does not appear to be recommended for scores near the mean because of the greater measurement error at this level of the score. This seems unusual; however, it may be reasonable in the measurement of resilience, given that this construct emerges or is clearly observable when the subject is exposed to adverse factors, and the expression of resilient behavior may show consistency in these extreme situations. A practical implication is that because all factors had a decrease in informativeness around the mean, if one wanted to improve the scale in terms of greater coverage in the range of latent scores, creating items that are informative at average levels of resilience would be appropriate. One practical implication for the use of the instrument is that the description of the resilience attribute may be less appropriate for groups at the middle level of the RESI-M and more accurate and consistent at both ends of the construct. Overall, in the future, practitioners using the RESI-M in caregivers of oncology patients could reliably determine whether the caregiver has high or

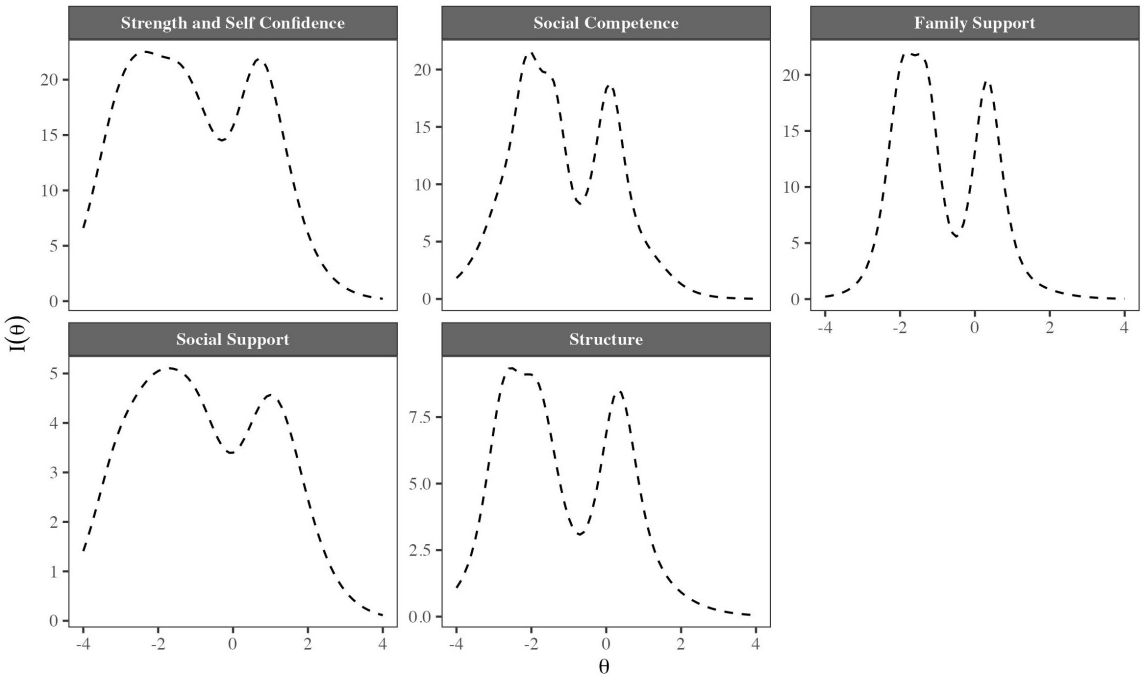


FIGURE 2
Test information functions (TIFs) for the five factors of the RESI-M.

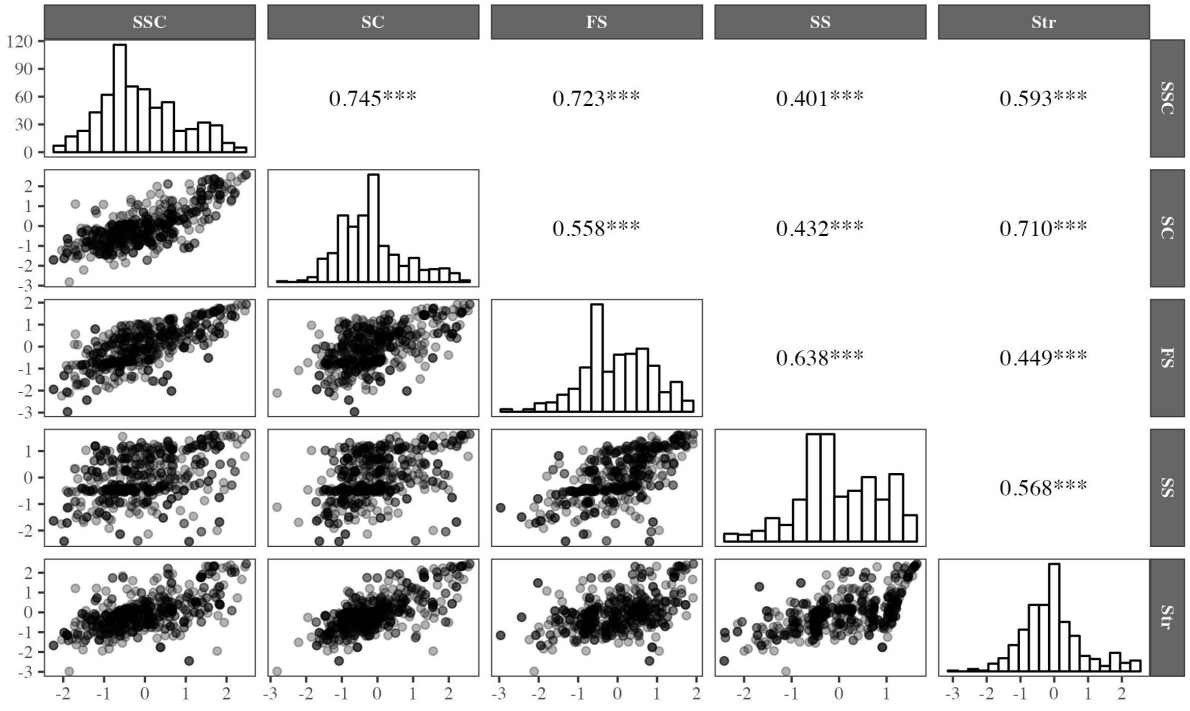


FIGURE 3
Estimation of individual factor scores, their dispersion and Pearson's correlation coefficient. SSC, Strength and Self-Confidence; SC, Social Competence; FS, Family Support; SS, Social Support; Str, Structure. ***$p < 0.001$.

low levels of the dimensions assessed without additional analysis; this is consistent with what has been reported in previous research about IRT utility, level of reporting, and test-retest reliability (53).

Regarding the convergent and divergent validity of the RESI-M with the total scores of the BDI, BAI, WHOQoL-BREF, and ZBI, the latent scores in the factors of the instrument correlated with the scores of scales to measure depression (BDI), anxiety (BAI), quality of life (WHOQoL-BREF) and caregiver overload (ZBI); therefore, the hypotheses of association with variables were satisfactorily fulfilled. An important finding of the present study is that regardless of the

TABLE 4 Standardized linear regression slopes to evaluate the validity of factor score measures in primary caregivers.

	Linear correlations				Descriptive information			
Variable	BDI	WHOQoL	BAI	ZBI	M	SD	Sk	K
SSC	−0.44	0.54	−0.23	−0.23	6.19	8.36	0.15	−0.47
SC	−0.34	0.46	−0.18	−0.20	22.98	3.97	0.36	0.13
FS	−0.45	0.53	−0.31	−0.23	19.91	3.25	−0.84	1.22
SS	−0.17	0.39	−0.13	−0.22	16.12	3.09	−0.84	1.39
Str	−0.24	0.34	−0.10	−0.17	14.43	2.44	0.19	0.78

SSC, Strength and Self Confidence; SC, Social Competence; FS, Family Support; SS, Social Support; Str, Structure; Sk, Fisher's skew coefficient; K, Fisher's excess kurtosis. Slopes were obtained controlling by sex and age of the caregiver. All p-values were statistically significant at $p < 0.05$.

number of items contained in each factor, the factors correlated congruently and statistically significantly with the scales. Therefore, the strongest correlations with the scores of the instruments were those of the Strength and Self-Confidence factor, which is the factor with the largest number of items; even the Structure factor, with only five items, correlated congruently and statistically significantly with the scales mentioned; therefore, we can conclude the validity of the estimates of the constructs that we obtained with IRT. The results of the correlations coincide with a previous study (27, 35) that evaluated the relationship between scores of the RESI-M factors, obtained with confirmatory factor analysis (CFA), and the total scores of depression and anxiety. In the present study, we also detected a negative association of RESI-M factors with depression and anxiety scores; however, we extended those findings by obtaining a positive correlation with the WHOQoL-BREF quality of life scores and a negative correlation with the ZBI scores. The theoretical congruence of the correlations and the correspondence with previous findings provide evidence for the validity of the RESI-M and its factor estimates in the IRT framework.

A limitation of the study is that objective measures of health were not used; therefore, future studies would benefit from establishing a relationship with measurements other than self-report, such as physiological measures or behavioral records. Another limitation refers to the non-probability sampling and the sample size of less than 1,000, which indicate that the estimates should be taken with due caution (although some simulation studies suggest that 500 participants may be adequate (54)), even within the population from which the sample was drawn (family caregivers of children undergoing cancer treatment at the Hospital Infantil de México Federico Gómez, National Institute of Health, in Mexico City).

5. Conclusion

The original five-dimensional structure of the RESI-M was confirmed. As a contrasting strategy, alternative structures were tested, specifically unidimensional and bifactor (one general dimension and five specific dimensions), but they were not strong enough to justify the use of a general score and interpret it theoretically. However, there are items with potential psychometric strength to create a possible general dimension, and future studies may confirm this psychometric property. The items of the subscales in general are shown to be representative of the measured dimensions and to contribute to the robust interpretation of their dimensions. The accuracy of the scores is high at the extremes, i.e., when the respondent scores below or above the mean. The overall reliability of the scores tends to be acceptable for group description and applied research purposes. Finally, the RESI-M scores show convergent validity in relation to the emotional responses of depression, anxiety and burden, as well as perceived quality of life.

Finally, we provide some suggestions for future lines of research. Due to the length of the instrument and imbalance in the content presentation of the subscales (number of items in each subscale), the moderate overall factor strength and the size of the interfactor correlations, an abbreviated version of the instrument could be developed. At the same time, in the present study, reliability by stability was not estimated, so it is suggested to estimate reliability at least at two different time points. Using a short-term and a long-term interval, the stability and dependability of the scores can be evaluated (55). Both aspects are conceptually different and provide different facets of score stability. The invariance or equivalence between groups was also not contrasted since the eligible samples were unbalanced; therefore, its evaluation (sex of the parents, sex of the oncology patient, etc.) is indicated from a non-proportional stratified sampling (with equiprobable or balanced strata). This type of contrast will help to establish the invariance of the estimated psychometric parameters or to describe differences.

Data availability statement

The original contributions presented in this study are included in this article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

This study is a part of the research project HIM/2015/017/SSA.1207 "Effects of mindfulness training on psychological distress and quality of life of the family caregiver", which was approved by the Research, Ethics, and Biosafety Commissions of the Hospital Infantil de México Federico Gómez, National Institute of Health, in Mexico City. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JM, SJ, and FT-T: conceptualization. SJ, CM-S, and FT-T: methodology and writing—original draft preparation. SJ: software

and formal analysis. RV-G and SJ: validation. CM-S and FT-T: research. FT-T and RV-G: resources. SJ and FT-T: data curation and project administration. FT-T, CM-S, and JM: writing—review and editing. SJ and CM-S: visualization. FT-T and JM: supervision and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding

This work presents some results of the research project HIM/2015/017/SSA.1207 "Effects of mindfulness training on psychological distress and quality of life of the family caregiver", Main researcher: FT-T. The present research was funded with federal funds for health research and was approved by the Commissions of Research, Ethics and Biosafety (Comisiones de Investigación, Ética y Bioseguridad), Hospital Infantil de México Federico Gómez, National Institute of Health. The funding agency had no control over the design of the study; the collection, analysis and interpretation of the data; or the writing of the manuscript.

Acknowledgments

We thank Angel Gabriel Uribe Zamorano for his support from him.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Toledano-Toledano F, Luna D, Moral De La Rubia J, Martínez Valverde S, Bermúdez Morón C, García M, et al. Psychosocial factors predicting resilience in family caregivers of children with cancer: a cross-sectional study. *Int J Environ Res Public Health.* (2021) 18:748. doi: 10.3390/ijerph18020748

2. Teixeira R, Remondes-Costa S, Graça Pereira M, Brandão T. The impact of informal cancer caregiving: a literature review on psychophysiological studies. *Eur J Cancer Care.* (2019) 28:e13042. doi: 10.1111/ecc.13042

3. Toledano-Toledano F, Moral De La Rubia J, McCubbin L, Cauley B, Luna D. Brief version of the coping health inventory for parents (CHIP) among family caregivers of children with chronic diseases. *Health Qual Life Outcomes.* (2020) 18:104. doi: 10.1186/s12955-020-01357-5

4. Toledano-Toledano F, Domínguez-Guedea M. Psychosocial factors related with caregiver burden among families of children with chronic conditions. *Biopsychosoc Med.* (2019) 13:6. doi: 10.1186/s13030-019-0147-2

5. Haegen MV, Luminet O. Stress, psychosocial mediators, and cognitive mediators in parents of child cancer patients and cancer survivors: attention and working memory pathway perspectives. *J Psychosoc Oncol.* (2015) 33:504–50. doi: 10.1080/07347332.2015.1067279

6. Cernvall M, Hovén E, Ljungman L, Ljungman G, Carlbring P, Von Essen L. Posttraumatic stress and attentional bias towards cancer-related stimuli in parents of children recently diagnosed with cancer. *PLoS One.* (2016) 11:e0152778. doi: 10.1371/journal.pone.0152778

7. Toledano-Toledano F, Moral de la Rubia J. Factors associated with anxiety in family caregivers of children with chronic diseases. *Biopsychosoc Med.* (2018) 12:20. doi: 10.1186/s13030-018-0139-7

8. Young C, Craig J, Clapham K, Williams S, Williamson A. Stressful life events and resilience among carers of aboriginal children in urban New South Wales: cross-sectional findings from the study of environment on aboriginal resilience and child health (SEARCH). *BMJ Open.* (2018) 8:e021687. doi: 10.1136/bmjopen-2018-021687

9. Khanna A, Prabhakaran A, Patel P, Ganjiwale J, Nimbalkar S. Social, psychological and financial burden on caregivers of children with chronic illness: a cross-sectional study. *Indian J Pediatr.* (2015) 82:1006–11. doi: 10.1007/s12098-015-1762-y

10. Toledano-Toledano F, de la Rubia J, Nabors L, Domínguez-Guedea M, Escudero G, Pérez E, et al. Predictors of quality of life among parents of children with chronic diseases: a cross-sectional study. *Healthcare.* (2020) 8:456. doi: 10.3390/healthcare8040456

11. Toledano-Toledano F, Luna D. The psychosocial profile of family caregivers of children with chronic diseases: a cross-sectional study. *Biopsychosoc Med.* (2020) 14:29. doi: 10.1186/s13030-020-00201-y

12. Toledano-Toledano F, Rodríguez-Rey R, Moral De La Rubia J, Luna D. A Sociodemographic variables questionnaire (Q-SV) for research on family caregivers of children with chronic disease. *BMC Psychol.* (2019) 7:85. doi: 10.1186/s40359-019-0350-8

13. Iranmanesh S, Shamsi A, Dehghan M. Post-Traumatic stress symptoms among iranian parents of children during cancer treatment. *Issues Ment Health Nurs.* (2015) 36:279–85. doi: 10.3109/01612840.2014.983622

14. Creswell P, Wisk L, Litzelman K, Allchin A, Witt W. Parental depressive symptoms and childhood cancer: the importance of financial difficulties. *Support Care Cancer.* (2014) 22:503–11. doi: 10.1007/s00520-013-2003-4

15. Gardner M, Mrug S, Schwebel D, Phipps S, Whelan K, Madan-Swain A. Demographic, medical, and psychosocial predictors of benefit finding among caregivers of childhood cancer survivors. *Psychooncology.* (2017) 26:125–32. doi: 10.1002/pon.4014

16. Toledano-Toledano F, de la Rubia J, Frometa R, Betanzos F, Guzmán L, García M. The social support networks scale (SSNS) for family caregivers of children with cancer: a psychometric evaluation. *Int J Environ Res Public Health.* (2020) 17:7820. doi: 10.3390/ijerph17217820

17. Granek L, Rosenberg-Yunger Z, Dix D, Klaassen R, Sung L, Cairney J, et al. Caregiving, single parents and cumulative stresses when caring for a child with cancer. *Child Care Health Dev.* (2014) 40:184–94. doi: 10.1111/cch.12008

18. Rubira E, Marcon S, Belasco A, Gaíva M, Espinosa M. Burden and quality of life of caregivers of children and adolescents with cancer under chemotherapy treatment. *ACTA Paul Enferm.* (2012) 25:567–73. doi: 10.1590/S0103-21002012005000020

19. Iranmanesh S, Shamsi A. The relationship between type of cancer and parent's psychosocial risks. *Asian J Nurs Educ Res.* (2014) 4:495–501.

20. Okado Y, Tillery R, Howard Sharp K, Long A, Phipps S. Effects of time since diagnosis on the association between parent and child distress in families with pediatric cancer. *Child Heal Care.* (2016) 45:303–22. doi: 10.1080/02739615.2014.996883

21. Juth V, Silver R, Sender L. The shared experience of adolescent and young adult cancer patients and their caregivers. *Psychooncology.* (2015) 24:1746–53. doi: 10.1002/pon.3785

22. Lakkis N, Khoury J, Mahmassani D, Ramia M, Hamadeh G. Psychological distress and coping strategies in parents of children with cancer in Lebanon. *Psychooncology.* (2016) 25:428–34. doi: 10.1002/pon.3934

23. Chen J, Liu Y, Cai Q, Liu Y, Wang T, Wang J, et al. Depression in parents of children with leukemia in southern China accompanied by the prevalence of type D personality. *Support Care Cancer.* (2014) 22:1277–86. doi: 10.1007/s00520-013-2082-2

24. Kim Y, Shaffer K, Carver C, Cannady R. Quality of life of family caregivers 8 years after a relative's cancer diagnosis: follow-up of the National Quality of Life Survey for Caregivers. *Psychooncology.* (2016) 25:266–74. doi: 10.1002/pon.3843

25. Teixeira R, Pereira M. Psychological morbidity, burden, and the mediating effect of social support in adult children caregivers of oncological patients undergoing chemotherapy. *Psychooncology.* (2013) 22:1587–93. doi: 10.1002/pon.3173

26. Hermanns M, Mastel-Smith B. Caregiving: a qualitative concept analysis. *Qual Rep.* (2012) 17:1–18. doi: 10.46743/2160-3715/2012.1727

27. Toledano-Toledano F, Moral De La Rubia J, Broche-Pérez Y, Domínguez-Guedea M, Granados-García V. The measurement scale of resilience among family caregivers of children with cancer: a psychometric evaluation. *BMC Public Health.* (2019) 19:1164. doi: 10.1186/s12889-019-7512-8

28. Windle G, Bennett K, Noyes J. A methodological review of resilience measurement scales. *Health Qual Life Outcomes.* (2011) 9:8. doi: 10.1186/1477-7525-9-8

29. Lever J, Valdez N. Desarrollo de una escala de medición de la resiliencia con mexicanos (RESI-M). [Construction of a measurement scale of resilience in Mexicans (RESI-M)]. *Interdiscipl Rev Psicol Ciencias Afines.* (2010) 27:7–22.

30. Connor K, Davidson J. Development of a new resilience scale: the connor-davidson resilience scale (CD-RISC). *Depress Anxiety.* (2003) 18:76–82. doi: 10.1002/da.10113

31. Friborg O, Hjemdal O, Rosenvinge J, Martinussen M. A new rating scale for adult resilience: what are the central protective resources behind healthy adjustment? *Int J Methods Psychiatr Res.* (2003) 12:65–76. doi: 10.1002/mpr.143

32. Camacho-Valadez D. Propiedades psicométricas de la escala de resiliencia mexicana en población del norte De México. *Enseñanza Investig En Psicol.* (2016) 21:78–83.

33. Sanjuan-Meza X, Landeros-Olvera E, Cossío-Torres P. Validity of a resilience scale (RESI-M) in indigenous women in Mexico. *Cad Saude Publica.* (2018) 34:e00179717. doi: 10.1590/0102-311x00179717

34. Lerma A, Ordónez G, Mendoza L, Salazar-Robles E, Rivero J, Pérez-Granados E, et al. Psychometric properties of the resilience scale in Mexican patients with chronic hemodialysis. *Salud Ment.* (2019) 42:121–9. doi: 10.17711/SM.0185-3325.2019.016

35. Toledano-Toledano F, Moral de la Rubia J, McCubbin L, Liebenberg L, Vera Jiménez J, Rivera-Rivera L, et al. Validity and reliability of the Mexican resilience measurement scale in families of children with chronic conditions. *Health Qual Life Outcomes.* (2017) 15:242. doi: 10.1186/s12955-017-0817-3

36. Reckase M. *Multidimensional item response theory*. New York, NY: Springer (2009). doi: 10.1007/978-0-387-89976-3

37. Samejima F. Graded response models. In: van der Linden WJ, Hambleton RK editors. *Handbook of modern item response theory*. New York, NY: Springer (2016). doi: 10.1201/9781315374512

38. Beck A, Steer R, Brown G. *Manual for the beck depression inventory-II*. San Antonio, TX: Psychol Corp (1996).

39. Beck A, Epstein N, Brown G, Steer R. An Inventory for Measuring Clinical Anxiety: psychometric Properties. *J Consult Clin Psychol.* (1988) 56:893–7. doi: 10.1037/0022-006X.56.6.893

40. González-Celis R, Sánchez-Sosa J. Efectos de un programa cognitivo-conductual para mejorar la calidad de vida en adultos mayores. *Rev Mex Psicol.* (2003) 20:43–58.

41. Zarit S, Reever K, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist.* (1980) 20:649–55. doi: 10.1093/geront/20.6.649

42. Toledano-Toledano F, Contreras-Valdez J. Validity and reliability of the beck depression inventory II (BDI-II) in family caregivers of children with chronic diseases. *PLoS One.* (2018) 13:e0206917. doi: 10.1371/journal.pone.0206917

43. Toledano-Toledano F, de la Rubia J, Domínguez-Guedea M, Nabors L, Barcelata-Eguiarte B, Rocha-Pérez E, et al. Validity and reliability of the beck anxiety inventory (BAI) for family caregivers of children with cancer. *Int J Environ Res Public Health.* (2020) 17:7765. doi: 10.3390/ijerph17217765

44. Flores-Terrones M, Galindo-Vázquez Ó, Jiménez-Genchi J, Rivera-Fong L, González-Rodríguez E. Validación de la entrevista de carga de zarit en cuidadores primarios informales de pacientes con diagnóstico de enfermedades mentales. *Psicol Salud.* (2018) 29:17–24. doi: 10.25009/pys.v29i1.2564

45. Sociedad Mexicana de Psicología. *Código Etico del Psicólogo*. Ciudad de México: Trillas (2017).

46. American Psychological Association. Ethical principles of psychologists and code of conduct. *Am Psychol.* (2002) 57:1060–73. doi: 10.1037/0003-066X.57.12.1060

47. World Medical Association. World medical association declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA.* (2013) 310:2191–4. doi: 10.1001/jama.2013.281053

48. Chalmers R. Mirt: a multidimensional item response theory package for the R environment. *J Stat Softw.* (2012) 48:1–29. doi: 10.18637/jss.v048.i06

49. Dueber D, Toland M. A bifactor approach to subscore assessment. *Psychol Methods.* (2021). [Epub ahead of print]. doi: 10.1037/met0000459

50. Rodriguez A, Reise S, Haviland M. Evaluating bifactor models: calculating and interpreting statistical indices. *Psychol Methods.* (2016) 21:137–50. doi: 10.1037/met0000045

51. Hayes A, Coutts J. Use omega rather than cronbach's alpha for estimating reliability. But… *Commun Methods Meas.* (2020) 14:1–24. doi: 10.1080/19312458.2020.1718629

52. Baker F. *The basics of item response theory. second edition. Ericedgov*. (2001). Available online at: http://eric.ed.gov/ERICWebPortal/recordDetail?accno=ED458219%5Cnpapers2://publication/uuid/53C840DD-C92B-4719-8EC3-AF2076EDCAB3 (accessed November 15, 2022).

53. Manapat P, Edwards M, MacKinnon D, Poldrack R, Marsch LAA. Psychometric analysis of the brief self-control scale. *Assessment.* (2021) 28:395–412. doi: 10.1177/1073191119890021

54. Jiang S, Wang C, Weiss D. Sample size requirements for estimation of item parameters in the multidimensional graded response model. *Front Psychol.* (2016) 7:109. doi: 10.3389/fpsyg.2016.00109

55. Revelle W, Condon D. Reliability from α to ω: a tutorial. *Psychol Assess.* (2019) 31:1395–411. doi: 10.1037/pas0000754

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 04 April 2023
DOI 10.3389/fpsyt.2023.928727

OPEN ACCESS

EDITED BY
Michael Noll-Hussong,
Saarland University,
Germany

REVIEWED BY
Wangjun Qin,
China-Japan Friendship Hospital,
China
Yueli Zhang,
Zhengzhou Central Hospital,
China

*CORRESPONDENCE
Jinju Duan
✉ duanjinju@163.com

SPECIALTY SECTION
This article was submitted to Anxiety and Stress Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 26 April 2022
ACCEPTED 07 March 2023
PUBLISHED 04 April 2023

CITATION
Wang S, Wang X, Liu X, Zhao C and Duan J (2023) Moderating effects of humanistic care and socioeconomic status on the relationship among pain intensity, psychological factors, and psychological function in adults with cancer pain from a province of China: A cross-sectional study. *Front. Psychiatry* 14:928727. doi: 10.3389/fpsyt.2023.928727

Moderating effects of humanistic care and socioeconomic status on the relationship among pain intensity, psychological factors, and psychological function in adults with cancer pain from a province of China: A cross-sectional study

Shuyun Wang[1], Xuyan Wang[2], Xiaohong Liu[3], Chenxing Zhao[4] and Jinju Duan[1]*

[1]Department of Pharmacy, Second Hospital of Shanxi Medical University, Taiyuan, Shanxi, China, [2]School of Traditional Chinese Medicine and Food Engineering, Shanxi University of Traditional Chinese Medicine, Taiyuan, Shanxi, China, [3]Department of Pharmacy, Yangquan First People's Hospital, Yangquan, Shanxi, China, [4]Department of Pharmacy, Linfen People's Hospital, Linfen, Shanxi, China

Objective: The objective of this study is to explore whether humanistic care practiced by clinical pharmacists and socioeconomic status moderate the associations among pain intensity, psychological factors (catastrophizing and resilience), and psychological function (depression and anxiety) in cancer patients with low levels of education and income in the Shanxi province in the Northwest of China.

Methods: Our sample comprised 123 adult inpatients with cancer pain. Demographic variables were obtained from the Hospital Information System of The Second Hospital of Shanxi Medical University. Pain intensity, psychological factors, and psychological functions were evaluated with four scales, and humanistic care was practiced with a part of the patients by clinical pharmacists. First, univariate analyses were conducted, followed by moderating effect models.

Results: The incidence of depression and anxiety in patients with cancer pain in our sample were 48.78 and 41.46%, respectively. Low levels of psychological resilience (63.37, SD 21.74) were in this study. Pain intensity was significantly associated with humanistic care and anxiety. Humanistic care practiced by clinical pharmacists moderated not only the association between resilience and pain intensity but also the association between pain intensity and anxiety. Education levels moderated the relationship between pain intensity and the psychological factors of catastrophizing and resilience. Income levels moderated the association between resilience and anxiety.

Conclusion: Humanistic care is essential in moderating the association among pain intensity, psychological factors, and psychological functions in Chinese cancer patients, especially those from lower-level counties and rural areas. Furthermore, socioeconomic statuses, such as education level and income, cannot easily change quickly. Still, proper humanistic care can relieve pain

more effectively, reminding us that medical staff should implement effective personalized interventions to reduce patients' pain intensity.

KEYWORDS

humanistic care, cancer pain, psychological factors, psychological functions, pain intensity

1. Introduction

Pain is one of the most prevalent consequences of cancer, although increasing attention on the assessment and management of it. Pain prevalence rates were reported to be 55.0% during cancer treatment and 66.4% in the advanced stages of cancer (1). Recently, the International Association for the Study of Pain (IASP) revised the definition of pain that had been in use for 40 years, explaining that pain is a personal experience that is affected by biological, psychological, and social factors to varying degrees; people can perceive pain through life experience; and pain may have an impact on patients' physical functions and social and mental health (2).

Sociodemographic factors have been extensively researched to identify trends within populations with chronic pain. In general, research has shown that African American individuals, those from rural and low-income communities, and individuals with lower levels of education and literacy are more vulnerable to the harmful effects of suffering (3, 4).

The existing literature (5–7) shows that pain catastrophizing, resilience, anxiety, and depression may affect individual pain perception and expression. Pain catastrophizing is a significant psychological factor involved in regulating behavioral responses to pain. It is defined as a belief system, coping strategy, and evaluation process when experiencing pain (6). Resilience can be defined as an individual's ability to recover or "bounce back" from negative events and maintain their function (or even thrive and grow) in the face of ongoing stress (8). Aside from this, research has suggested that pain is related to mental health problems in patients with cancer, but the possible causation and direction of these associations are not clear (9, 10). The intensity of pain and the states of anxiety and depression also interact with each other; for example, the severity of depressive symptoms is associated with the frequency of pain complaints (11).

In addition, socioeconomic status could moderate the impact of psychological factors (catastrophizing and resilience) on pain intensity and psychological functions (depression and anxiety). A study in Nepal found that both pain intensity and income moderated the association between resilience and physical function in individuals with chronic pain, while income moderated the association between resilience, catastrophizing, and depression (6). Another study on a population of patients with chronic pain in rural Alabama indicated that age notably mediated the relationship between catastrophizing, depression, and pain (3). Robert et al. (12) also found that the relationship between catastrophizing and pain intensity was significantly moderated by education and social functioning in patients with rheumatoid arthritis in the United States.

Humanistic care involves a fundamental belief in the power of the care process to produce growth and change for people (13). Humanistic care can help patients to eliminate fear in multiple dimensions, improve their psychological threshold for pain, and become aware of pain control measures, thus enabling them to better cooperate with the treatment (14). Clinical pharmacists are professionals who are licensed pharmacists with specialized advanced training and provide patients with comprehensive drug management and related care in all medical areas (15, 16). Humanistic care is one of the intervention contents of clinical pharmacists. Pharmacist-led interventions have yielded excellent results and have been shown to play a positive role in many areas, such as when including pharmacists in cancer pain multidisciplinary management teams (17).

Humanistic care may moderate the relationship between pain intensity, psychological factors (catastrophizing and resilience), and psychological functions (depression and anxiety). Furthermore, most previous studies have focused on the moderating effect of socioeconomic status (e.g., education level and income), which cannot easily be changed in a short time, on the relationships among pain intensity, psychological factors, and psychological functioning.

In January 2018, three clinical pharmacists with professional qualifications in pain were assigned to the oncology department to provide multifaceted interventions for pain management, humanistic care is included in it. The multifaceted interventions included: (1) daily ward round: made ward rounds with the physician every day (working days only) to assess the patient's pain, medication, and laboratory results, and advised the physician to determine the optimal drug treatment; (2) regular review of medical orders: checked each patient's temporary and long-term medical orders and gave feedback and explanation of the problematic orders to the physician; and (3) humanistic care: humanistic care was defined as providing patients with necessary one-on-one and face-to-face medication guidance and education for patients when they are receptive and able to cooperate. To illustrate, when patients did not accept using opiates because of concerns about its addictive properties, the clinical pharmacists would tell patients that, with the correct use, addiction would not occur. When patients had a poor emotional state, the clinical pharmacists would talk with them and teach them some methods to change their perceptions. When patients struggled with the belief that their pain was uncontrollable, the clinical pharmacists would educate them that, with reasonable treatment, the pain could be relieved. Patients who do not accept or cannot cooperate were not given humane care.

In view of this fact, the objective of this study is to explore whether humanistic care practiced by clinical pharmacists and socioeconomic status moderate the association among pain intensity, psychological factors (catastrophizing and resilience), and psychological functions (depression and anxiety) in patients with cancer with low levels education and income in the Shanxi province in the Northwest of China. In 2020, China's average *per capita* GDP value is 114,808 yuan, and Shanxi Province, with a *per capita* GDP of 50,528 yuan.

2. Research methods

2.1. Sample and setting

This was a cross-sectional study and was performed with a sample of inpatients with cancer pain between August 2018 and August 2021 at The Second Hospital of Shanxi Medical University, a 2,700-bed academic teaching hospital in Taiyuan, China. The sample size was estimated by the statistical calculation formula of a cross-sectional survey of related factors (5).

We included patients who met the following criteria: (1) hospital inpatients; (2) aged ≥ 18 years; (3) diagnosed with cancer; (4) conscious, could communicate independently, and could express their wishes clearly; (5) suffered from cancer pain for at least 1 week; (6) live in Taiyuan City or its surrounding areas, including county towns and rural areas.

The exclusion criteria were: (1) diagnosed with psychiatric or mental disorders, such as schizophrenia, bipolar disorder, and depression by the physician; (2) cognitive disorders; and (3) being unable to complete the questionnaires.

One clinical pharmacist recorded all these works. It is important to note that humane care, which was routine work only studied as a moderator, like socioeconomic status, not as an intervention in this study. Another clinical pharmacist identified potentially eligible patients by reviewing their medical records and psychiatric history. The eligible patients were first informed about the purpose and protocol of the study. Secondly, they verbally told consent to participate in the research if they agreed to participate; at the same time, they informed them that all information would be protected. For participants who could not read or write, the investigator read out the questionnaire items word by word without any further explanation and completed the questionnaires based on the patient's responses.

The study adopted the 5th day of participants' pain score and provided participants with questionnaires. The entire investigation may last 10–20 min. When they completed the questionnaires, investigators checked and asked participants to fill in any missing items.

2.2. Measures

Demographic variables were obtained from the Hospital Information System (HIS) of The Second Hospital of Shanxi Medical University. The variables of interest were age, gender, income, marital status, education level, living area, the primary site of cancer, degree of disease progression, and type of pain. Humanistic care, pain intensity, psychological factors, and psychological functions were evaluated by clinical pharmacists using five scales during the daily ward rounds of the multifaceted pharmacist-led guidance team.

2.2.1. Pain intensity

The Faces Pain Rating Scale (FPS-R; IASP, 2001), used with permission from the IASP, is a self-reported pictorial scale that consists of six faces showing increasing levels of pain. The respondents are asked to select a face that best represents their level of pain at the time of assessment (2).

2.2.2. Resilience

Psychological resilience was assessed using the Chinese version of the Conner and Davidson resilience scale (CD-RISC). The 25-item CD-RISC contains three subscales, namely tenacity (13 items), strength (8 items), and optimism (4 items). It is rated using a 5-point Likert scale from 0 (not true at all) to 4 (true all the time), with a total score of 0–100. Higher scores indicate higher levels of psychological resilience. The Cronbach's α coefficient of the scale in the present study was 0.927 (18).

2.2.3. Pain catastrophizing

The Chinese version of the Pain Catastrophizing Scale (PCS) was used to assess patient reports of catastrophic thinking. The 13-item scale asks respondents to rate the degree to which they have certain thoughts and feelings when experiencing pain using a 5-point Likert scale ranging from 0 (not at all) to 4 (all the time). The total score for overall catastrophizing is equal to the sum of the raw scores. Higher scores indicate greater levels of catastrophic thinking. The Cronbach's α coefficient of the scale in the present study was 0.91 (19).

2.2.4. Anxiety and depression

Anxiety and depression levels were assessed with the Hospital Anxiety and Depression Scale (HADS), which is a 14-item inventory used to examine the degree of anxiety and depression of patients in nonpsychiatric hospitals. The HADS has two subscales—the anxiety subscale (HADS-A) and depression subscale (HADS-D)—each consisting of seven items. A 4-point Likert scale (0–3) is used to rate the items. Higher scores represent more severe psychological distress. This instrument is widely used in clinical settings, and the Chinese version used in the current study has sound reliability, with a Cronbach's alpha coefficient of 0.832. The Cronbach's alpha coefficients of the HADS-A and HADS-D subscales were 0.753 and 0.764, respectively (20).

2.2.5. Humanistic care

The humanistic care ability of clinical pharmacists was assessed with the Humanistic Care Scale (HCS), which is a 5-item to evaluate the humanistic care ability of clinical pharmacists by patients. This scale was referenced to the Watson Caritas Patient Score (WCPS). A 7-point Likert scale (1–7) is used to rate the items. The items empirically assess the patient's subjective experience of receiving humanistic care. The items refer to such indicators as loving kindness, trust, dignity, a healing environment, and honoring beliefs and values. The total score ranged from 5 to 35, with higher scores indicating better humanistic care ability. The Cronbach's α coefficient of the scale in the present study was 0.835.

2.3. Data analysis

All statistical analyses were performed using IBM SPSS 25.0 (IBM Corp., Armonk, NY, United States). Due to the methods of data collection, missing data were minimal, and thus, data imputations were not utilized in this analysis. Descriptive statistical analysis: the patients' general demographic data and clinically relevant data were described by percentage.

Correlation Test of Social Factors, Humane Care (Independent Variable) and Pain Intensity, Psychological Factors and Psychological Function (Dependent Variable): When the dependent variable is a continuous variable, the independent variable is categorical, One-Way ANOVA (multivariate variable) and t-test (binary variable) are used.

Correlation test between Pain Intensity, Psychological Factors, and Psychological Function: (1) taking pain intensity as the

dependent variable and psychological factors as the independent variable; (2) taking pain intensity as the dependent variable and psychological function as the independent variable; (3) taking a psychological function as the dependent variable and psychological factors as the independent variable, using Pearson correlation analysis or Spearman correlation analysis.

Moderating effect tests: (1) we performed moderating effect tests of socioeconomic status on the relationship among pain intensity, psychological factors, and psychological functioning with hierarchical regression analysis. We used pain intensity and anxiety as the dependent variables. Subsequently, we performed moderating effect tests of socioeconomic status on the relationship between pain intensity, psychological factors, and psychological functioning. In the first step, we entered anxiety when testing pain intensity and entered pain intensity when testing anxiety to control the potential confounding effects on both the predictor and criterion variables. In the second step, we entered the socioeconomic variables of education level and income. In the third step, we entered the psychological variables of pain catastrophizing and resilience. In the fourth step, we entered the 12 interaction terms representing income × anxiety, income × depression, income × catastrophizing, income × resilience interaction effects, education × anxiety, education × depression, education × catastrophizing, and education × resilience interaction effects stepwise. (2) we performed moderating effect tests of humanistic care on the relationship between pain intensity, psychological factors, and psychological functioning. In the first step, we entered humanistic care. In the second step, we entered the psychological variables of pain catastrophizing and resilience. In the third step, we entered five interaction terms representing humanistic care × catastrophizing, humanistic care × resilience, humanistic care × pain intensity, humanistic care × anxiety, and humanistic care × depression interaction effects stepwise. Statistical significance was set at the level of 0.05 or less (two-tailed).

Outliers and missing data were not found in our study. All variables were normally distributed. The data met the necessary hierarchical regression analysis.

The study was approved by the research ethics committee of The Second Hospital of Shanxi Medical University (2021–242).

3. Results

3.1. Participant attributes

We enrolled 51 male and 72 female patients in the study ($N = 123$). Their average age was 56.26 years, with an SD of 19.09 years. More than half of the patients had 6 years of education or less ($n = 69$, 56.1%). Most of the participants ($n = 120$, 97.6%) had medical insurance. The highest incidence of carcinoma was chest tumors ($n = 37$, 30.1%), followed by abdominal tumors ($n = 29$, 23.6%). In total, 65.0% of the patients were locally advanced, and 61.8% of them were suffering from mixed pain. The demographic and clinical characteristics are summarized in Table 1.

TABLE 1 Description of the study sample (N=123).

Characteristic	*N* (%)
Age (years)	
≤ 50	33 (26.8)
51–70	63 (51.2)
> 70	27 (22.0)
Sex	
Male	51 (41.5)
Female	72 (58.5)
Educational level	
None	9 (7.3)
Primary/below (≤ 6 years)	60 (48.8)
Middle (7–12 years)	39 (31.7)
High	15 (12.2)
Income	
0	9 (7.3)
≤ 1,000	51 (41.5)
1,000–3,000	45 (36.6)
≥ 3,000	18 (14.6)
Medical insurance type	
Provincial/city insurance	105 (85.4)
Resident health insurance	15 (12.2)
Own expense	3 (2.4)
Living area	
City (TaiYuan)	38 (30.9)
County seat	33 (26.8)
Rural area	52 (42.3)
Primary cancer site	
Abdominal tumor	29 (23.6)
Urinary tumor	3 (2.4)
Chest tumor	37 (30.1)
Cervical cancer	15 (12.2)
Osteosarcoma	12 (9.8)
Leukemia and lymphoma	3 (2.4)
Head and neck	11 (8.9)
Breast cancer	13 (10.6)
Extent of disease	
First stage of cancer	43 (35.0)
Locally advanced	80 (65.0)
Type of pain	
Nociceptive pain	44 (35.8)
Neuropathic pain	3 (2.4)
Mixed pain	76 (61.8)

3.2. The correlation between variables.

The univariate correlations among the study variables are presented in Tables 2, 3. As can be seen, education, sex, and age in years were not significantly related to any of the standard variables. Humanistic care was significantly related to depression and marginally statistically associated with anxiety and pain intensity. Income had a significant correlation with resilience ($p < 0.05$). Anxiety levels showed

a statistically significant moderate positive correlation with pain intensity ($r = 0.361$, $p < 0.05$). There was a statistically significant moderate negative correlation between both anxiety and depression and resilience ($r = -0.346$, $p < 0.05$ and $r = -0.423$, $p < 0.01$, respectively). Catastrophizing showed a statistically significant moderate negative correlation with resilience ($r = -0.435$, $p < 0.01$). There was a statistically significant strong positive correlation between both anxiety and depression and catastrophizing ($r = 0.702$, $p < 0.01$ and $r = 0.597$, $p < 0.01$, respectively).

3.3. Moderating effects of humanistic care and socioeconomic status on the relationship among pain intensity and psychological factors and psychological function

The results of the moderating effect test are presented in Tables 4, 5. In the first step, anxiety made a statistically significant contribution to pain intensity, and pain intensity made a statistically significant contribution to anxiety. As can be seen, education moderated the associations of resilience and pain catastrophizing with pain intensity. Pain intensity and depression moderated the association of pain catastrophizing with anxiety, and income moderated the association between resilience and anxiety (Table 4). Furthermore, humanistic care moderated not only the association between resilience and pain intensity but also the association between pain intensity and anxiety (Table 5).

4. Discussion

The key finding from this study was that humanistic care practiced by clinical pharmacists moderated the associations among pain intensity, psychological factors, and psychological functions, which has rarely been studied previously. From another perspective, these findings suggest that pharmacist-led interventions play a positive role in cancer pain multidisciplinary management teams.

The frequencies of depression and anxiety are higher in cancer patients, but prevalence rates vary greatly between studies. In patients with cancer, estimated prevalence rates range between 11 and 57% for depression and between 6.5 and 23% for anxiety (21, 22). The results of our study showed that the incidence of depression in patients with cancer pain was 48.78%, within the range of previous literature reports. However, the incidence of anxiety was 41.46%, which is higher than the previously reported range. Naser et al. (22) found that anxious symptomatology was more prevalent in patients with lung cancer in inpatient settings. Similarly, the most common cancer type in our study was lung cancer (27.9%). Additionally, the frequency of depression was higher than anxiety in our study, which is consistent with other studies (23, 24). Patients who were in advanced disease stages were particularly susceptible to suffering from depression, and 65.9% of our patients were in advanced disease stages. Our study reported a low level of psychological resilience (63.37 ± 21.74), which was similar to the level found in Chinese cancer patients in a previous study (65.46 ± 13.93) (25). Low resilience is linked to mood disorders (18), and this may, thus, be a reason for the high rates of anxiety and depression detected in our sample.

Through a univariate analysis, we found that pain intensity was notably associated with anxiety. Unseld et al. (21) highlighted that most studies suggest that depression may be more frequently related to pain than anxiety, but the results are controversial. The possible reason for pain intensity being associated with anxiety in this study is that our sample included a wide range of cancer types, while the samples of those previously reported studies focused on specific cancer types, such as colorectal cancer, breast cancer, or lung cancer.

Pain catastrophizing is considered one of the most important modifiable psychosocial predictors of pain intensity (26). Our analysis revealed that pain catastrophizing was not notably associated with pain intensity, which is inconsistent with prospective studies (27), which have found that pain catastrophizing is a robust predictor of greater pain severity. However, other studies also highlight that, although pain catastrophizing is commonly associated with pain intensity, there is limited evidence showing that changes in pain catastrophizing causes changes in pain (26, 28). Rizzo et al. (26) performed longitudinal assessments for the mediating effect of pain

TABLE 2 The *p*-value of comparisons between categorical variables.

	Catastrophizing	Anxiety	Depression	Resilience	Pain intensity
Educational level	0.209	0.216	0.418	0.774	0.347
Income	0.155	0.115	0.157	**0.035**	0.350
Age (years)	0.630	0.556	0.846	0.075	0.491
Sex	0.750	0.529	0.498	0.135	0.274
Humanistic care	0.479	0.061	**0.043**	0.478	0.059

When $p < 0.05$, the values have been highlighted in bold.

TABLE 3 Mean and SD values of the continuous variables and correlation coefficients between the continuous variables.

	Mean±SD	Catastrophizing	Anxiety	Depression	Resilience	Pain intensity
Catastrophizing	26.61 ± 13.32		0.702***	0.597***	−0.435***	0.293
Anxiety	7.54 ± 4.89				−0.346**	0.361**
Depression	7.85 ± 5.30				−0.423***	0.138
Resilience	63.37 ± 21.74					−0.172

**$p < 0.05$ and ***$p < 0.01$.

TABLE 4 The moderating effect of socioeconomic status on the relationship among pain intensity, psychological factors, and psychological functioning.

	Total R^2	ΔR^2	F-ΔR^2	Standardized beta coefficient (B)	*p*-Value
*Pain Intensity as the criterion variable, Educational Level*Resilience as the interaction term*					
	0.147	0.147	6.591		
Anxiety				0.084	0.694
	0.220	0.073	3.412		
Income				0.176	0.303
Education level				−1.812	0.008
	0.222	0.002	1.994		
Resilience				−1.547	0.339
Catastrophizing				0.190	0.003
	0.380	0.158	3.311		
Educational Level*Resilience				2.511	0.002
*Pain Intensity as the criterion variable, Educational Level*Catastrophizing as the interaction term*					
	0.147	0.147	6.591		
Anxiety				0.322	0.118
	0.220	0.073	3.412		
Income				0.167	0.338
Education level				1.207	0.003
	0.222	0.002	1.994		
Resilience				1.273	0.009
Catastrophizing				−0.005	0.974
	0.408	0.186	3.909		
Educational Level*Catastrophizing				−1.627	0.006
*Anxiety as the criterion variable, Income*Resilience as the interaction term*					
	0.147	0.147	6.591		
Pain Intensity				0.224	0.047
	0.275	0.127	4.543		
Income				0.147	0.227
Education level				−1.033	0.001
	0.575	0.301	9.177		
Resilience				−0.779	0.005
Catastrophizing				0.564	0.000
	0.672	0.097	11.523		
Income*Resilience				1.300	0.003

catastrophizing on pain intensity and drew the conclusion that the timing of the assessment influenced the mediating role of pain catastrophizing on pain intensity. However, we did not conduct the self-report measures of pain catastrophizing with patients at a fixed time because of the absence of patients when we made ward rounds. This may explain why pain catastrophizing was not notably associated with pain intensity in our investigation.

The results of the moderating effect test showed that neither pain catastrophizing nor resilience made statistically significant independent contributions to the prediction of pain intensity. However, when adding the moderator of education level, both pain catastrophizing and resilience had statistically significant relationships with pain intensity. Importantly, the finding that education level moderated the relationship between pain catastrophizing and pain intensity is consistent with previous studies, which found that high pain catastrophizing was linked to low education, which, in turn, led to inappropriate pain-coping strategies (29). Indeed, Cano et al. suggested that numerous pain-coping strategies, such as the ability to distract and reinterpret, may rely on cognitive skills that are potentially enhanced by higher education and primary literacy (3). Individuals with lower levels of literacy may have fewer cognitive resources available to navigate the management of chronic pain, thus increasing the risk for distress and negative thinking patterns and ultimately exacerbating the pain condition (3). Furthermore, cognitive flexibility

TABLE 5 The moderating effect of humanistic care on the relationship among pain intensity, psychological factors, and psychological functioning.

	Total R^2	ΔR^2	F-ΔR^2	Standardized beta coefficient(B)	p-Value
*Pain Intensity as the criterion variable, Humanistic care*Resilience as the interaction term*					
	0.147	0.147	6.591		
Anxiety				0.405	0.092
	0.150	0.002	3.339		
Humanistic care				−1.096	0.071
	0.151	0.001	1.594		
Resilience				−1.048	0.045
Catastrophizing				−0.077	0.734
	0.253	0.102	2.265		
Educational Level*Resilience				1.307	0.035
*Anxiety as the criterion variable, Humanistic care*Pain Intensity as the interaction term*					
	0.147	0.147	6.591		
Anxiety				−0.828	0.011
	0.560	0.413	23.771		
Humanistic care				−0.464	0.122
	0.627	0.067	15.055		
Resilience				−0.130	0.210
Catastrophizing				0.359	0.008
	0.718	0.092	17.770		
Humanistic care*Pain Intensity				1.519	0.002

is reported to be a critical factor in preventing negative outcomes and suicidal behavior in response to stressful life events (30). Overall, individuals with high levels of literacy may have more resources available to cope with stress and the burden of illness (18, 31). When patients with chronic diseases have higher mental resilience, they show higher degrees of acceptance of the disease, higher compliance with the treatment plan, and better prognoses (18, 32).

Regarding depression and anxiety, depression has received more attention from researchers, and its adverse effects on physical functioning and quality of life are well established (33). However, we chose to discuss anxiety, which has been studied less frequently, as a predictor of pain intensity, because pain intensity was not significantly associated with depression in the current study. The results of this study indicated that higher income contributed to a higher level of psychological resilience in patients with cancer pain, which supports the theory proposed by Wister et al. (34), and income significantly moderated the association between resilience and anxiety (income × resilience interaction; $\beta = 1.300$, $p = 0.003$). These data are consistent with reports describing the prediction of depression. However, income was not significantly associated with anxiety in our investigation, which may have resulted from the fact that nearly half of the sample were unemployed or farmers, whose incomes are at low levels; indeed, such drastic poverty may function as a leveling factor (29). People with low incomes experience negative emotions, which in turn affect resilience levels (35).

More importantly, considering the moderators of education level and income cannot be changed easily in a short time, we further investigated the moderating effect of humanistic care. In the present investigation, humanistic care practiced by clinical pharmacists moderated not only the association between resilience and pain intensity but also the association between pain intensity and anxiety. This suggests that, with patients with low socioeconomic status, medical staff should focus more on humanistic care to reduce their negative emotions and relieve their pain intensity. A previous study suggested that health knowledge education could work in the short term, especially when patients were seriously ill or had severe pain (14). Additionally, Edwards et al. confirmed that pharmacist educational interventions for cancer pain patients showed promise in reducing pain intensity (36). A number of publications have indicated that the multifaceted pharmacist-led guidance team intervention successfully decreases drug-related problems and shows both initial and prolonged pain relief (37). In summary, humanistic care practiced by clinic pharmacists could improve patients' awareness of cancer pain to enable them to overcome their fears and build confidence, thus making pain management more humanized, scientific, and comprehensive to effectively relieve pain.

4.1. Study limitations

The findings of the current study have a number of limitations that should be considered when interpreting the results. Firstly, this study used cross-sectional data, which limits the conclusions that can be drawn with respect to causal relationships. The underlying reasons for the associations found in the present analyses remain to be fully understood. It is possible to use longitudinal measurements to examine the relationship between mediator and outcome variables and allow inferences of causality in further research. Secondly, the

sample's demographic homogeneity is a potential limitation; to determine whether rurality itself is a predictor of poorer pain outcomes, it would be important to compare the findings of this rural population with low socioeconomic status to those of an urban population with similar demographic features. Thirdly, the sample was obtained from a single institution during a limited study period, and, thus, the results may not be widely representative or generalizable.

4.2. Clinical implications

Our research emphasizes the importance of humanistic care practiced by clinical pharmacists for patients with cancer and low levels of education and income in the Northwest of China. Clinical pharmacists could better provide patients with cancer pain with cognitive resources to reduce their negative thoughts and improve their awareness in order to overcome fear, build confidence, and increase their mental resilience in a short time. Furthermore, this would improve their acceptance of pain, enhance their compliance with treatment plans, and enhance the therapeutic effects.

Additionally, the results of this study highlight the need to pay more attention to screening for psychiatric disorders, such as depression and anxiety, in inpatients with cancer pain. To optimize treatment, a positive screening result should be followed by a thorough psychiatric diagnostic interview conducted face-to-face. Therefore, adequate pain-related treatment should be discussed by a multidisciplinary team, which may include doctors, clinical pharmacists, and nurses.

5. Conclusion

This study found that humanistic care plays an important role in moderating the associations among pain intensity, psychological factors, and psychological functions in Chinese patients with cancer, especially for those from counties and rural areas with lower levels of income. From another perspective, this study shows that pharmacist-led interventions play a positive role in cancer pain multidisciplinary management teams.

Furthermore, in this study, there was a high incidence of both anxiety and depression, and pain intensity was significantly associated with humanistic care and anxiety. After adjusting for these associations, the results showed that education levels moderate the relationship between pain intensity and both pain catastrophizing and resilience. Additionally, income moderates the relationship between resilience and anxiety.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by The research ethics committee of the Second Hospital of Shanxi Medical University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

SW, XW, XL, and CZ conducted the literature review and did the statistical analysis. SW, XW, XL, CZ, and JD contributed to the interpretation of data and to the study concept and design. SW, XL, and JD contributed to drafting the paper. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Second Hospital of Shanxi Medical University Science Foundation (Grant number 202202-7) and Shanxi Provincial Health Commission Science Foundation (Grant number 2019118).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. van den Beuken-van Everdingen MH, Hochstenbach LM, Joosten EA, Tjan-Heijnen VC, Janssen DJ. Update on prevalence of pain in patients with cancer: systematic review and meta-analysis. *J Pain Symptom Manag.* (2016) 51:1070–1090e9. doi: 10.1016/j.jpainsymman.2015.12.340

2. Raja SN, Carr DB, Cohen M, Finnerup NB, Flor H, Gibson S, et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. *Pain.* (2020) 161:1976–82. doi: 10.1097/j.pain.0000000000001939

3. Newman AK, Van Dyke BP, Torres CA, Baxter JW, Eyer JC, Kapoor S, et al. The relationship of sociodemographic and psychological variables with chronic pain variables in a low-income population. *Pain.* (2017) 158:1687–96. doi: 10.1097/j.pain.0000000000000964

4. Goode AP, Freburger JK, Carey TS. The influence of rural versus urban residence on utilization and receipt of care for chronic low back pain. *J Rural Health.* (2013) 29:205–14. doi: 10.1111/j.1748-0361.2012.00436.x

5. Xu X, Ou M, Xie C, Cheng Q, Chen Y. Pain acceptance and its associated factors among cancer patients in mainland China: a cross-sectional study. *Pain Res Manag.* (2019) 2019:1–7. doi: 10.1155/2019/9458683

6. Sharma S, Pathak A, Jha J, Jensen MP. Socioeconomic factors, psychological factors, and function in adults with chronic musculoskeletal pain from rural Nepal. *J Pain Res.* (2018) 11:2385–96. doi: 10.2147/JPR.S173851

7. Poulin PA, Romanow HC, Rahbari N, Small R, Smyth CE, Hatchard T, et al. The relationship between mindfulness, pain intensity, pain catastrophizing, depression, and

quality of life among cancer survivors living with chronic neuropathic pain. *Support Care Cancer*. (2016) 24:4167–75. doi: 10.1007/s00520-016-3243-x

8. Sturgeon JA, Zautra AJ. Resilience: a new paradigm for adaptation to chronic pain. *Curr Pain Headache Rep*. (2010) 14:105–12. doi: 10.1007/s11916-010-0095-9

9. Morasco BJ, Lovejoy TI, Lu M, Turk DC, Lewis L, Dobscha SK. The relationship between PTSD and chronic pain: mediating role of coping strategies and depression. *Pain*. (2013) 154:609–16. doi: 10.1016/j.pain.2013.01.001

10. Lillis TA, Gerhart J, Bouchard LC, Cvengros J, O'Mahony S, Kopkash K, et al. Sleep disturbance mediates the association of post-traumatic stress disorder symptoms and pain in patients with cancer. *Am J Hosp Palliat Care*. (2018) 35:788–93. doi: 10.1177/1049909117739299

11. Von Korff M, Dworkin SF, Le Resche L, Kruger A. An epidemiologic comparison of pain complaints. *Pain*. (1988) 32:173–83. doi: 10.1016/0304-3959(88)90066-8

12. Edwards RR, Giles J, Bingham CO 3rd, Campbell C, Haythornthwaite JA, Bathon J. Moderators of the negative effects of catastrophizing in arthritis. *Pain Med*. (2010) 11:591–9. doi: 10.1111/j.1526-4637.2010.00804.x

13. Shiau SJ, Chen CH. Reflection and critical thinking of humanistic care in medical education. *Kaohsiung J Med Sci*. (2008) 24:367–72. doi: 10.1016/S1607-551X(08)70134-7

14. Shi M, Zhang P, Xia L, Wei Z, Bi F, Xu Y, et al. Application of multimode health education combined with humanistic care in pain management of patients with femoral fracture and its influence on VAS score. *J Healthc Eng*. (2021) 2021:1242481. doi: 10.1155/2021/1242481

15. Du Y, Li J, Wang X, Peng X, Wang X, He W, et al. Impact of a multifaceted pharmacist-led intervention on antimicrobial stewardship in a gastroenterology ward: a segmented regression analysis. *Front Pharmacol*. (2020) 11:442. doi: 10.3389/fphar.2020.00442

16. Liu J, Wang C, Chen X, Luo J, Xie J, Li S, et al. Evaluation of pharmacist interventions as part of a multidisciplinary cancer pain management team in a Chinese academic medical center. *J Am Pharm Assoc*. (2003) 60:76–80. doi: 10.1016/j.japh.2019.09.005

17. Chisholm-Burns MA, Spivey CA, Sherwin E, Wheeler J, Hohmeier K. The opioid crisis: origins, trends, policies, and the roles of pharmacists. *Am J Health Syst Pharm*. (2019) 76:424–35. doi: 10.1093/ajhp/zxy089

18. Qiu Y, Huang Y, Wang Y, Ren L, Jiang H, Zhang L, et al. The role of socioeconomic status, family resilience, and social support in predicting psychological resilience among Chinese maintenance hemodialysis patients. *Front Psych*. (2021) 12:723344. doi: 10.3389/fpsyt.2021.723344

19. Xu X, Wei X, Wang F, Liu J, Chen H, Xiong Y, et al. Validation of a simplified Chinese version of the pain catastrophizing scale and an exploration of the factors predicting catastrophizing in pain clinic patients. *Pain Physician*. (2015) 18:E1059–72. doi: 10.36076/ppj.2015/18/E1059

20. Yang Y, Ding R, Hu D, Zhang F, Sheng L. Reliability and validity of a Chinese version of the HADS for screening depression and anxiety in psycho-cardiological outpatients. *Compr Psychiatry*. (2014) 55:215–20. doi: 10.1016/j.comppsych.2013.08.012

21. Unseld M, Zeilinger EL, Fellinger M, Lubowitzki S, Krammer K, Nader IW, et al. Prevalence of pain and its association with symptoms of post-traumatic stress disorder, depression, anxiety and distress in 846 cancer patients: a cross sectional study. *Psychooncology*. (2021) 30:504–10. doi: 10.1002/pon.5595

22. Naser AY, Hameed AN, Mustafa N, Alwafi H, Dahmash EZ, Alyami HS, et al. Depression and anxiety in patients with cancer: a cross-sectional study. *Front Psychol*. (2021) 12:585534. doi: 10.3389/fpsyg.2021.585534

23. Teunissen SC, de Graeff A, Voest EE, de Haes JC. Are anxiety and depressed mood related to physical symptom burden? A study in hospitalized advanced cancer patients. *Palliat Med*. (2007) 21:341–6. doi: 10.1177/0269216307079067

24. Ciaramella A, Poli P. Assessment of depression among cancer patients: the role of pain, cancer type and treatment. *Psychooncology*. (2001) 10:156–65. doi: 10.1002/pon.505

25. Yu XN, Lau JT, Mak WW, Zhang J, Lui WW, Zhang J. Factor structure and psychometric properties of the Connor-Davidson resilience scale among Chinese adolescents. *Compr Psychiatry*. (2011) 52:218–24. doi: 10.1016/j.comppsych.2010.05.010

26. Rizzo R, Lee H, Cashin AG, Costa L, Gustin SM, McAuley JH. The mediating effect of pain catastrophizing on pain intensity: the influence of the timing of assessments. *Eur J Pain*. (2021) 25:1938–47. doi: 10.1002/ejp.1810

27. Kovacs FM, Seco J, Royuela A, Corcoll-Reixach J, Peña-Arrebola A. The prognostic value of catastrophizing for predicting the clinical evolution of low back pain patients: a study in routine clinical practice within the Spanish National Health Service. *Spine J*. (2012) 12:545–55. doi: 10.1016/j.spinee.2012.06.002

28. Lee H, Mansell G, McAuley JH, Kamper SJ, Hübscher M, Moseley GL, et al. Causal mechanisms in the clinical course and treatment of back pain. *Best Pract Res Clin Rheumatol*. (2016) 30:1074–83. doi: 10.1016/j.berh.2017.04.001

29. Day MA, Thorn BE. The relationship of demographic and psychosocial variables to pain-related outcomes in a rural chronic pain population. *Pain*. (2010) 151:467–74. doi: 10.1016/j.pain.2010.08.015

30. de Berardis D, Fornaro M, Valchera A, Cavuto M, Perna G, di Nicola M, et al. Eradicating suicide at its roots: preclinical bases and clinical evidence of the efficacy of ketamine in the treatment of suicidal behaviors. *Int J Mol Sci*. (2018) 19:2888. doi: 10.3390/ijms19102888

31. Karadag E, Kilic SP, Metin O. Relationship between fatigue and social support in hemodialysis patients. *Nurs Health Sci*. (2013) 15:164–71. doi: 10.1111/nhs.12008

32. García-Martínez P, Ballester-Arnal R, Gandhi-Morar K, Castro-Calvo J, Gea-Caballero V, Juárez-Vela R, et al. Perceived stress in relation to quality of life and resilience in patients with advanced chronic kidney disease undergoing hemodialysis. *Int J Environ Res Public Health*. (2021) 18:536. doi: 10.3390/ijerph18020536

33. Brown LF, Kroenke K, Theobald DE, Wu J, Tu W. The association of depression and anxiety with health-related quality of life in cancer patients with depression and/or pain. *Psychooncology*. (2010) 19:734–41. doi: 10.1002/pon.1627

34. Wister AV, Coatta KL, Schuurman N, Lear SA, Rosin M, MacKey D. A lifecourse model of multimorbidity resilience: theoretical and research developments. *Int J Aging Hum Dev*. (2016) 82:290–313. doi: 10.1177/0091415016641686

35. Tompkins DA, Hobelmann JG, Compton P. Providing chronic pain management in the "fifth vital sign" era: historical and treatment perspectives on a modern-day medical dilemma. *Drug Alcohol Depend*. (2017) 173:S11–21. doi: 10.1016/j.drugalcdep.2016.12.002

36. Edwards Z, Ziegler L, Craigs C, Blenkinsopp A, Bennett MI. Pharmacist educational interventions for cancer pain management: a systematic review and meta-analysis. *Int J Pharm Pract*. (2019) 27:336–45. doi: 10.1111/ijpp.12516

37. Bruhn H, Bond CM, Elliott AM, Hannaford PC, Lee AJ, McNamee P, et al. Pharmacist-led management of chronic pain in primary care: results from a randomised controlled exploratory trial. *BMJ Open*. (2013) 3:e002361. doi: 10.1136/bmjopen-2012-002361

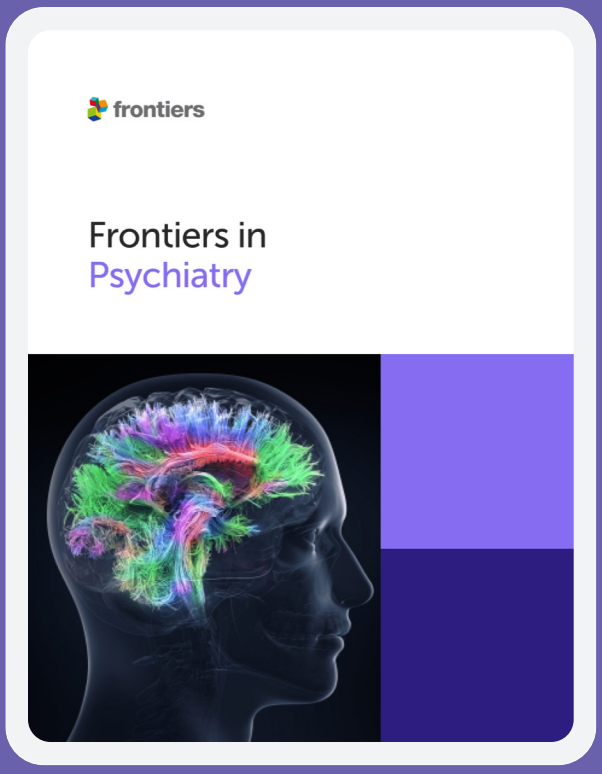
Frontiers in Psychiatry

Explores and communicates innovation in the field of psychiatry to improve patient outcomes

The third most-cited journal in its field, using translational approaches to improve therapeutic options for mental illness, communicate progress to clinicians and researchers, and consequently to improve patient treatment outcomes.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact