



# Brief screening tools to facilitate referral of people with emotional disorders in primary care: A Cohort Study on the use of the Overall Depression and Anxiety Severity and Impairment Scales

Ó. Peris-Baquero<sup>1,2</sup> · A. Fadrique-Jiménez<sup>1</sup> · L. Martínez-García<sup>1,2</sup> · I. Prieto-Rollan<sup>1</sup> · Jorge Osma<sup>1,2,3</sup>

Accepted: 14 March 2025  
© The Author(s) 2025

## Abstract

Emotional disorders (EDs) represent a significant volume of consultations in Primary Care (PC), where family physicians do not have the time or tools to assess them effectively. Therefore, individuals are referred to Mental Health Units (MHUs) regardless of the severity of their symptoms. The aim of this study is to explore the screening ability of two brief measures, ODSIS and OASIS. A cohort study was conducted with 256 individuals referred from PC to Spanish MHUs with a main diagnosis of EDs. The variables assessed at pre-psychological treatment, post and up to the 15-month follow-up were anxiety, depression, neuroticism, extraversion, quality of life, interference, and emotional regulation. Participants were retrospectively divided into three groups based on their baseline scores on the severity and interference of depressive (ODSIS) and anxious (OASIS) symptomatology. The results showed statistically significant differences between the three groups in all variables and times assessed, with effect sizes ranging from moderate to large ( $\eta^2$  values between 0.42 and 0.95). Similar results were obtained in the linear mixed models, a main effect of the variable “group” was observed in all variables with effect sizes ranging from small to medium (Cohen’s  $d$  values between 0.20 and 0.49). However, no “Time\*Group” interaction effects were found (except for ODSIS and neuroticism), nor Time\*Group\*Sessions\* so no different evolution was observed over time between groups or based on the number of sessions. ODSIS and OASIS allow an easy and effective discrimination of individuals regarding their emotional symptoms and their differences remain stable over time while receiving a psychological treatment. Trial NCT03064477 (March 10, 2017).

**Keywords** ODSIS · OASIS · Emotional disorder · Primary care · Screening measures

For many individuals, the first place of contact for mental health care is the Primary Care (PC) setting. In the Spanish public health system, this service is organized in territorially delimited community centers, with multidisciplinary teams, which offer health care (Aguilera-Martín et al., 2022).

According to the Annual Report of the Spanish National Health System (NHS) (Ministerio de Sanidad, 2022), approximately 37% of the individuals treated in PC settings in 2021 presented mental health problems. The most frequent problems seen in PC settings in the adult population were anxiety disorders (12.69%), sleep disorders (7.88%) and depressive disorders (4.68%) (Ministerio de Sanidad, 2022). These data are similar to those found in other countries where the higher prevalence of mental disorders is reflected in PC consultations (Smit et al., 2020).

In addition to the PC settings, the Spanish NHS has also specialized Mental Health Units (MHUs) that form the core of the mental health care network. The MHUs are composed of mental health professionals teams (psychiatrists, psychologists, nurses, and social workers mostly) whose function is to provide mental care services for individuals with severe mental disorders (e.g., bipolar disorders, major depressive

✉ Jorge Osma  
osma@unizar.es

<sup>1</sup> Department of Psychology and Sociology, Faculty of Social and Human Sciences, University of Zaragoza, Teruel, Spain

<sup>2</sup> Health Research Institute of Aragon (IIS Aragon), Biomedical Research Center of Aragon (CIBA), Zaragoza, Spain

<sup>3</sup> Departamento de Psicología y Sociología, Facultad de Ciencias Sociales y Humanas, Universidad de Zaragoza, C/ Ciudad Escolar s/n. 44003, Teruel, Spain

disorders, post-traumatic stress disorders, obsessive-compulsive disorders, among others) (Instituto de Información Sanitaria, 2007). Both types of settings, PC and MHUs, are free to the population and are financed through public taxation (Aguilera-Martín et al., 2022).

However, given the increase in the prevalence of emotional disorders (EDs; e.g., anxiety, depressive and related disorders; Bullis et al., 2019) and their high comorbidity (Jurado-González et al., 2024), in recent years MHUs have collapsed as evidenced by: (a) the long waiting lists to receive the first session of psychological treatment, with an average waiting time of four months for ordinary cases and one month for preferential cases, and (b) the long time between sessions (one and a half months on average), which make it impossible to dedicate sufficient time to people who require psychological treatment (Cuéllar-Flores et al., 2022; Peris-Baquero & Osmá, 2023). Furthermore, if we analyze the severity of the symptoms of individuals who are referred to a specialized MHU, we observe that only 18% are referred preferentially, meaning that their severity requires immediate and intensive care, while 82% are referred despite presenting milder or moderate emotional symptomatology (Díaz et al., 2017). These data show that most referrals are characterized by a milder symptomatology that takes several months to be treated, thus causing a worsening of the mental health status of individuals and the collapse of the waiting lists of MHUs (Díaz et al., 2017).

Nevertheless, the current reality shows that the professionals working in PC are mostly family physicians (FP), nurses and social workers, and very few are psychologists (Aguilera-Martín et al., 2022). This situation is problematic, since EDs constitute a substantial volume of consultations (616 cases per 1000 people attended; BDCAP, 2022), and these health professionals do not always receive adequate training to diagnose, refer and treat mental disorders, nor do they have the time to do so efficiently (Huprich, 2018). Then, many of the individuals receive only pharmacological treatment instead of receiving brief psychological support and advice or evidence based psychological treatments as recommended (Fonseca, 2021).

In this regard, clinical guidelines state that the main approach in PC is supportive psychotherapy (e.g., Díaz González, 2008). According to these guidelines, care can be structured in three levels of intervention according to the individual's symptomatology severity: a first level based on emotional support, a second level based on counselling, and a third level based on regulated psychotherapy when individuals present a more severe mental disorder (Díaz González, 2008). In addition, international clinical guidelines such as American Academy of Family Physicians (AAFP) and US Preventive Services Task Force (USPSTF) also highlight the relevance of screening and treatment of

mental health problems to address the needs of individuals attending in PC centres (Barry et al., 2023; Colorafi et al., 2017).

In turn, FP are in charge of screening between those problems with mild or moderate interference (e.g., anxiety, minor depression and adjustment disorders) and those more severe psychological disorders. The former are usually taken on and treated by the FP mostly with psychotropic drugs, and the latter are referred, depending on the severity, to the MHUs or specialized units. Nonetheless, this referral process is relatively complex given that FPs have limited time to see each individual (Cuéllar-Flores et al., 2022), they do not have sufficient training in mental health, and, in some cases, they do not have specialized protocols for PC settings for the assessment of individuals with EDs. Consequently, this has an impact on the ability to detect, treat and refer the different psychological problems (Confederación Salud Mental España, 2023). Considering the barriers mentioned before, it is necessary to design brief and effective mental health assessment and referral protocols to help FPs in their decision-making process.

Several instruments have been evaluated as potential screening tools in primary care, such as the Patient Health Questionnaire-9 (PHQ-9), a 9-item tool used to measure the severity of depressive symptoms in adults (Costantini et al., 2021), or the shorter 4-item version, PHQ-4, which assesses both the severity of depressive symptoms and the severity of anxiety symptoms (Muñoz-Navarro et al., 2017). However, these tools are generally not specific to EDs, are typically administered separately, do not assess the interference caused by symptoms, and many of the studies lack longitudinal evaluations to assess the long-term effectiveness of screening. In recent years, new brief evaluation measures have emerged that have demonstrated good psychometric properties and evidence of validity. Among the possible assessment measures for EDs, the ODSIS (Overall Depression Severity and Impairment Scale; Bentley et al., 2014) and OASIS (Overall Anxiety Severity and Impairment Norman et al., 2006) measures have emerged. Both are short self-administered questionnaires (5 items), they are validated in Spanish (Osmá et al., 2019), and their marking is simple (the sum of the scores for each item provides a total score).

In addition, these measures have a clinical cut-off point, a score greater than or equal to 10 in both ODSIS and OASIS (Osmá et al., 2019), and a normative cut-off point, scores equal to or greater than 4 in the case of OASIS and equal to or greater than 5 in the case of ODSIS (Osmá et al., 2019), so both measures allow us to quantify the severity and interference caused by symptoms of anxiety and depression in adult individuals. Moreover, these instruments are commonly used to assess the evolution of clinical symptomatology

throughout the intervention (Eustis et al., 2020; Timulak et al., 2022), as they allow, in a quick (with only 5 items per measure), simple and reliable way, to assess the severity of anxious and depressive symptoms and their interference. Finally, they have been translated and validated in multiple languages, such as Japanese (Ito et al., 2015), Portuguese (Pereira et al., 2020), Czech (Sandora et al., 2021), German (Hiller et al., 2023) and Danish (Madsen et al., 2024). Given the characteristics of these measures, a screening and referral protocol could be carried out, establishing cut-off points from which individuals could be referred to one service or another, depending on the severity of the symptoms and with the aim of optimizing the available resources.

This approach would be in line with the “step-care” movement, a staged approach to mental health service delivery. This approach is characterized by two main features; on the one hand, the recommended treatment should be the least restrictive and impose the least possible burden on the individual while still provide health benefits. On the other hand, it is a self-correcting model, so treatment can be adapted based on the needs of individuals (Bower & Gilbody, 2005). In addition, stepped care models are evidence-based and have been shown to be cost-effective alternatives to conventional mental health care (Haugh et al., 2019).

In this regard, considering the high prevalence of EDs in primary care and in line with clinical guidelines (Díaz González, 2008), a cost-effective approach would be to intensify psychological interventions as the individual’s distress and symptom severity increase. This approach requires the use of tools to assess symptom severity, such as the ODSIS and OASIS scales. To our knowledge, there are no studies that have analyzed the usefulness of the ODSIS and OASIS as screening tools in the Spanish public health system.

Therefore, the aim of the present study is: (1) to explore the discriminatory capacity of the ODSIS and OASIS measures in individuals with a diagnosis of EDs referred by FPs from PC settings and treated in specialized MHUs in Spain, and: (2) to analyze whether the differential severity groups created through the pre-treatment scores of these measures can predict the evolution of symptomatology after receiving psychological treatment. Finally, we expect to find that the cut-off points established to differentiate the emotional severity of the participants’ symptoms, based on their pre-treatment severity/interference scores, will allow us to create heterogeneous groups with different characteristics and evolution throughout the intervention.

## Method

### Participants

The sample of this study consisted of 256 participants, all of them met the study criteria (see Osma et al., 2018), being outpatients referred to different Spanish MHUs from PC settings and with a main diagnosis of ED. Most of the participants were women (76.2%), with a mean age of 41.93 years ( $DT=13.22$ , range 18–77). The remaining socio-demographic data can be found in Table 1, and the flowchart of the participants throughout the intervention in Fig. 1.

### Instruments

#### Primary outcomes

Socio-demographic data questionnaire. It includes information on gender, educational level, marital status, job status, primary, and secondary psychological diagnoses (ADIS-IV diagnostic interview; Di Nardo et al., 1994), and number of treatment sessions received.

Depressive symptomatology was assessed through the ODSIS (Bentley et al., 2014; Osma et al., 2019) which consists of five items with a 5-point Likert-type response scale (0=Little/Not at all; 4=Extreme), and the Beck Depression Inventory (BDI-II; Beck et al., 1996; Sanz et al., 2005) which consists of 21 items and uses a 4-point Likert-type response scale (0=Low severity of symptoms; 3=High severity of symptom). The internal consistency of the measures’ scores was adequate, with a Cronbach’s alpha of 0.95 and 0.91, respectively.

Anxious symptomatology was assessed through the OASIS (Norman et al., 2006; Osma et al., 2019) which consists of five items with a 5-point Likert-type response scale (0=Little/Not at all; 4=Extreme), and the Beck Anxiety Inventory (BAI; Beck & Steer, 1993; Sanz et al., 2012) which consists of 21 items and uses a 4-point Likert-type response scale (0=Low severity of symptoms; 3=High severity of symptom). The internal consistency of the measures’ scores was adequate, with a Cronbach’s alpha of 0.89 and 0.92, respectively.

#### Secondary outcomes

Neuroticism and extraversion personality dimensions were assessed through the 12 items for each dimension of the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1999), which uses a 5-point Likert-type response scale (0=strongly disagree; 4=strongly agree). The instrument scores obtained an adequate internal consistency with a

**Table 1** Socio-demographic characteristics of the participants ( $N=256$ )

	Group 1 ( $n=33$ )	Group 2 ( $n=93$ )	Group 3 ( $n=130$ )	Comparison		Total ( $n=256$ )
	$n$ (%)	$n$ (%)	$n$ (%)	$\chi^2$	$p$	$n$ (%)
Gender				0.37	0.831	
Women	25 (75.8)	69 (74.2)	101 (77.7)			195 (76.2)
Men	8 (24.2)	24 (24.8)	29 (22.3)			61 (23.8)
Educational level				19.84	0.011	
More than 12 years of education	18 (54.5)	60 (64.5)	76 (58.5)			154 (60.2)
University studies	6 (18.2)	26 (28.0)	29 (22.3)			61 (23.8)
Vocational training	9 (27.3)	20 (21.5)	35 (26.9)			64 (25.0)
High school	3 (9.1)	14 (15.1)	12 (9.2)			29 (11.3)
Less than 12 years of education	15 (45.5)	33 (35.5)	54 (41.5)			102 (39.8)
Primary studies or less	3 (9.1)	17 (18.3)	39 (30.0)			59 (23.0)
Secondary studies	12 (36.4)	16 (17.2)	15 (11.5)			43 (16.8)
Marital status				9.45	0.150	
Married/living with partner	14 (42.4)	47 (50.5)	71 (54.6)			132 (51.6)
Not Married/not living with partner	19 (57.6)	46 (49.5)	59 (45.4)			124 (48.4)
Single	17 (51.5)	33 (35.5)	40 (30.8)			90 (35.2)
Separated/ Divorced	1 (3.0)	13 (14.0)	15 (11.5)			29 (11.3)
Widowed	1 (3.0)	0 (0.0)	4 (3.1)			5 (2.0)
Job status				15.46	0.116	
Not working	15 (45.5)	47 (50.5)	83 (63.8)			145 (48.4)
Unemployed	7 (21.2)	23 (24.7)	31 (23.8)			61 (23.8)
Sick leave	3 (9.1)	11 (11.8)	30 (23.1)			44 (17.2)
Student	1 (3.0)	8 (8.6)	10 (7.7)			19 (7.9)
Home-maker	3 (9.1)	1 (1.1)	9 (6.9)			13 (5.1)
Retired	1 (3.0)	4 (4.3)	3 (2.3)			8 (3.1)
Working	18 (54.5)	46 (49.5)	47 (36.2)			111 (43.4)
Primary diagnoses				46.75	0.004	
Anxiety and related disorders	15 (45.5)	46 (49.5)	51 (39.2)			117 (45.7)
Generalized anxiety disorder	2 (6.1)	13 (14.0)	14 (10.8)			29 (11.3)
Panic disorder with agoraphobia	1 (3.0)	8 (8.6)	11 (8.5)			20 (7.8)
Non-specific anxiety disorder	2 (6.1)	10 (10.8)	8 (6.2)			20 (7.8)
Panic disorder without agoraphobia	2 (6.1)	7 (7.5)	5 (3.8)			14 (5.5)
Agoraphobia	2 (6.1)	0 (0.0)	5 (3.8)			7 (2.7)
Obsessive-compulsive disorder	2 (6.1)	2 (2.2)	7 (5.4)			11 (4.3)
Hypochondria	2 (6.1)	2 (2.2)	1 (0.8)			5 (2.0)
Social anxiety	2 (6.1)	4 (4.3)	0 (0.0)			6 (2.3)
Posttraumatic stress disorder	0 (0.0)	0 (0.0)	5 (3.8)			5 (2.0)
Mood disorders	7 (21.2)	21 (22.6)	43 (33.1)			71 (27.7)
Major depressive disorder	2 (6.1)	17 (18.3)	31 (23.8)			50 (19.5)
Dysthymia	1 (3.0)	3 (3.2)	10 (7.7)			14 (5.5)
Unspecified mood disorder	4 (12.1)	1 (1.1)	2 (1.5)			7 (2.7)
Mixed disorders	11 (33.3)	26 (28.0)	31 (23.8)			68 (26.6)
Adjustment disorder	11 (33.3)	26 (28.0)	31 (23.8)			68 (26.6)
Diagnostic comorbidity				1.37	0.505	
Yes	4 (12.1)	24 (25.8)	33 (25.4)			61 (23.8)
No	29 (87.9)	69 (74.2)	97 (74.6)			195 (76.2)
Secondary diagnoses				41.04	0.339	
Anxiety and related disorders	2 (1.5)	17 (13.1)	16 (12.3)			35 (26.9)
Non-specific anxiety disorder	1 (3.0)	5 (5.4)	7 (5.4)			13 (5.1)
Generalized anxiety disorder	0 (0.0)	2 (2.2)	3 (2.3)			5 (2.0)
Agoraphobia	0 (0.0)	2 (2.2)	2 (1.5)			4 (1.6)
Obsessive-compulsive disorder	1 (3.0)	3 (3.2)	1 (0.8)			5 (2.0)
Panic disorder without agoraphobia	0 (0.0)	1 (1.1)	3 (2.3)			4 (1.6)
Social anxiety	0 (0.0)	1 (1.1)	0 (0.0)			1 (0.4)

**Table 1** (continued)

	Group 1 ( <i>n</i> =33)	Group 2 ( <i>n</i> =93)	Group 3 ( <i>n</i> =130)	Comparison		Total ( <i>n</i> =256)
	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	$\chi^2$	<i>p</i>	<i>n</i> (%)
Hypochondria	0 (0.0)	1 (1.1)	0 (0.0)			2 (0.8)
Panic disorder with agoraphobia	0 (0.0)	1 (1.1)	0 (0.0)			1 (0.4)
Posttraumatic stress disorder	0 (0.0)	1 (1.1)	0 (0.0)			1 (0.4)
Mood disorders	2 (6.1)	2 (1.5)	16 (12.3)			20 (7.8)
Major depressive disorder	0 (0.0)	1 (1.1)	12 (9.2)			13 (5.1)
Dysthymia	2 (6.1)	1 (1.1)	3 (2.3)			6 (2.3)
Unspecified mood disorder	0 (0.0)	0 (0.0)	1 (0.8)			1 (0.4)
Mixed disorders	0 (0.0)	5 (5.4)	1 (0.8)			6 (2.3)
Adjustment disorder	0 (0.0)	5 (5.4)	1 (0.8)			6 (2.3)

Group 1: ODSIS/OASIS scores 0–6; Group 2: ODSIS/OASIS scores 7–12; Group 3: ODSIS/OASIS scores > 13

Cronbach's alpha value of 0.78 for neuroticism and 0.79 for extraversion.

Quality of life was assessed through the 10 items of the Quality of Life Index (QLI; Mezzich et al., 2000), which uses a 11-point Likert-type response scale (0=poor; 10=excellent). Instrument scores showed adequate psychometric properties with Cronbach's alpha of 0.88.

Maladjustment Inventory (MI; Echeburúa et al., 2000; Osma et al., 2024). Interference was assessed through the six items of the MI, which use a 6-point Likert scale (0=Not at all; 5=Extremely). Instrument scores showed adequate psychometric properties with Cronbach's alpha of 0.85.

Difficulties in Emotional Regulation Questionnaire (DERS; Gratz & Roemer, 2004; Hervás & Jódar, 2008). Emotional dysregulation was assessed through the 28 items of the DERS instrument, which employs a 5-point Likert-type response scale (1=Almost never; 5=Almost always). Instrument scores showed adequate psychometric properties with Cronbach's alpha of 0.91.

## Procedure

This study is a secondary analysis of a main study, which consisted of a randomised multicentre clinical trial that aimed to analyze the efficacy and cost-effectiveness of a transdiagnostic intervention, the Unified Protocol for the Transdiagnostic Treatment of Emotional Disorders (UP; Barlow et al., 2018), applied in group compared to the treatment as usual (TAU) condition (disorder-specific non-structured cognitive behavioral treatment, including techniques such as psychoeducation, cognitive flexibility, cognitive techniques, relaxation strategies, among others), delivered in an individual format with a typical session duration of 30–40 min and the usual frequency of the care unit (approximately one session every month or month and a half) (Osma et al., 2018). Specifically, those participants who were assigned to the TAU condition in different Spanish MHUs were those who took part in this study. This is a cohort study, where

participants were retrospectively divided based on their initial (pre-treatment) scores on the severity and interference of depressive and anxious symptomatology assessed through the ODSIS and OASIS. For this purpose, we used half a standard deviation above and below the clinical cut-off point (cut-off point=10) established in the validation of the measures in Spanish (mean=0.88, *SD*=5.14 for ODSIS and mean=10.45, *SD*=4.49 for OASIS, Osma et al., 2019), with the aim of establishing three differentiated groups, as can be seen in appendix 1. Group 1 would present a mild initial severity, with ODSIS and OASIS scores between 0 and 6; Group 2 would present a medium-moderate initial severity with scores between 7 and 12 and, finally, Group 3 would present a moderate-severe initial severity with scores equal to or higher than 13. The choice of these questionnaires as screening measures was due to their brevity (five items), simplicity of structure (unidimensional), and correction (sum of the items), which favors their use in contexts with limited resources such as PC settings.

In addition, evaluations were carried out at pre-treatment, and at 3-, 6-, 9- and 15-month follow-up. Information regarding the setting, sampling size, the period of recruitment and data collection can be found in Osma et al. (2018).

Finally, all participants signed the informed consent form, and the project was approved by the research ethics committees of the collaborating centers (Osma et al., 2018). This study follows the STROBE guidelines.

## Data analysis

First, descriptive statistical analyses were carried out to analyze the sociodemographic characteristics of the participants. Then, analyses of variance (ANOVA) were carried out to analyze the differences in scores between the three comparison groups. Specifically, since the group sizes are different and there is no heterogeneity of variances, a Welch ANOVA test was performed, which is less sensitive to unequal group sizes.

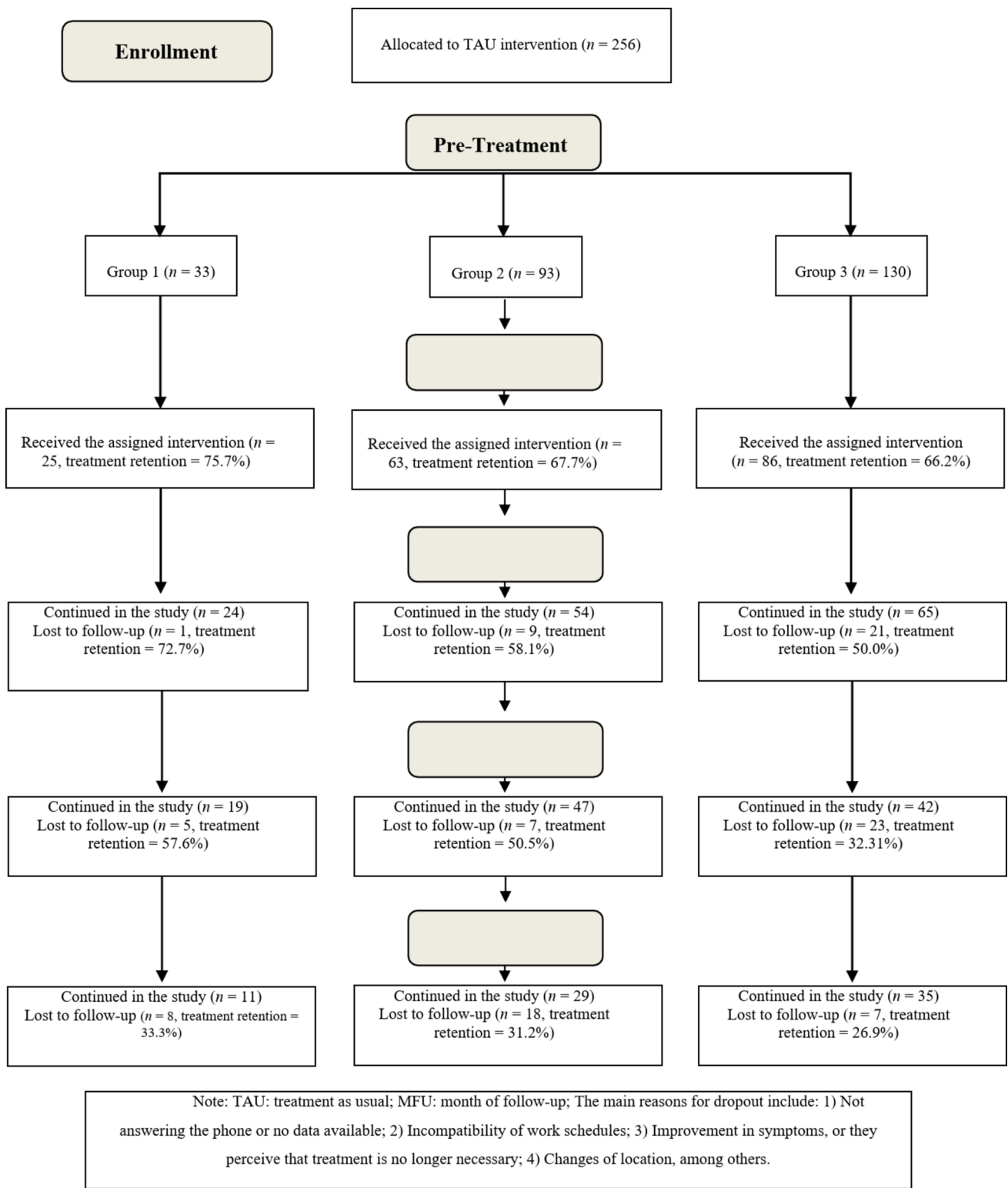


Fig. 1 Flow chart of study participants according to the consort guidelines

Similarly, comparisons were carried out using the chi-square test to analyze differences between groups on categorical variables, specifically including differences in sociodemographic data. Finally, linear mixed models were performed. Specifically, the linear mixed models included “Time” (intrasubject: T1, T2, T3, T4, T5) and “Group” (intersubject: Group 1, Group 2, Group 3) as main effects, as well as the interactions “Time\*Group” and “Time\*Group\*Sessions”, with the aim of analyzing the evolution of scores over time as a function of group and number of sessions received. Linear mixed models were used in the analysis to handle missing data, as these models are more reliable and efficient than traditional analysis of variance, especially when analyzing data from repeated measurements taken over time on the same participants (Baayen et al., 2008). Linear mixed models are multilevel models that allow for the study of individual-level trends, rather than estimating a single average model for all participants, as is done in repeated measures analysis of variance, which leads to a loss of variability. These models also employ more efficient estimation methods and are more flexible in handling missing data, a common issue in studies with multiple time points (Verbeke & Molenberghs, 2000).

Effect sizes were also calculated using Cohen’s *d* statistic, interpreted as small ( $d=0.2$ ), medium ( $d=0.5$ ), and large ( $d=0.8$ ) and eta squared ( $\eta^2$ ), interpreted as small ( $\eta^2 = 0.01$ ), medium ( $\eta^2 = 0.06$ ), and large ( $\eta^2 = 0.14$ ) (Cohen, 1988). All statistical analyses were carried out using SPSS software (version 25.0) (IBM Corp, 2017).

## Results

### Results of comparison between groups

The results showed statistically significant differences ( $p < .05$ ) between the three groups in the primary diagnoses, with a greater number of anxiety disorders among participants in groups 1 and 2, and also differences in all study variables at all assessment times (except for the DERS measure at the 9-month follow-up:  $F=4.67$ ,  $p=.059$ , and 15-month follow-up:  $F=2.20$ ,  $p=.129$ , as can be seen in Table 2). The effect sizes of the differences were moderate-to-large ( $\eta^2$  values ranging from 0.42 to 0.95), tending to decrease over time and with the smallest effect sizes found in the variables of extraversion ( $\eta^2$  values ranging from 0.42 to 0.57) and DERS ( $\eta^2$  values ranging from 0.50 to 0.64), which implies large differences in the severity of symptomatology between the groups.

The highest scores on anxious and depressive symptomatology, and secondary variables such as neuroticism, interference or emotional dysregulation were reported by

group 3, followed by group 2, while the lowest scores were reported by group 1. Regarding the number of sessions received, no statistically significant differences were found between the groups at any evaluation time ( $p > .05$ ). Finally, the treatment adherence of each group can be seen in Fig. 1. The results showed statistically significant differences in adherence between the groups at the 9-month follow-up ( $\chi^2=11.06$ ,  $p=.004$ ), with group 3 having the lowest treatment retention (32.31%).

### Scores’ evolution over time

The linear mixed models showed a main effect of time ( $p < .05$ ) in BDI-II, BAI, neuroticism, QLI and MI, as can be seen in Table 3. In addition, a main effect of the variable “group” was also observed in all the variables under study. As for the number of sessions, the results showed no significant effect ( $p > .05$ ). Finally, the results showed Time\*Group interaction effects on the variables ODSIS and neuroticism, indicating different evolutions of these variables over time depending on the condition. Specifically, the largest changes in the ODSIS variable were observed in groups 3 and 2 ( $F=18.28$ ,  $p < .001$ , *Cohen’s d*=0.58, and  $F=2.76$ ,  $p=.029$ , *Cohen’s d*=0.34, respectively), with medium effect sizes, while no changes were observed in group 1 ( $F=2.41$ ,  $p < .001$ , *Cohen’s d*=0.31). Regarding the neuroticism variable, statistically significant changes were found in all three groups (group 1:  $F=3.39$ ,  $p=.014$ , *Cohen’s d*=0.37; group 2:  $F=11.67$ ,  $p < .001$ , *Cohen’s d*=0.70; group 3:  $F=8.24$ ,  $p < .001$ , *Cohen’s d*=0.59), with large effect sizes in group 3, followed by group 2, and medium effect sizes in group 1.

In terms of the main effects of time according to each specific group, the results showed a statistically significant main effect of time in group 1 with small effect sizes for BDI ( $F=5.48$ ,  $p=.001$ , *Cohen’s d*=0.21), neuroticism ( $F=3.40$ ,  $p < .001$ , *Cohen’s d*=0.16) and MI ( $F=3.32$ ,  $p=.019$ , *Cohen’s d*=0.16). As for group 2, statistically significant effects were found with small-medium effect sizes in BDI ( $F=21.46$ ,  $p < .001$ , *Cohen’s d*=0.41), ODSIS ( $F=2.76$ ,  $p=.029$ , *Cohen’s d*=0.15), BAI ( $F=15.66$ ,  $p < .001$ , *Cohen’s d*=0.35), OASIS ( $F=7.54$ ,  $p < .001$ , *Cohen’s d*=0.24) Neuroticism ( $F=11.67$ ,  $p < .001$ , *Cohen’s d*=0.30), QLI ( $F=9.20$ ,  $p < .001$ , *Cohen’s d*=0.27), MI ( $F=6.97$ ,  $p < .001$ , *Cohen’s d*=0.23) and DERS ( $F=10.28$ ,  $p < .001$ , *Cohen’s d*=0.28). Finally, in group 3, effects were found with small-medium effect sizes for the variables BDI ( $F=15.30$ ,  $p < .001$ , *Cohen’s d*=0.35), ODSIS ( $F=18.25$ ,  $p < .001$ , *Cohen’s d*=0.38), BAI ( $F=7.91$ ,  $p < .001$ , *Cohen’s d*=0.25), OASIS ( $F=13.27$ ,  $p < .001$ , *Cohen’s d*=0.32) Neuroticism ( $F=8.24$ ,  $p < .001$ , *Cohen’s d*=0.25), QLI ( $F=10.29$ ,  $p < .001$ , *Cohen’s d*=0.28), MI ( $F=9.86$ ,  $p < .001$ , *Cohen’s d*=0.28) and DERS ( $F=3.17$ ,  $p=.016$ , *Cohen’s d*=0.16).

**Table 2** Means, standard deviations and Welch ANOVA test between groups results (N=256)

Screening groups	Normative data†			Pre-T			3-MFU			6-MFU			9-MFU			15-MFU			
	M (SD)	M (SD)	F	p	$\eta^2$	M (SD)	F	p	$\eta^2$	M (SD)	F	p	$\eta^2$	M (SD)	F	p	$\eta^2$		
Primary Outcomes	Group 1	13.32 (6.42)	9.4 (7.7)	<b>87.52</b>	<0.001	0.86	6.93 (7.24)	<b>58.27</b>	<0.001	0.77	6.31 (5.28)	<b>33.70</b>	<0.001	0.77	4.58 (5.87)	<b>24.70</b>	<0.001	0.74	5.29 (4.46)
	Group 2	24.63 (9.46)					14.79 (9.83)				12.78 (8.66)				11.57 (8.67)				12.38 (8.95)
	Group 3	34.55 (11.1)					26.59 (13.87)				24.68 (15.31)				24.45 (16.07)				21.69 (13.53)
ODSIS	Group 1	2.21 (1.99)	2.79 (4.06)	<b>334.87</b>	<0.001	0.93	1.94 (2.91)	<b>64.07</b>	<0.001	0.86	2.85 (4.13)	<b>32.23</b>	<0.001	0.83	4.22 (4.15)	<b>9.34</b>	<0.001	0.76	1.91 (1.86)
	Group 2	7.02 (3.78)					6.59 (5.50)				5.75 (4.72)				6.16 (4.49)				5.27 (5.10)
	Group 3	13.96 (3.36)					11.09 (5.05)				11.12 (5.07)				9.95 (5.72)				9.56 (6.41)
BAI	Group 1	16.16 (12.65)	9.6 (7.8)	<b>38.32</b>	<0.001	0.72	8.93 (10.13)	<b>33.09</b>	<0.001	0.60	6.05 (5.63)	<b>22.58</b>	<0.001	0.59	5.83 (6.62)	<b>13.10</b>	<0.001	0.55	4.86 (5.08)
	Group 2	23.49 (11.46)					15.90 (10.26)				12.95 (8.52)				14.26 (11.18)				12.59 (10.92)
	Group 3	33.25 (11.9)					26.48 (14.39)				25.14 (15.76)				22.61 (16.04)				23.29 (14.3)
OASIS	Group 1	2.65 (1.85)	3.92 (4.13)	<b>360.49</b>	<0.001	0.95	2.52 (3.35)	<b>50.70</b>	<0.001	0.88	3.18 (3.78)	<b>29.28</b>	<0.001	0.87	4.50 (3.75)	<b>13.70</b>	<0.001	0.80	2.82 (2.44)
	Group 2	9.18 (2.29)					7.70 (3.93)				7.58 (4.30)				7.45 (4.03)				5.88 (4.03)
	Group 3	13.92 (2.92)					11.10 (4.41)				11.87 (5.99)				10.95 (5.33)				10.02 (6.18)

Table 2 (continued)

Screening groups	Normative data†			Pre-T			3-MFU			6-MFU			9-MFU			15-MFU		
	M(SD)	M(SD)	F	p	$\eta^2$	M(SD)	F	p	$\eta^2$	M(SD)	F	p	$\eta^2$	M(SD)	F	p	$\eta^2$	
Secondary Neuroticism	26.63 (9.15)	20.53 (7.46)	<b>41.61</b>	<0.001	0.76	19.8 (6.72)	<b>23.98</b>	<0.001	0.56	18.72 (6.85)	<b>23.35</b>	<0.001	0.57	18.00 (9.12)	<b>12.72</b>	<0.001	0.51	22.00 (6.63)
Group 1	32.34 (6.13)	29.15 (6.58)				29.15 (6.58)				27.24 (7.91)				25.74 (7.45)				25.68 (8.15)
Group 2	34.86 (6.67)	32.47 (7.32)				32.47 (7.32)				30.63 (8.83)				29.86 (9.52)				28.2 (8.21)
Group 3	28.88 (6.46)	31.72 (6.74)	<b>18.31</b>	<0.001	0.57	29.67 (6.56)	<b>13.04</b>	<0.001	0.44	30.81 (5.27)	<b>12.32</b>	<0.001	0.48	32.33 (4.23)	<b>5.52</b>	<b>0.007</b>	0.47	27.57 (5.88)
Extra-version	22.57 (7.71)					25.35 (7.34)				26.05 (8.01)				26.42 (7.46)				26.29 (6.49)
Group 1	19.19 (8.07)					20.79 (9.14)				21.20 (9.44)				21.82 (9.85)				22.01 (10.03)
Group 2	6.14 (1.17)	6.98 (1.11)	<b>79.60</b>	<0.001	0.87	7.36 (1.46)	<b>33.57</b>	<0.001	0.83	7.52 (0.94)	<b>26.22</b>	<0.001	0.79	7.99 (1.16)	<b>15.07</b>	<0.001	0.72	7.96 (0.73)
Group 1	4.89 (1.4)					5.88 (1.49)				6.25 (1.34)				6.49 (1.31)				6.42 (1.38)
Group 2	3.78 (1.47)					4.73 (1.9)				4.91 (1.84)				5.20 (2.11)				5.28 (2.03)
Group 3	12.23 (6.42)	8.56 (6.98)	<b>68.19</b>	<0.001	0.84	6.29 (7.09)	<b>37.44</b>	<0.001	0.82	5.25 (5.78)	<b>12.44</b>	<0.001	0.81	5.00 (4.51)	<b>5.41</b>	<b>0.013</b>	0.75	4.40 (5.13)
MI	17.87 (5.72)					12.68 (6.82)				8.91 (7.25)				8.08 (7.01)				9.77 (7.15)
Group 1	21.48 (5.42)					18.30 (6.38)				16.74 (7.61)				16.4 (9.12)				15.78 (8.5)
Group 2	64.56 (13.65)	58.40 (17.60)	<b>23.87</b>	<0.001	0.64	51.00 (12.92)	<b>9.08</b>	<b>0.002</b>	0.58	50.33 (11.06)	<b>9.47</b>	<b>0.002</b>	0.62	58.00 (11.75)	<b>4.67</b>	0.059	0.57	46.00 (15.62)
Group 1	80.47 (16.36)					67.35 (15.77)				65.88 (20.88)				63.86 (19.66)				59.05 (17.89)
Group 2	94.68 (19.25)					82.57 (23.49)				78.53 (24.93)				79.23 (26.11)				76.40 (22.17)
Group 3	-	-	-	-	-	2.52 (1.99)	0.13	0.876	3.65 (1.69)	0.27	0.763	5.07 (2.01)	0.27	7.08 (2.63)	0.088	0.923		
Number of Sessions	-	-	-	-	-	2.17 (1.71)	-	-	3.72 (1.75)	-	-	5.22 (2.28)	-	7.12 (3.37)	-	-	-	
Group 1	-	-	-	-	-	2.43 (2.07)	-	-	4.06 (1.98)	-	-	5.63 (2.44)	-	7.37 (2.70)	-	-	-	

Group 1: ODSIS/OASIS scores 0–6; Group 2: ODSIS/OASIS scores 7–12; Group 3: ODSIS/OASIS scores > 13; MFU: month of follow-up; BDI-II: Beck Depression Inventory; ODSIS: Overall Depression Severity and Impairment Scale; BAI: Beck Anxiety Inventory; OASIS: Overall Anxiety Severity and Impairment Scale; QLI: Quality of Life Index; In bold p-values < 0.05. †Normative data have been extracted from: BDI: Sanz et al. (2003); BAI: Sanz and Navarro (2003); ODSIS: Osma et al. (2021); OASIS: Osma et al. (2021); NEO-FFI: Costa and McCrae (1999); QLI: Lorente et al. (2002); MI: Osma et al., 2024, and DERS: Hervás and Jódar (2008)

No main effects of time on extraversion were observed in either group ( $p > .05$ ).

## Discussion

The aim of this study was to explore the discriminatory capacity of the ODSIS and OASIS measures in individuals with a diagnosis of EDs referred from PC settings to be treated in specialized MHUs, and to analyze whether they presented a different evolution of symptomatology after receiving psychological treatment based on the severity/interference groups created through the pre-treatment scores of these measures. Our hypothesis was that the proposed cut-off points based on the initial severity and interference of the participants would allow for the creation of heterogeneous groups. The results obtained in this study have supported this hypothesis, specifically, the data show how the three groups present statistically significant differences in all the variables assessed; depression (BDI and ODSIS), anxiety (BAI and OASIS), temperament (neuroticism and extraversion), quality of life (QLI), interference (MI), and emotional regulation (DERS) at all times assessed; pre-treatment, at 3-, 6-, 9- and 15-month follow-up (except in DERS at 15-month follow-up). These results show that the initial severity/interference with which individuals started in each of the groups will determine the course of their symptomatology throughout the psychological intervention, with these differences remaining stable between groups over time.

In this line, individuals in groups 1 and 2 started with non-clinical symptomatology (according to ODSIS and OASIS) in depression and anxiety and, after the intervention, the scores were reduced. On the other hand, participants in group 3 began with clinical symptomatology in depression and anxiety and, after the intervention, they showed non-clinical symptomatology scores in depression ( $M=9.56$ ) and clinical symptomatology scores in anxiety ( $M=10.02$ ). Overall, these results show that the differences between the three groups at the beginning of the intervention remain constant until the end of the intervention. These results seem to indicate that it is necessary to adjust the care services offered to individuals according to the symptomatology and interference presented by each individual.

In the case of group 1, the symptomatology was milder and anxiety predominated, in accordance with the high prevalence of anxiety cases (12.69%) seen in PC settings (Ministerio de Sanidad, 2022). Previous literature has indicated that mild anxiety symptoms can precede more serious clinical psychological symptoms (Fava et al., 2000; Wigman et al., 2012). Therefore, and according to the recommendations of clinical guidelines, these individuals could

benefit from short care interventions in PC settings, such as follow-up appointments or counselling (Díaz González, 2008). This would help to reduce the burden of care in specialized mental health care.

Participants in group 2 were characterized by an initial profile with moderate anxiety and depression symptoms that improved to mild after treatment, which could indicate that this profile of individuals requires a more extended and structured treatment than participants in group 1, in accordance with the recommendations of clinical guidelines in PC settings (Díaz González, 2008). Given the characteristics of these participants, short intensive psychological interventions could be offered from PC settings.

Finally, in the case of group 3, the participants presented severe symptoms of anxiety and depression that improved after treatment, reaching non-clinical scores in the case of depressive symptoms; however, in the case of anxiety, they continued to present moderate clinical symptoms. This could indicate that even though the treatment received by this group of individuals helped reduce the severity of their symptoms, therapeutic work is still necessary in order to further reduce their symptoms and interference. Considering that, 15 months after the start of the treatment, they had all only received an average of seven sessions, it may be necessary to increase the frequency with which they receive sessions, through more intensive intervention formats that would allow to reduce the intervention time; for example, through group formats that would allow professionals to treat a larger number of people, in a shorter period of time (Norton, 2012).

These findings support the stepped care approach, highlighting the importance of intensifying psychological interventions as the emotional symptomatology of individuals becomes more severe (Bower & Gilbody, 2005). This underscores the need to adapt the intensity and duration of treatment in response to symptom severity.

In terms of changes after the psychological intervention, the data show that, over time, all individuals achieved improvements in depression, anxiety, neuroticism, quality of life, and interference. However, there were still differences between the three groups.

Regarding the evolution of the scores, the results showed a statistically different evolution over time in the variables ODSIS and neuroticism. Specifically, changes in ODSIS were only observed in groups 2 and 3, with participants in group 3 obtaining the greatest changes. On the other hand, no statistically significant changes were observed in the participants of group 1. These results suggest that, after the psychological intervention, the individuals who benefited most were those who presented more severe symptomatology, especially greater interference related to depressive symptomatology. These findings are consistent with the stepped

**Table 3** Main effects of the linear mixed models (N=256)

	Secondary Outcomes																											
	Primary Outcomes				OASIS				Neuroticism				Extraversion				QLI				MI				DERS			
	BDI-II		ODSIS		BAI		F		F		F		F		F		F		F		F		F		F		F	
Time	7.61	<0.001	0.24	1.58	0.179	0.11	2.79	0.026	0.15	1.92	0.105	0.12	3.74	0.005	0.17	0.28	0.887	0.05	7.53	<0.001	0.24	5.24	<0.001	0.20	0.36	0.837	0.05	
Group	25.27	<0.001	0.44	18.99	<0.001	0.38	10.14	<0.001	0.28	26.72	<0.001	0.46	13.45	<0.001	0.32	10.06	<0.001	0.28	30.56	<0.001	0.49	11.27	<0.001	0.30	5.38	0.005	0.20	
Sessions	1.17	0.280	0.09	2.34	0.124	0.13	1.21	0.272	0.10	0.58	0.447	0.07	0.66	0.416	0.07	0.02	0.881	0.04	2.31	0.129	0.13	1.30	0.254	0.10	0.11	0.745	0.03	
Time*Group	0.87	0.539	0.08	3.09	0.002	0.16	0.72	0.675	0.07	1.42	0.185	0.11	2.16	0.029	0.13	0.79	0.610	0.08	1.14	0.333	0.09	0.83	0.574	0.08	1.73	0.102	0.11	
Time*Group*Sessions	1.19	0.300	0.09	1.34	0.222	0.10	0.56	0.814	0.07	0.72	0.674	0.07	1.34	0.223	0.10	0.84	0.568	0.08	1.63	0.112	0.11	0.50	0.859	0.06	1.17	0.319	0.09	

d: Cohen's d BDI-II: Beck Depression Inventory; ODSIS: Overall Depression Severity and Impairment Scale; BAI: Beck Anxiety Inventory; OASIS: Overall Anxiety Severity and Impairment Scale; QLI: Quality of Life Index; In bold *p*-values < 0.05

care model and align with primary care clinical guidelines, which recommend more intensive and structured treatment for individuals with more severe symptomatology, while possibly suggesting a shorter duration and fewer sessions for those with milder cases (Díaz González, 2008).

Regarding neuroticism, the results were similar, with the greatest changes being observed in groups 2 and 3. However, neuroticism scores at 15-month follow-up in these groups were higher than the normative mean.

Furthermore, in the case of extraversion, there were no statistically significant changes in either group after the intervention. A similar result was found with difficulties in emotional regulation. Although participants in groups 2 and 3 seemed to show a trend towards improvement, the results did not show statistically significant changes over time. This is especially important, since emotion dysregulation has been associated with the etiology and maintenance of EDs (Antuña-Cambor et al., 2024). Based on this outcome, the TAU offered in MHUs in Spain, this is a non-structured psychological intervention for specific disorders, and the frequency of sessions provided must be reconsidered. In this regard, the use of emotion regulation based psychological interventions could be a solution to improve emotion regulation strategies in individuals with EDs such as the Unified Protocol for transdiagnostic treatment of EDs (UP; Barlow et al., 2018). This cognitive-behavioral intervention has been already applied in the Spanish mental health system with good cost-efficacy outcomes, both in PC settings (Corpas et al., 2021) and MHUs (Peris-Baquero et al., 2022; Peris-Baquero & Osma, 2023). We must consider also the positive outcomes of improving neuroticism but also extraversion scores after applying the UP (Osma et al., 2021; Sauer-Zavala et al., 2021) making it a good alternative in comparison with disorder-specific interventions commonly used in these settings.

Regarding the interference and quality of life variables, groups 2 and especially 3, were the ones that improved the most. Group 3 presented great severity and severe interference at the beginning of treatment, and after treatment, achieved greater improvements. These results show that these groups with more severe symptoms require and benefit to a greater extent from a more structured psychological intervention, in line with the recommendations of clinical guidelines (Díaz González, 2008).

Finally, the results have shown that there were no differences in the number of sessions received between the groups, and that the number of sessions did not have a main effect or interaction effect on the results of the psychological intervention. Therefore, these results highlight that it may not be necessary for everyone to receive the same number of psychological sessions. The number of sessions has not been shown to be related to the evolution of the scores, therefore,

in order to be more cost-effective and in line with the step-care model and personalization of the treatments (Haugh et al., 2019), we should adjust the number of sessions and the treatment offered to individuals according to the severity and interference of their symptom and interference. In this way, people could be referred to the most appropriate care resource given their symptomatology and interference, as shown in Fig. 2 (Appendix 1). This referral proposal would be in line with the clinical guidelines that establish three levels of intervention according to severity symptoms (Díaz González, 2008).

In view of all the above, a possible solution to improve the efficiency of the mental health resources and the effectiveness of the psychological interventions provided to those in need may be to establish a clear referral process based on the severity of the emotional symptom of the individuals treated in PC, the first step of public health care systems. In this regard, the findings of this study highlight that ODSIS and OASIS, in addition to assessing the evolution of emotional symptomatology as used in previous studies (Eustis et al., 2020; Timulak et al., 2022), are also able to discriminate the severity of the symptom of individuals. These measures are characterized by being brief, simple to complete, and easy to correct and interpret. This would allow FPs working in PC to quickly assess emotional symptoms and facilitate their clinical decision making based on the established scores, thus facilitating referral of individuals to the appropriate care resources based on the severity of the symptoms.

## Limitations

This study is not without limitations; the first limitation relates to the context in which the study was carried out, the specialized mental health services in Spain, which influence the generalization of the results to other countries and contexts, due to differences in health policies, the structure of the public healthcare system, available human and material resources, and even the training and therapeutic approaches of professionals working within the Spanish healthcare system. In addition, although MHUs form part of the national health service, they have different resources and function differently to PC services, which also affects the effectiveness of interventions (Finazzi & MacBeth, 2022). Another limitation is that the number of participants in each group is different, which may also affect the comparability of the groups as they are not homogeneous, and also impact the statistical power of the results. Finally, there may be a potential selection bias due to the retrospective grouping, and the lack of comparison with other instruments that have similar characteristics. Similarly, since these are self-administered questionnaires, the results may be biased due to factors such

as social desirability or difficulty in understanding some items, which could significantly affect the accuracy of the obtained results.

## Future research directions

In line with the limitations previously mentioned, future studies should involve a larger number of participants, with a similar number of participants in each group, and replicate the results of this study in a primary care (PC) context or in health systems of other countries. This will increase the representativeness and generalization of the results obtained in this study, allowing for a more precise analysis of the effectiveness of the ODSIS and OASIS instruments in evaluating and screening the severity of symptoms in EDs. Additionally, future studies should employ a prospective assignment, assigning participants to different groups prior to the intervention in order to minimize selection bias.

Nevertheless, the results of this study can serve as a guide for future studies in other contexts within the Spanish public health system and in the public health systems of other countries with similar characteristics. Furthermore, future studies should explore the discriminatory capacity of ODSIS and OASIS compared to other similar measures in this context (e.g., PHQ-9: Costantini et al., 2021; PHQ-4 or GAD-7: Muñoz-Navarro et al., 2017, among others) and other brief measures such as the SCL-K-9 that include other types of symptoms, such as obsessive-compulsive behaviors, hostility, paranoid tendencies, and psychoticism (Imperatori et al., 2020), which could also be suitable for use as screening and predictive measures in PC.

The use of brief and precise evaluation instruments helps reduce assessment time, thereby optimizing the resources available in public health systems. Additionally, these instruments can be integrated into clinical practice guidelines to automate the processes of assessment, screening, and referral in primary care. Finally, future studies should evaluate the cost-effectiveness of implementing these instruments, taking into account both the costs of training professionals and other associated expenses (e.g., paper printing costs), with the aim of determining whether their implementation is efficient for both the healthcare system and the users.

## Conclusions

There is an urgent need to improve the efficiency and effectiveness of the assessment and referral of individuals with mental health demands in the public mental health system in Spain. Psychological problems constitute a substantial volume of demands in PC, where FPs do not have the time

or tools to evaluate individuals effectively. This causes the referral of a huge number of individuals, regardless of the severity of their emotional symptom, to the MHUs, generating long waiting lists to be treated. The results of this study have shown that the ODSIS and OASIS allow an effective segregation of individuals according to the severity of their anxiety, depression, and interference symptoms and their distinct evolution over time while receiving a psychological treatment. Therefore, training family physicians in the effective use of these measures would be recommended for use in PC services. Knowing that, we can personalize more efficiently the referral process to different professionals and interventions; for example, those individuals with milder emotional symptoms could perhaps benefit from a non-psychological intervention, such as brief follow-up sessions conducted by the FP; those with medium-moderate symptomatology could be suitable for brief psychological interventions in PC conducted by psychologists, but these interventions could perhaps be optimized by reducing the frequency, time, and number of sessions. Finally, individuals with severe emotional symptomatology may probably need more intensive psychological treatment and more specialized assistance conducted in MHUs. These measures offer a brief and easy alternative for detecting EDs in PC, improving the referral process and reducing the burden of care in the MHU, thus optimizing the functioning of mental health services.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12144-025-07715-0>.

**Acknowledgements** We would like to thank all the individuals and professionals who have participated in the realization of this project.

**Author contributions** All authors contributed to the study conception and design. Material preparation and data collection were performed by Óscar Peris-Baquero, Laura Martínez-García and Jorge Osma. Analysis were performed by Óscar Peris-Baquero and Iván Prieto-Rollan. The first draft of the manuscript was written by Óscar Peris-Baquero and Alba Fadrique-Jiménez, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

This work was supported by the ISCIII-General Sub-Directorate for Evaluation and Promotion of Research of the European Regional Development Fund (ERDF), “A way of doing Europe” (Grant Number PI17/00320). Co-funding by Gobierno de Aragón (Department of Science, University and Society knowledge) [Grant Number Research team S31\_23R].

**Data availability** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Ethics approval** This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of all collaborating centers.

**Consent to participate** Informed consent was obtained from all individual participants included in the study.

**Competing interests** The authors have no financial or proprietary interests in any material discussed in this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aguilera-Martín, Á., Gálvez-Lara, M., Cuadrado, F., Moreno, E., García-Torres, F., Venceslá, J. F., Corpas, J., Jurado-González, F. J., Muñoz-Navarro, R., González-Blanch, C., Ruiz-Rodríguez, P., Barrio-Martínez, S., Prieto-Vila, M., Carpallo-González, M., Cano-Vindel, A., & Moriana, J. A. (2022). Cost-effectiveness and cost-utility evaluation of individual vs. group transdiagnostic psychological treatment for emotional disorders in primary care (PsicAP-Costs): A multicentre randomized controlled trial protocol. *Bmc Psychiatry*, *22*(1), 99. <https://doi.org/10.1186/s12888-022-03726-4>
- Antuña-Cambor, C., Gómez-Salas, F. J., Burgos-Julián, F. A., González-Vázquez, A., Juarros-Basterretxea, J., & Rodríguez-Díaz, F. J. (2024). Emotional regulation as a transdiagnostic process of emotional disorders in therapy: A systematic review and Meta-Analysis. *Clinical Psychology & Psychotherapy*, *31*(3), e2997. <https://doi.org/10.1002/cpp.2997>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.
- Barlow, D. H., Farchione, T. J., Sauer-Zavala, S., Latin, H. M., Ellard, K. K., Bullis, J. R., Bentley, K. H., Boettcher, H. T., & Cassiello-Robbins, C. (2018). *Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide*. Oxford University Press.
- Barry, M. J., Nicholson, W. K., Silverstein, M., Coker, T. R., Davidson, K. W., Davis, E. M., Donahue, K. E., Jaén, C. R., Li, L., & Ogedegbe, G. (2023). Screening for anxiety disorders in adults: US preventive services task force recommendation statement. *Journal of the American Medical Association*, *329*(24), 2163–2170.
- Base de Datos Clínicos de Atención Primaria [BDCAP] (2022). *Problemas de salud. Aparato y Sección*. [Dataset]. <https://pestadistic.o.inteligenciadegestion.sanidad.gob.es/publicoSNS/C/base-de-datos-de-clinicos-de-atencion-primaria-bdcap/problemas-de-salud>
- Beck, A. T., & Steer, R. A. (1993). *Beck Anxiety Inventory Manual*. Psychological Corporation.

- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *BDI-II. Beck Depression Inventory-Second Edition manual*. Psychological Corporation.
- Bentley, K. H., Gallagher, M. W., Carl, J. R., & Barlow, D. H. (2014). Development and validation of the overall depression severity and impairment scale. *Psychological Assessment*, 26(3), 815–830. <https://doi.org/10.1037/a0036216>
- Bower, P., & Gilbody, S. (2005). Stepped care in psychological therapies: Access, effectiveness and efficiency: Narrative literature review. *British Journal of Psychiatry*, 186(1), 11–17. <https://doi.org/10.1192/bjp.186.1.11>
- Bullis, J. R., Boettcher, H., Sauer-Zavala, S., Farchione, T. J., & Barlow, D. H. (2019). What is an emotional disorder? A transdiagnostic mechanistic definition with implications for assessment, treatment, and prevention. *Clinical Psychology: Science and Practice*, 26(2), e12278. <https://doi.org/10.1111/cpsp.12278>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge.
- Colorafi, K., Vanselow, J., & Nelson, T. (2017). Treating anxiety and depression in primary care: Reducing barriers to access. *Family Practice Management*, 24(4), 11–16.
- Confederación Salud Mental España (2023). *La situación de la Salud Mental en España*. <https://www.consaludmental.org/publicaciones/Estudio-situacion-salud-mental-2023.pdf>
- Corpas, J., Moriana, J. A., Venceslá, J. F., & Gálvez-Lara, M. (2021). Brief psychological treatments for emotional disorders in primary and specialized care: A randomized controlled trial. *International Journal of Clinical and Health Psychology*, 21(1), 100203. <https://doi.org/10.1016/j.ijchp.2020.11.002>
- Costa, P. T., Jr., & McCrae, R. R. (1999). *NEO PI-R, Inventario de Personalidad NEO Revisado. NEOFFI, Inventario NEO reducido de Cinco Factores [Revised NEO Personality Inventory (NEOPI-R) and NEO Five-Factor Inventory (NEO-FFI)]*. TEA Ediciones.
- Costantini, L., Pasquarella, C., Odone, A., Colucci, M. E., Costanza, A., Serafini, G., Aguglia, A., Belvederi Murri, M., Brakoulias, V., Amore, M., Ghaemi, S. N., & Amerio, A. (2021). Screening for depression in primary care with patient health Questionnaire-9 (PHQ-9): A systematic review. *Journal of Affective Disorders*, 279, 473–483. <https://doi.org/10.1016/j.jad.2020.09.131>
- Cuéllar Flores, I., Fernández Garzón, L., Félix-Alcántara, M. P., Manzano Olivares, B., De La Vega Rodríguez, I., Ferreira González, M., Albarsanz, P., Vázquez, M. L., Rivera, S., & Maldonado Belmonte, M. J. (2022). Indicadores asistenciales y Estándares de Calidad asistencial Para la psicología clínica En Los Centros de Salud mental Del sistema Madrileño de Salud evaluados Por Sus profesionales. *Apuntes De Psicología*, 40(2), 71–86. <https://doi.org/10.55414/ap.v40i2.1414>
- Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM-IV: Lifetime Version (ADIS-IV-L)*. Psychological Corporation.
- Díaz, J., Díaz-de-Neira, M., Jarabo, A., Roig, P., & Román, P. (2017). Estudio de derivaciones de Atención primaria a Centros de Salud mental En Pacientes Adultos En La Comunidad de Madrid. *Clinica Y Salud*, 28(2), 65–70. <https://doi.org/10.1016/j.clysa.2017.03.001>
- Díaz González, L. (2008). *Salud mental en atención primaria recomendaciones para el abordaje de los trastornos más prevalentes*. Dirección General de Programas Asistenciales.
- Echeburúa, E., Corral, P., & Fernández-Montalvo, J. (2000). Maladjustment inventory (MI): Psychometric properties in clinical contexts. *Análisis Y Modificación De Conducta*, 26, 325–340.
- Eustis, E. H., Gallagher, M. W., Tirpak, J. W., Nauphal, M., Farchione, T. J., & Barlow, D. H. (2020). The unified protocol compared with diagnosis-specific protocols for anxiety disorders: 12-month follow-up from a randomized clinical trial. *General Hospital Psychiatry*, 67, 58–61.
- Fava, M., Rankin, M. A., Wright, E. C., Alpert, J. E., Nierenberg, A. A., Pava, J., & Rosenbaum, J. F. (2000). Anxiety disorders in major depression. *Comprehensive Psychiatry*, 41(2), 97–102. [https://doi.org/10.1016/S0010-440X\(00\)90140-8](https://doi.org/10.1016/S0010-440X(00)90140-8)
- Finazzi, E., & MacBeth, A. (2022). Service users experience of psychological interventions in primary care settings: A qualitative meta-synthesis. *Clinical Psychology & Psychotherapy*, 29(2), 400–423. <https://doi.org/10.1002/cpp.2650>
- Fonseca, E. (2021). *Manual de tratamientos psicológicos en adultos*. Pirámide.
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, 26(1), 41–54. <https://doi.org/10.1023/B:JOBA.0000007455.08539.94>
- Haugh, J. A., Herbert, K., Choi, S., Petrides, J., Vermeulen, M. W., & D’Onofrio, J. (2019). Acceptability of the stepped care model of depression treatment in primary care patients and providers. *Journal of Clinical Psychology in Medical Settings*, 26(4), 402–410. <https://doi.org/10.1007/s10880-019-09599-2>
- Hervás, G., & Jódar, R. (2008). Adaptación al castellano de la Escala de Dificultades en la Regulación Emocional. *Clinica y Salud*, 19(2), 139–156. <https://scielo.isciii.es/pdf/clinsa/v19n2/v19n2a01.pdf>
- Hiller, T. S., Hoffmann, S., Teismann, T., Lukaschek, K., & Gensichen, J. (2023). Psychometric evaluation and Rasch analyses of the German overall anxiety severity and impairment scale (OASIS-D). *Scientific Reports*, 13(1), 6840.
- Huprich, S. K. (2018). Personality pathology in primary care: ongoing needs for detection and intervention. *Journal of Clinical Psychology in Medical Settings*, 25, 43–54. <https://doi.org/10.1007/s10880-017-9525-8>
- IBM Corp. (2017). *IBM SPSS Statistics para Windows (Version 25.0) [Computer software]*. IBM Corp.
- Imperatori, C., Bianciardi, E., Niolu, C., Fabbricatore, M., Gentileschi, P., Di Lorenzo, G., Siracusano, A., & Innamorati, M. (2020). The Symptom-Checklist-K-9 (SCL-K-9) discriminates between overweight/obese patients with and without significant binge eating pathology: Psychometric properties of an Italian version. *Nutrients*, 12(3), 674. <https://doi.org/10.3390/nu12030674>
- Instituto de Información Sanitaria (2007). *Salud Mental: Organización y dispositivos*. [https://www.sanidad.gob.es/gl/estadEstudios/estadisticas/docs/siap/SALUD\\_MENTAL\\_ORGANIZACION\\_Y\\_DISPOSITIVOS.pdf](https://www.sanidad.gob.es/gl/estadEstudios/estadisticas/docs/siap/SALUD_MENTAL_ORGANIZACION_Y_DISPOSITIVOS.pdf)
- Ito, M., Bentley, K. H., Oe, Y., Nakajima, S., Fujisato, H., Kato, N., & Barlow, D. H. (2015). Assessing depression related severity and functional impairment: The overall depression severity and impairment scale (ODSIS). *PLoS ONE*, 10(4), e0122969.
- Jurado-González, F., García-Torres, F., Contreras, A., Muñoz-Navarro, R., González-Blanch, C., Medrano, L. A., Ruiz-Rodríguez, P., Cano-Vindel, A., & Moriana, J. A. (2024). Mapping bridges between anxiety, depression, and somatic symptoms in primary care patients: A network perspective. *Current Psychology*, 43(5), 4292–4304. <https://doi.org/10.1007/s12144-023-04657-3>
- Lorente, E., Ibáñez, M. I., Moro, M., & Ruipérez, M. A. (2002). Índice de Calidad de Vida: Estandarización y características psicométricas en una muestra española. *Psiquiatría y Salud*, 2(2), 45–50. <https://medworksmidia.com/wp-content/uploads/2017/10/45-50-PSI-V2-N2-LORENTE.pdf>
- Madsen, C., Reinholt, N., Arendt, M., Eskildsen, A., Bryde, A., Hvenegaard, M., & Arnfred, S. (2024). Validation of the Danish version of the overall depression severity and impairment scale (ODSIS): A reliable instrument for assessing depression symptoms and functional impairment. *Clinical Psychology & Psychotherapy*, 31(5), e3049.

- Mezzich, J. E. M. D., Ruipérez, M. A., Pérez, C., Yoon, G., Liu, J., & Mahmud, S. (2000). La versión española del índice de calidad de vida: Presentación y Validación. 188(5):p 301–305, mayo de 2000. *The Journal of Nervous and Mental Disease*, 188(5), 301–305.
- Ministerio de Sanidad. (2022). *Informe anual Del sistema Nacional de Salud 2020–2021*. Ministerio de Sanidad, centro de publicaciones. [Informes, estudios e investigación].
- Muñoz-Navarro, R., Cano-Vindel, A., Ruiz-Rodríguez, P., Adrián Medrano, L., González-Blanch, C., Moriana, J. A., Bonet, C., A., & Dongil-Collado, E. (2017). Modelo Jerárquico de Diagnóstico y derivación de Los Trastornos mentales comunes En Centros de Atención primaria. Una propuesta a partir Del Ensayo clínico Psic-AP. *Ansiedad Y Estrés*, 23(2–3), 124–129. <https://doi.org/10.1016/j.anyes.2017.10.002>
- Norman, S. B., Cissell, H., Means-Christensen, S., A. J., & Stein, M. B. (2006). Development and validation of an overall anxiety severity and impairment scale (OASIS). *Depression and Anxiety*, 23(4), 245–249. <https://doi.org/10.1002/da.20182>
- Norton, P. J. (2012). *Group cognitive-behavioral therapy of anxiety: A transdiagnostic treatment manual*. Guilford Press.
- Osma, J., Peris-Baquero, O., Suso-Ribera, C., Sauer-Zavala, S., & Barlow, D. H. (2021). Predicting and moderating the response to the unified protocol: Do baseline personality and affective profiles matter?? *Cognitive Therapy and Research*, 45(4), 817–830. <https://doi.org/10.1007/s10608-021-10208-6>
- Osma, J., Quilez-Orden, A., Suso-Ribera, C., Peris-Baquero, O., Norman, S. B., Bentley, K. H., & Sauer-Zavala, S. (2019). Psychometric properties and validation of the Spanish versions of the overall anxiety and depression severity and impairment scales. *Journal of Affective Disorders*, 252, 9–18. <https://doi.org/10.1016/j.jad.2019.03.063>
- Osma, J., Quilez-Orden, A., Ordóñez, J., Peris-Baquero, Ó., & Echeburúa, E. (2024). Psychometric properties of the maladjustment inventory in a Spanish clinical and student sample. *BMC Psychology*, 12(1), 666.
- Osma, J., Suso-Ribera, C., García-Palacios, A., Crespo-Delgado, E., Robert-Flor, C., Sánchez-Guerrero, A., Ferreres-Galan, V., Pérez-Ayerra, L., Malea-Fernández, A., & Torres-Alfosea, M. A. (2018). Efficacy of the unified protocol for the treatment of emotional disorders in the Spanish public mental health system using a group format: Study protocol for a multicenter, randomized, non-inferiority controlled trial. *Health and Quality of Life Outcomes*, 16(1), 46. <https://doi.org/10.1186/s12955-018-0866-2>
- Pereira, M., Melo, C., Fonseca, A., & Canavarro, M. C. (2020). European Portuguese version of the Overall Depression Severity and Impairment Scale (ODSIS). *Faculty of Psychology and Educational Sciences*. University of Coimbra, Portugal.
- Peris-Baquero, O., & Osma, J. (2023). Unified protocol for the transdiagnostic treatment of emotional disorders in group format in Spain: Results of a noninferiority randomized controlled trial at 15 months after treatment onset. *Depression and Anxiety*, 2023, 1–16. <https://doi.org/10.1155/2023/1981377>
- Peris-Baquero, Ó., Moreno, J. D., & Osma, J. (2022). Long-term cost-effectiveness of group unified protocol in the Spanish public mental health system. *Current Psychology*. <https://doi.org/10.1007/s12144-022-03365-8>
- Sandora, J., Novak, L., Brnka, R., van Dijk, J. P., Tavel, P., & Malinakova, K. (2021). The abbreviated overall anxiety severity and impairment scale (Oasis) and the abbreviated overall depression severity and impairment scale (odsis): Psychometric properties and evaluation of the Czech versions. *International Journal of Environmental Research and Public Health*, 18(19), 10337.
- Sanz, J., & Navarro, M. E. (2003). Propiedades psicométricas de una versión española del Inventario de Ansiedad de Beck (BAI) en estudiantes universitarios. *Ansiedad y Estrés*, 9(1), 59–84. <https://hdl.handle.net/20.500.14352/98522>
- Sanz, J., García-Vera, M. P., Espinosa, R., Fortún, M., & Vázquez, C. (2005). Adaptación Española Del inventario Para La depresión de Beck-II (BDI-II): 3. Propiedades Psicométricas En Pacientes Con Trastornos Psicológicos. *Clinica Y Salud*, 2(16), 121–142.
- Sanz, J., García-Vera, M. P., & Fortún, M. (2012). The Beck anxiety inventory (BAI): Psychometric properties of the Spanish version in patients with psychological disorders. *Behavioral Psychology/ Psicología Conductual*, 20(3), 563–583.
- Sanz, J., Perdígón, A. L., & Vázquez, C. (2003). Adaptación Española Del inventario Para La depresión de Beck-II (BDI-II): 2. Propiedades Psicométricas En Población general. *CLÍNICA Y SALUD*, 14(3), 249–280.
- Sauer-Zavala, S., Fournier, J. C., Steele, J., Woods, S., Wang, B. K., Farchione, M., T. J., & Barlow, D. H. (2021). Does the unified protocol really change neuroticism? Results from a randomized trial. *Psychological Medicine*, 51(14), 2378–2387. <https://doi.org/10.1017/S0033291720000975>
- Smit, D., Hill, L., Walton, I., Kendall, S., & De Lepeleire, J. (2020). European forum for primary care: Position paper for primary care mental health: Time for change, now more than ever! *Primary Health Care Research & Development*, 21, e56. <https://doi.org/10.1017/S1463423620000304>
- Timulak, L., Richards, D., Bhandal-Griffin, L., Healy, P., Azevedo, J., Connon, G., Martin, E., Kearney, A., O’Kelly, C., Enrique, A., Eilert, N., O’Brien, S., Harty, S., González-Robles, A., Eustis, E. H., Barlow, D. H., & Farchione, T. J. (2022). Effectiveness of the internet-based unified protocol transdiagnostic intervention for the treatment of depression, anxiety and related disorders in a primary care setting: A randomized controlled trial. *Trials*, 23(1), 721. <https://doi.org/10.1186/s13063-022-06551-y>
- Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. Springer Science & Business Media.
- Wigman, J. T. W., Van Nierop, M., Vollebergh, W. A. M., Lieb, R., Beesdo-Baum, K., Wittchen, H. U., & Van Os, J. (2012). Evidence that psychotic symptoms are prevalent in disorders of anxiety and depression, impacting on illness onset, risk, and Severity—Implications for diagnosis and Ultra-High risk research. *Schizophrenia Bulletin*, 38(2), 247–257. <https://doi.org/10.1093/schbul/sbr196>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.