J. Bobes[1]
C. García-Calvo[2]
R. Prieto[2]
M. García-García[3]
F. Rico-Villademoros[3]
y Grupo Español de Trabajo para la validación de la versión española de la Escala de Detección del Trastorno de Ansiedad Generalizada según DSM-IV (Escala de TAG de Carrol y Davidson)*

Psychometric Properties of the Spanish Version of the Screening Scale for DSM-IV Generalized Anxiety Disorder of Carroll and Davidson

[1] University of Oviedo
Psychiatry Department
Oviedo (Spain)
[2] Laboratories Wyeth Farma, S.A.
San Sebastián de los Reyes (Madrid) (Spain)
[3] Biométrica
Barcelona (Spain)

Introduction. The aim was to validate the Spanish version of the screening scale for DSM-IV General Anxiety Disorder of Carroll and Davidson for use in research and clinical practice in Spain for screening and assessing specific anxiety symptoms of patients with Generalized Anxiety Disorder (GAD).

Methods. Observational, prospective, multisite, study comparing between patients with DSM-IV diagnosis of GAD (group A), starting or switching treatment (group A1) or stable patients (group A2), followed-up for 6 months (group A1) or 2 weeks (group A2) versus healthy control subjects (group B), assessed in a single visit.

Results. Among 223 valuable subjects the scale showed: *a)* adequate feasibility with a mean time of administration: 6,53 and 4,49 minutes (DT: 5.48 and 3.56) in groups A and B, and percentage of patients without response < 5%; *b)* adequate reliability (Kuder-Richardson coefficient: 0.85 and 0.79 in groups A1 and A2, and CCI coefficient: 0.89 in group A2); *c)* adequate validity, showing capability for discriminating between patients and controls, with area under curve AUC: 0.9713 (IC 95%: 0.9510-0.9917), and obtaining a high correlation with HARS (r = 0.88) and ICG-G (r = 0.87) scales, y *d)* adequate sensitivity to clinical changes from start and end of treatment (SES: –1.6, –3.1 and –3.8 after 1, 3 and 6 months), spite of the high percentage of patients with highest score in group A1 (38.6%).

Conclusion. The Spanish version of the screening scale for DSM-IV GAD showed adequate psychometric properties for use in research and clinical practice in Spain as well as an screening as evaluative measure for patients with GAD, spite of the ceiling effect showed in severe patients.

Key words:
Validation. Anxiety. Disorders depression. Venlafaxine.

Actas Esp Psiquiatr 2006;34(2):83-93

Propiedades psicométricas de la versión española de la Escala de Detección del Trastorno de Ansiedad Generalizada según DSM-IV de Carroll y Davidson

Introducción. El objetivo era validar en español la Escala del TAG de Carroll y Davidson para su uso en la práctica e investigación clínica en España para detectar y evaluar los síntomas específicos de ansiedad de los pacientes con trastorno de la ansiedad generalizada (TAG).

Método. Estudio observacional, prospectivo, multicéntrico, comparativo entre pacientes con diagnóstico DSM-IV de TAG (grupo A) que iniciaron o cambiaron de tratamiento (grupo A1) o estables (grupo A2), seguidos durante 6 meses (grupo A1) o 2 semanas (grupo A2), frente a controles sanos (grupo B) evaluados en una ocasión.

Resultados. La escala mostró en 223 sujetos valorables: *a)* adecuada factibilidad con tiempo de administración medio: 6,53 y 4,49 minutos (DT: 5,48 y 3,56) en grupos A y B, y porcentaje de pacientes sin respuesta < 5%; *b)* adecuada fiabilidad (coeficientes Kuder-Richardson: 0,85 y 0,79 en grupos A1 y A2, y CCI: 0,89 en grupo A2); *c)* adecuada validez, confirmándose su capacidad discriminante entre pacientes y controles, con área bajo la curva AUC: 0,9713 (IC 95%: 0,9510-0,9917), y su alta correlación con escalas HARS (r = 0,88) e ICG-G (r = 0,87), y *d)* adecuada sensibilidad para detectar cambios clínicos

*Work group (in alphabetical order): Pedro Antón[1], M. Teresa Bascarán[2], Lorenzo Chamorro[3], Alfonso Chinchilla[4], Consuelo de Dios[5], Olga García[6], Javier García-Campayo[7], Patxi Gil[8], Antonio González[9], M. Paz García-Portilla[2], Andrés Herrán[10], Miguel Ángel Jiménez[11], José Ángel Macías[12], Guillem Massana[13], Antonio Pérez[14], José Ramón Pigem[15], Jordi Pujol[16], José Juan Rodríguez[17], José M. Sala[18], Ángel Sanchez[19], Manuel Serrano[20], Pedro Sopelana[21] and Yolanda Zapico[22].

[1] Clínica Mediterránea de Neurociencias (Alicante), [2] Universidad de Oviedo. Área de Psiquiatría (Oviedo), [3] Hospital General (Guadalajara), [4] Hospital Ramón y Cajal (Madrid), [5] C. S. M. Castroviejo (Madrid); [6] C. S. M. Sur (Sevilla); [7] C. S. Torrero (Zaragoza), [8] C. S. M. Sestao (Pamplona), [9] Equipo de Salud Mental (Badajoz), [10] Hospital Valdecilla (Santander), [11] C. S. M. Arganzuela (Madrid), [12] Hospital Clínico Universitario (Valladolid), [13] Hospital Clínico (Barcelona), [14] Centro de Salud de San Bernardo (Salamanca), [15] Clínica Bellavista (Lérida), [16] C. S. M. Mora de Ebro (Mora de Ebro, Tarragona), [17] C. S. M. Colmenar (Colmenar Viejo, Madrid), [18] Hospital Clínico Lozano Blesa (Zaragoza), [19] Hospital Virgen de la Arrixaca (Murcia), [20] Hospital Universitario Juan Canalejo (A Coruña), [21] C. S. M. Reyes Magos (Alcalá de Henares, Madrid) and [22] Hospital del Bierzo (Ponferrada, León).

This study has been sponsored by Wyeth Farma, S.A. San Sebastián de los Reyes (Madrid).

Correspondence:
Julio Bobes
Universidad de Oviedo
Departamento de Psiquiatría
Julián Clavería, 6
33006 Oviedo (Spain)
E-mail: bobes@uniovi.es

entre antes y después del tratamiento (SES: –1,6, –3,1 y –3,8 en meses 1, 3 y 6) pese al elevado porcentaje de pacientes con puntuación máxima en grupo A1 (38,6 %).

Conclusiones. La Escala del TAG posee adecuadas propiedades psicométricas para su uso en la práctica e investigación clínica en España como instrumento de cribaje y evaluativo con pacientes con TAG pese al efecto techo que presenta en pacientes graves.

Palabras clave:
Validación. Trastornos de ansiedad. Depresión. Venlafaxina.

INTRODUCTION

The nosological idea of Generalized Anxiety Disorder (GAD) has considerably changed from the DSM-III[1] to the DSM-IV-R[2]. Over the different editions, the GAD has gone from being a residual category (in the DSM-III1) to reaching in the DSM-III-R[3] the category of diagnostic entity *per se*, whose nuclear and differential element was worry (apprehensive expectation). DSM-IV[4] limited the disorder even more on requiring the subjective perception of difficulty of control on this concern. Furthermore, the temporal and sociolaboral impact criteria became stronger while the accessory-somatic symptoms lost relevance. However, these diagnostic criteria are subject to debate[5,6] for the DSM-V-2011[7], since they exclude a substantial proportion of subjects with a psychosocial deterioration level comparable to that of the patients who totally fulfill the diagnostic criteria for GAD[8].

The nosological changes have made the scientific community develop specific measurement instruments, that are psychometrically adequate to identify and quantify GAD intensity according to its present concept idea and to assess the efficacy of our psychosocial and psychopharmacological interventions in these patients.

Traditionally, the most used assessment instrument to evaluate this disorder was the Hamilton Anxiety Rating Scale (HARS)[9]. However, the HARS was designed to assess the intensity of anxiety, considered as a combination of non-specific symptoms and not that of the generalized anxiety disorder as a specific nosological and well-defined entity. Two solutions have been used to overcome the conceptual differences between the object to be measured (GAD) and construct that really measures the instrument (anxiety). The first consists in assuming that the global anxiety measure of the HARS may be divided into two components[10]; psychic anxiety, made up by the items 1-6 and 14 and physical anxiety (items 7 to 13), the former being the one that most approaches the present GAD concept. The second solution consists in only using the first two items (anxious mood and tension), since they are the only two items on the scale that measure the central GAD aspects. The first solution has been used most in the studies conducted up to now to demonstrate the efficacy of the antidepressive drugs in the GAD treatment (paroxetine[11] and venlafaxine[12-15]).

In the decade of the 90's, specific instruments were developed to measure the construct of pathological worry, the Penn State Worry Questionnaire (PSWQ)[16] being one of the self-administered instruments used most to measure worry and GAD. More recently, a generation of screening and evaluation instruments of the different anxiety disorders has arisen. These instruments are also useful for use in the Primary Health Care setting. This is the case of the instrument developed by Allgulander and Nilsson[17], of the WHAT IF questionnaire[18], or of the SFD[19].

Only two instruments, the Wittchen and Boyer[20] Anxiety Screening Questionnaire (ASQ-15) and the Carroll and Davidson[21] Screening Scale for DSM-IV Generalized Anxiety Disorder, have been designed to assess each and every one of the manifestations that presently define the GAD. The ASQ-15, developed from the Munich Composite International Diagnostic Interview, makes it possible to detect the generalized anxiety disorder according to DSM-IV and ICD-10 criteria, and other anxiety syndromes. On its part, the Carroll and Davidson instrument allows for the identification of patients with generalized anxiety disorder according to the DSM-IV criteria.

Faced with the growing interest of the GAD and the lack of adequate instruments for its detection and evaluation in our setting, the present study was designed to adapt and validate the Carroll and Davidson Screening Scale for DSM-IV Generalized Anxiety Disorder in Spanish for its use in the clinical practice and investigation in Spain.

METHOD

Study design and sample

The present assessment study of the psychometric properties of the GAD Scale was an observational, prospective, multicenter, comparative study among a group of patients with generalized anxiety disorder (GAD) diagnosis (group A). It was, in turn, divided between patients who initiated or switched treatment for the GAD (group A1) and stable patients (group A2) and a healthy control group (group B). These three different groups of subjects were considered in the study design in order to be able to evaluate the different psychometric properties of the GAD Scale being studied. All the subjects had to be ≥ 18 years of age, with the minimum physical and mental aptitudes necessary to understand and fill out the psychometric scales and must have given their written informed consent to participate in the study. Group A patients should also fulfill the following screening criteria: patients with primary diagnosis of GAD according to DSM-IV criteria; in situation of initiation or switching of drug treatment for the GAD (group A1) or in clinically and therapeutically stable situation (with scores ≤ 14 on the HARS scale and a ≤ 3 score on the CGI scale, without changes > 5 points on the HARS scale, without improvement on the CGI scale of change, and without chan-

ges in treatment for the 2 weeks of follow-up) (group A2); without diagnosis of any other relevant psychiatric disease that could interfere in the study objectives, including bipolar disorder, psychotic disorders, mental retardation and dementias. Besides the general criteria mentioned above, the healthy control subjects of group B should not have any relevant medical or psychiatric disease and a score of ≤ 5 on the HARS scale.

Description of the Screening Scale for DSM-IV Generalized Anxiety disorder according to Carroll and Davidson (see appendix 1)

This scale was developed in order to create a simple and rapid instrument for the identification of patients with generalized anxiety disorder according to the DSM-IV criteria. It is a self-administered scale, formed by 12 dichotomic answer items (yes/no) that determine the presence or absence of the DSM-IV criteria for the GAD. Eight of the 12 items refer to psychic anxiety (nervousness, worry, restlessness, concentration), one to sleep difficulties, on to muscle tension and the last 2 evaluate interference with daily life and the need to request help. It clearly establishes a timeframe, most of the days of the last 6 months, that correspond with the temporal criterion for the DSM-IV generalized anxiety disorder.

Appendix 1 | **Escala del TAG de Carroll y Davidson**

Instrucciones: Estas cuestiones son para preguntarle sobre cosas que puede haber sentido la mayoría de los días en los últimos 6 meses:

1. La mayoría de los días me siento nervioso (a)
   ❒ Sí ❒ No
2. La mayoría de los días me preocupo por muchas cosas
   ❒ Sí ❒ No
3. La mayoría de los días no puedo parar de preocuparme
   ❒ Sí ❒ No
4. La mayoría de los días me resulta difícil controlar mis preocupaciones
   ❒ Sí ❒ No
5. Me siento inquieto (a), intranquilo (a) o con los nervios de punta
   ❒ Sí ❒ No
6. Me siento cansado (a) fácilmente
   ❒ Sí ❒ No
7. Tengo problemas para concentrarme
   ❒ Sí ❒ No
8. Me enfado o irrito fácilmente
   ❒ Sí ❒ No
9. Mis músculos están tensos y agarrotados
   ❒ Sí ❒ No
10. Tengo problemas de sueño
    ❒ Sí ❒ No
11. Las cosas que ha señalado anteriormente, ¿afectaron su vida diaria (en el hogar, en el trabajo o en su tiempo libre) o le causaron mucho malestar?
    ❒ Sí ❒ No
12. Las cosas que ha señalado anteriormente, ¿fueron suficientemente molestas como para que pensara en buscar ayuda para ellas?
    ❒ Sí ❒ No

Information collection

Group A1 patients were evaluated for a 6 month period, with a baseline control and a control at months 1, 3 and 6 after having initiated or switched treatment. During the follow-up, not only the GAD Scale is used to evaluate the severity of GAD but also the Clinical Global Impression of Severity (CGI-S) and Improvement (CGI-I) of the anxiety disorder[22] and the 14 item Hamilton Anxiety Rating Scale (HARS)[23]. In addition, in each visit, the specialist questioned the patient on the use made of the drug prescribed, insisting on verifying if the dosage regime indicated had been complied with. All the subjects were informed of the study characteristics and written consent for their participation was obtained. Group A2 patients were only evaluated for 2 weeks, with a baseline control and an evaluation at the end of 2 weeks of being clinically and therapeutically stable, in which the GAD, CGI-S and CGI-I and HARS scales were assessed. Finally, the healthy control subjects of group B were only evaluated at baseline by the administration of the GAD, CGI-S and CGI-I and HARS scales.

Data analysis

To evaluate the feasibility of the questionnaire, the percentage of patients without response for each one of the scale items and for the total scale was calculated. Furthermore, distribution of the total scores obtained with this questionnaire was studied. To do so, the percentage of patients with each one of the different possible total scores was calculated.

To evaluate the capacity of the GAD Scale as a screening instrument between patients and healthy controls, the ROC curve analysis method was used and the sensitivity, specificity, and positive predictive value (PPV) and Negative one (NPV) were evaluated.

Reliability was analyzed in terms of internal consistency, with the Kuder-Richardson coefficient 20 for the total of the scale in the baseline evaluation (groups A and B) and in terms of test-retest reliability, with the use of the intraclass correlation coefficient (ICC) for the scores of the scale in the baseline evaluation and at the end of two weeks (only in group A2 of stable patients).

Two types of analysis were conducted to evaluate validity. In the first place, correlational analyses were conducted between the scores obtained on the scale and those obtained in the HARS and CGI scales to evaluate convergent/divergent validity. The Spearman correlation coefficient was used for this. In the second place, the statistical tests of comparison of non-parametric groups (Mann-Whitney and Kruskal-Wallis) were used to evaluate discriminating validity. This is understood as the capacity of the scale to discriminate between patients and healthy controls and between subgroups of patients with different evolution time, different concomitant psychiatric diagnosis and different anxiety disorder severity.

In order to assess sensitivity to change: *a)* the pre-post-intervention changes in the scale score were analyzed individually by statistical tests of comparison of scores for paired data, using non-parametric tests (McNemar, Friedman and Wilcoxon) according to the behavior of the variables, and *b)* magnitude of the effect collected by the scale was measured by the calculation of the Standardized Effect Size (SES), dividing the changes in the mean pre-post-treatment score with the standard deviation of the scores at baseline and/or by the calculation of the Standardized Response Mean (SRM), dividing the changes in the mean pre-post-treatment score with the standard deviation of these changes[24].

The *p* values referenced in this manuscript correspond to the statistical significance of the two tail tests. Values inferior to or equal to 0.05 were considered statistically significant. Once the study data were listed and quality control conducted, all the analyses were performed with the SPSS version 11.5.1 statistical program.

RESULTS

Sample description

Figure 1 describes the follow-up of the sample of subjects studied. It details the enrolled, evaluable and excluded patients for the three different groups in which the analyses were conducted: group A1 of patients of initiation or switching of treatment, group A2 of clinically and therapeutically stable patients and group B of healthy control subjects. In all, a sample of 261 patients and healthy subjects were enrolled. A total of 223 evaluable patients and healthy subjects were obtained from this.

Table 1 describes the main sociodemographic characteristics of the patient study sample (group A) and healthy controls (group B). This was mainly made up by women (71.4 % and 73.3 %), mean age 42.2 years (SD: 14) and 4.8

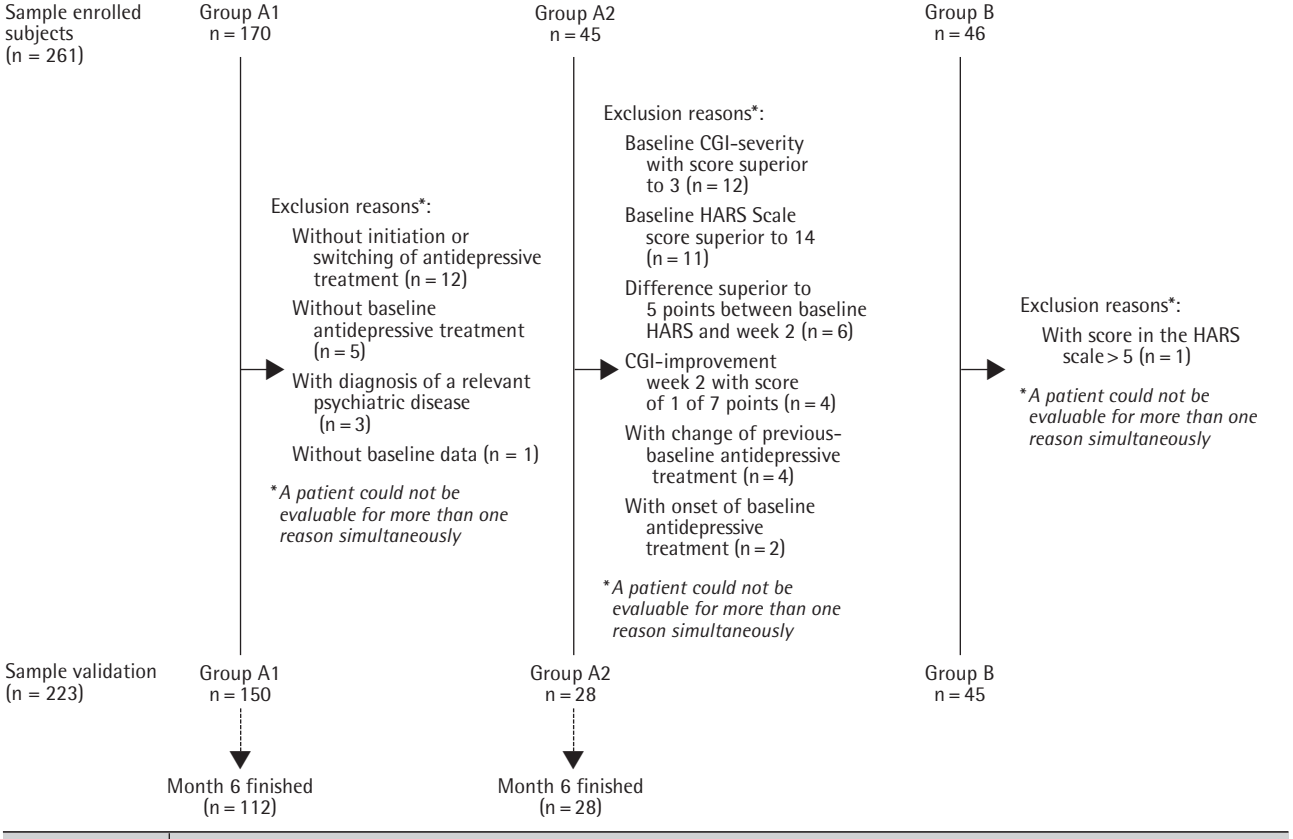

Figure 1 *Diagam of enrolled, evaluable and excluded patients.*

Table 1 Sociodemographic characteristics of the subjects (groups A and B)

	Group A patients	Group A1 patients	Group A2 patients	Group B subjects
Gender (n, %)[a,b,c,d]	175 (100 %)	149 (100 %)	26 (100 %)	45 (100 %)
Man	50 (28.6 %)	40 (26.8 %)	10 (38.5 %)	12 (26.7 %)
Woman	125 (71.4 %)	109 (73.2 %)	16 (61.5 %)	33 (73.3 %)
Age (m, SD)[a,b,c,d]	42.2 (14)	41.5 (13.9)	46 (14.2)	34.8 (11.1)
Stable partner (n, %)[a,b,c,d]	167 (100 %)	141 (100 %)	26 (100 %)	43 (100 %)
With stable partner	110 (65.9 %)	92 (65.2 %)	18 (69.2 %)	34 (79.1 %)
Without stable partner	57 (34.1 %)	49 (34.8 %)	8 (30.8 %)	9 (20.9 %)
Civil status (n, %)[a,b,c,d]	178 (100 %)	150 (100 %)	28 (100 %)	45 (100 %)
Single	49 (27.5 %)	42 (28 %)	7 (25 %)	23 (51.1 %)
Married or living together	104 (58.4 %)	86 (57.3 %)	18 (64.3 %)	20 (44.4 %)
Divorced or separated	14 (7.9 %)	13 (8.7 %)	1 (3.6 %)	0 (0.0 %)
Widow (er)	11 (6.2 %)	9 (6 %)	2 (7.1 %)	2 (4.4 %)
Occupation situation (n, %)[a,b,c,d]	171 (100 %)	144 (100 %)	27 (100 %)	45 (100)
Student	10 (5.8 %)	8 (5.6 %)	2 (7.4 %)	5 (11.1 %)
Housewife	45 (26.3 %)	40 (27.8 %)	5 (18.5 %)	4 (8.9 %)
Active worker	84 (49.1 %)	69 (47.9 %)	15 (55.6 %)	35 (77.8 %)
Unemployed	7 (4.1 %)	6 (4.2 %)	1 (3.7 %)	0 (0 %)
Sick leave	17 (9.9 %)	17 (11.8 %)	0 (0 %)	0 (0 %)
Disability pension	2 (1.2 %)	2 (1.4 %)	0 (0 %)	0 (0 %)
Retirement	7 (4.1 %)	3 (2.1 %)	4 (14.8 %)	1 (2.2 %)
Smoking habit (n, %)[a,b,c,d]	169 (100 %)	142 (100 %)	27 (100 %)	45 (100 %)
Non-smoker	90 (53.3 %)	75 (52.8 %)	15 (55.6 %)	23 (51.1 %)
Ex-smoker	22 (13 %)	19 (13.4 %)	3 (11.1 %)	10 (22.2 %)
Smokes less than 20 cig/day	34 (20.1 %)	29 (20.4 %)	5 (18.5 %)	10 (22.2 %)
Smokes 20 or more cig/day	23 (13.6 %)	19 (13.4 %)	4 (14.8 %)	2 (4.4 %)

[a] There are some cases for which the datum was not specified. [b] Total patients who supplied the datum. [c] Percentage/mean calculated on the total of patients who supplied the datum. [d] Non statistically significant differences (chi squared test; $p > 0.05$), except in age (Mann-Whitney test; $p < 0.05$) were found between the patients of groups A and the group B of healthy controls.

years (SD: 11.1) with stable partner (65.9 % and 79.1 %). There were no statistically significant differences between patients and healthy controls regarding any of the sociodemographic characteristics collected (chi squared test; $p > 0.05$).

There were also no statistically significant differences between patients (group A) and healthy controls (group B), except for age, weight and mean BMI (Mann-Whitney test, $p < 0.05$). It was found in this that the patients of group A had higher age, weight and BMI (m: 42.2 years; SD: 14.0; m: 67.7 kg; SD: 13.6; m: 24,5 kg/m^2; SD: 4.1) than the group B patients (m: 34.8 years; SD: 11.1; m: 63.1 kg; SD: 10.8; m: 22,6 kg/m^2; SD: 2.9). Table 2 describes the principal clinical characteristics of the patient group, there being no statistically significant differences between the patients A1 and A2 groups.

Score feasibility and distribution

In regards to the feasibility of the scale, mean administration time obtained in Group A of patients with GAD was 6.53 minutes (SD: 5.48), while that of group B of the healthy control subjects was 4.49 minutes (SD: 3.56). This difference was statistically significant (Mann-Whitney test, $p < 0.05$). Furthermore, a low percentage of patients without response was observed both for each item and for all the scale (less than 5 % in all the groups analyzed).

Figure 2 shows the distribution of the scale scores, observing a percentage of patients with minimum score on the scale (floor effect) less than 15 % in all the groups. It stands out that 38.6 % of the patients who initiated or switched treatment and thus have greater severity (group A1) obtained the maximum score (ceiling effect).

Table 2 Clinical characteristics of the patients (group A)

	Group A patients	Group A1 patients	Group A2 patients
Total patients with some concomitant diagnosis (n, %)[b,c]	75 (42.1%)	67 (44.7%)	8 (28.6%)
Anxiety disorder (n, %)[a,b,c]	35 (19.7%)	32 (21.3%)	3 (10.7%)
Anxiety disorder without agoraphobia	9 (5.1%)	8 (5.3%)	1 (3.6%)
Anxiety disorder with agoraphobia	6 (3.4%)	6 (4%)	0 (0%)
Agoraphobia without history of anxiety disorder	1 (0.6%)	0 (0%)	1 (3.6%)
Specific phobia	3 (1.7%)	3 (2%)	0 (0%)
Social phobia	7 (3.9%)	6 (4%)	1 (3.6%)
Obsessive-compulsive disorder	2 (1.1%)	2 (1.3%)	0 (0%)
Post-traumatic stress disorder	3 (1.7%)	3 (2%)	0 (0%)
Acute stress disorder	3 (1.7%)	3 (2%)	0 (0%)
Anxiety disorder due to medical disease	0 (0%)	0 (0%)	0 (0%)
Substance induce anxiety disorder	0 (0%)	0 (0%)	0 (0%)
Unspecified anxiety disorder	2 (1.1%)	2 (1.3%)	0 (0%)
Emotional state disorders (n, %)[a,b,c]	36 (20.2%)	31 (20.7%)	5 (17.9%)
Major depressive disorder, single episode	15 (8.4%)	14 (9.3%)	1 (3.6%)
Major depressive disorder, recurrent	5 (2.8%)	4 (2.7%)	1 (3.6%)
Dysthymic disorder	15 (8.4%)	12 (8%)	3 (10.7%)
Unspecified depressive disorder	1 (0.6%)	1 (0.7%)	0 (0%)
Other disorders	0 (0%)	0 (0%)	0 (0%)
Disorders related with substances (n, %)[a,b,c]	1 (0.6%)	1 (0.7%)	0 (0%)
Other concomitant psychiatric diagnoses (n, %)[a,b,c]	5 (2.8%)	5 (3.3%)	0 (0%)
Age of onset of disorder (years) (m, DE)[b,c]	36.3 (14.3)	35.6 (14)	40.3 (15.1)
Duration of disorder (years) (m, DE)[b,c]	6.1 (6.7)	6.2 (7)	5.7 (4.8)
Duration of present disorder (days) (m, DE)[b,c]	6 (6.8)	6.0 (7.1)	6 (4.7)

[a] A patient could have more than one diagnosis simultaneously. [b] Percentage/mean calculated on the total of patients of each group (A: 178; A1: 150, y A2: 28). [c] No statistically significant differences were found between the patients of groups A1 and A2 regarding number of patients with some concomitant diagnosis or for any of the four specified diagnostic categories (chi squared test; $p > 0.05$), or in regards to age of onset or duration of disorder OR duration of present episode (Mann-Whitney test; $p > 0.05$).

Analyses of ROC curves, sensitivity, specificity and positive and negative predictive value

Figure 3 describes the capacity of the scale to distinguish between patients and controls. An AUC (area under the ROC curve) of 0.9713 (95% CI: 0.9510-0.9917) is observed. The analysis of the ROC curves showed that the most optimum cut-off in which the best properties of sensitivity, specificity and positive and negative predictive value were obtained was the score > 3 (GAD), properties that are described in table 3.

Reliability

The reliability of the scale in terms of internal consistency was elevated for all the patients with GAD (Kuder-Richardson coefficient: 0.85), and especially for the stable patient group (Kuder-Richardson coefficient: 0.79). However, the internal consistency in the patient group that initiated or switched treatment was somewhat less than that recommended (Kuder-Richardson coefficient: 0.56). Test-retest reliability of the scale, that was only analyzed in the group of 28 stable patients, was very high for the total score of the scale (ICC: 0.89), although items 3 («I can't stop worrying most days») and 8 («I get angry or irritated easily») are more unstable over time below the recommended coefficient value of 0.70 (ICC: 0.68 and ICC: 0.63, respectively).

Validity

Table 4 shows the results obtained in the correlations between the GAD Scale and the HARS and CGI-Severity scales, that report elevated Spearman correlation coefficients with both the total HARS scale scores ($r = 0.8769$) and the psychic anxiety subscale (made up of items 1-6 and 14) ($r = 0.8821$), as well as the CGI-Severity scale ($r = 0.8726$). Furthermore, an item to item correlational analysis was conducted between the GAD Scale and the total scores of the HARS scale and of the subscale of psychic anxiety of this scale. Standing out in this is the high correlation obtained between the items 1, 2, 3, 5, and 12 of the GAD Scale

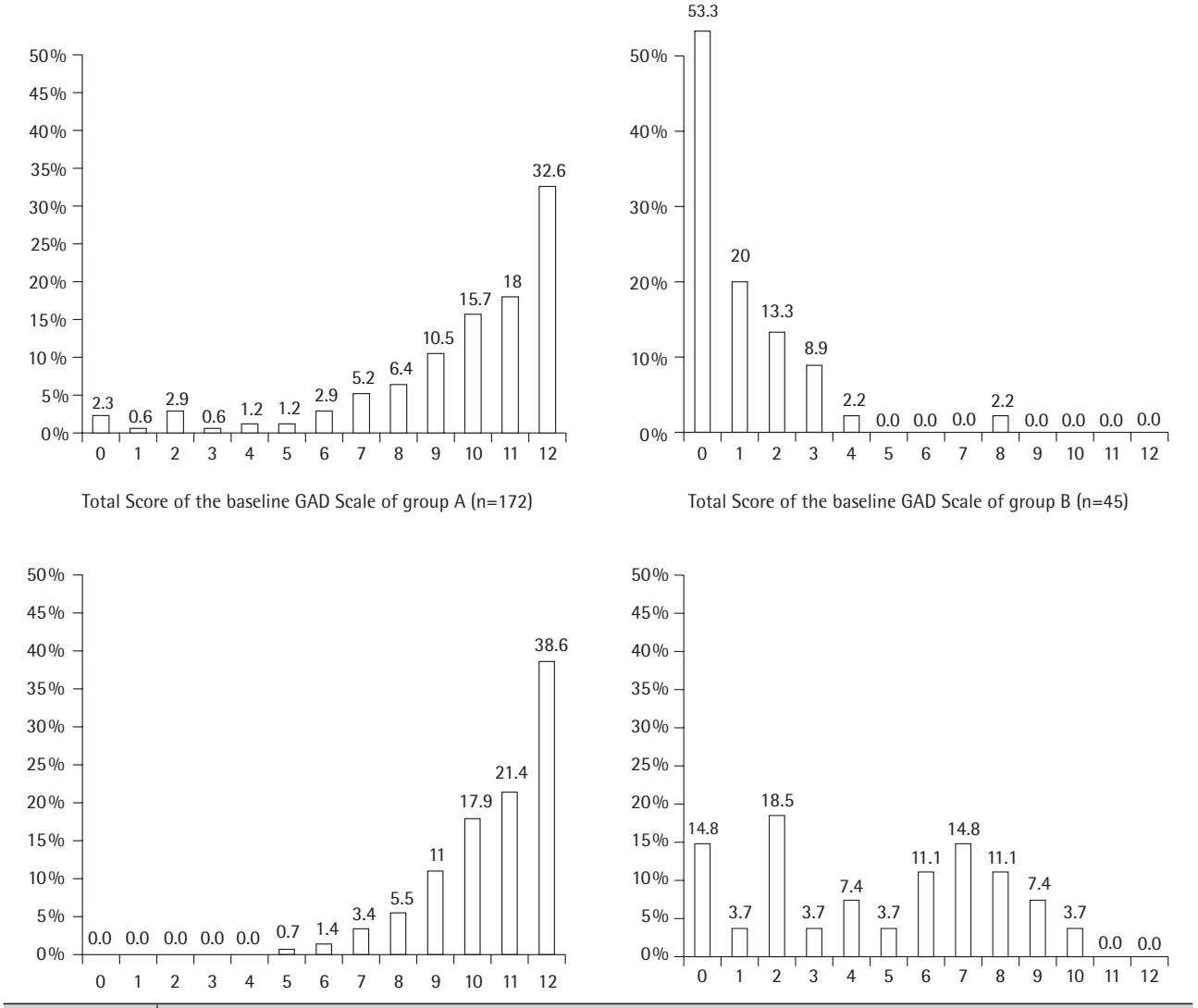

Figure 2 *Distribution of the total scores of the GAD Scale: percentages of patients with different total score.*

and the following items of the HARS Scale 1 («anxious mood»), 2 («tension»), 5 («intellectual functions»), 6 («depressive mood») y 14 («behavior during the interview»), with Spearman correlation coefficients greater than r = 0.70.

Regarding the discriminating validity of the GAD Scale, this property was examined in 2 ways: *a)* comparing patients and controls, whose results have already been described; and *b)* comparing subgroups of patients established according to age of initiation and duration of GAD disorder, duration of the present episode, presence of other concomitant disorders (anxiety, related with concomitant substances or mood state) and different baseline scores on the HARS and CGI-Severity scales. Comparing subgroups of patients, table 5 shows that the GAD Scale was capable of discriminating between patients with different grade of anxiety, baseline clinical severity (according to scores on HARS and CGI-S Scales, respectively) and patient's age (Kruskall-Wallis test; p < 0,05).

Sensitivity to change

Finally, table 6 shows the results of the sensitivity analyses to change of the GAD Scale. According to them, there were statistically significant differences regarding the baseline visit in all the follow-up visits (Wilcoxon test; p < 0.05). In addition, it should be stressed that the scale was sufficiently sensitive from the first month to record a mean standardized effect size ≥ 1.6 points and standardized response mean ≤ 0.9 points.

CONCLUSIONS

The purpose of the present study was to validate the GAD Scale of Carroll and Davidson (Screening Scale for DSM-IV General Anxiety Disorder) in Spanish and according to the standard recommended methods[25], for its use in clinical

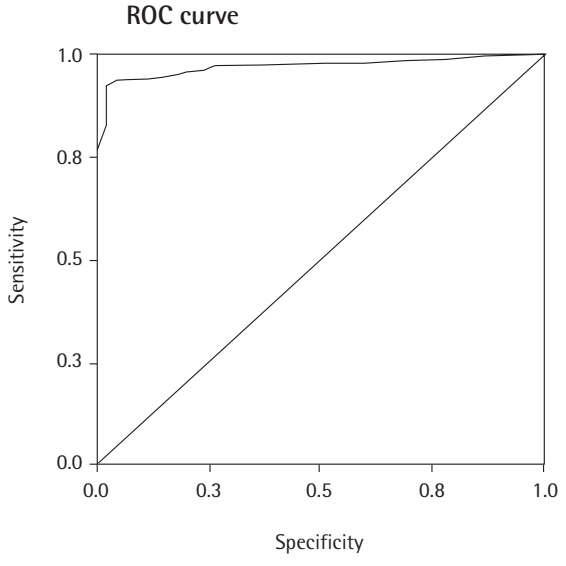

Figure 3 *ROC curves (groups A and B).*

practice and investigation in Spain. To do so, the different properties of this instrument were studied, obtaining the results subjected to discussion in the following.

Feasibility

One of the problems of the specific evaluation instruments used in the evaluation of psychiatric disorders is their lack of feasibility in the clinical practice, some because they are too long and others because they have scarce diffusion and clinical application. This characteristic has been evaluated in the present study, analyzing the administration time of the GAD Scale and the percentage of patients who left some item without an answer. The result is that it is seen as a specific instrument with elevated feasibility.

Table 3 Sensitivity, specificity and positive and negative predictive value of the GAD Scale with cut-off > 3 for screening of patients with GAD

	95% CI		
	Value	Min.	Max.
Sensitivity	0.9360	0.8995	0.9726
Specificity	0.9556	0.8953	1.0158
Predictive value+	0.9877	0.9708	1.0046
Predictive value–	0.7963	0.6889	0.9037

Table 4 Convergent/divergent validity: correlations between scores of the GAD, HARS and CGI-S Scales (groups A y B)

	HARS Scale total score	HARS Psychic Anxiety Scale score	CGI-S Scale total score
Total score GAD Scale			
Coef.	0.88	0.89	0.87
Sig.	<0.0001	<0.0001	<0.0001
n	213	216	213
Total score HARS Scale			
Coef.	–	0.96	0.90
Sig.	–	<0.0001	<0.0001
n	–	218	213
Score-HARDS Psychic Anxiety scale			
Coef.	–	–	0.89
Sig.	–	–	<0.0001
n	–	–	217

Coef.: correlation coefficient (bivariant); Sig.: significance level of Spearman correlation test; n: number of patients.

Reliability

The reliability of the scale in terms of internal consistency was elevated for the total patients with GAD (Kuder-Richardson coefficient: 0.85), and especially for the stable patient group (Kuder-Richardson coefficient: 0.79). However, internal consistency in the group of patients who initiated or switched treatment was somewhat less than that recommended (Kuder-Richardson coefficient: 0.56). The test-retest reliability of the scale, that was only analyzed in the group of 28 stable patients, was very high for the total score of the scale (ICC: 0.89), although items 3 («I can't stop worrying most days») and 8 («I get angry or irritated easily») were more unstable over time below the recommended coefficient value of 0.70 (ICC: 0.68, and ICC: 0.63, respectively).

Utility of the GAD Scale as screening instrument and its validity

The GAD diagnosis was not included in the psychiatric nomenclature until the publication of the third edition of

Table 5 Discriminating validity: comparison of the total score on the GAD Scale between subgroups of patients (group A)

	No.*	Mean	SD	Min.	Max.	*p*
Onset age of disorder						
Less than 25 years	34	9.2	2.3	4	12	
Between 25 and 50 years	101	10.1	2.6	0	12	<0.05**
More than 50 years	29	8.7	4.0	0	12	
Disorder duration						
Less than 3 years	55	10.1	2.4	0	12	
Between 3 and 5 years	45	9.1	3.4	0	12	>0.05**
More than 5 years	60	9.6	2.9	2	12	
Present episode duration						
Less than 7 days	6	10.2	2.1	7	12	>0.05**
7 or more days	147	9.8	2.9	0	12	
Concomitant anxiety disorder						
Without disorder	137	9.5	3.0	0	12	>0.05**
With disorder	35	10.3	2.5	2	12	
Concomitant disorders related with substances						
Without disorders	171	9.6	2.9	0	12	>0.05**
With disorders	1	11.0	–	11	11	
Concomitant mood state disorders						
Without disorder	34	10.2	2.6	2	12	>0.05**
With disorder	138	9.5	2.9	0	12	
Baseline HARS						
Low score (< 20)	53	7.1	3.6	0	12	<0.05**
Mean score (≥20 and ≤40)	113	10.8	1.3	6	12	
High score (>40)	3	11.7	0.6	11	12	
Baseline CGI-S						
Without disease	4	3.0	3.5	0	8	
Very mildly or mildly ill	39	6.6	3.4	0	12	<0.05**
Moderately ill	81	10.5	1.4	7	12	
Markedly seriously or extremely ill	46	11.1	1.4	6	12	

*There were some cases for which the datum was not specified **No statistically significant differences (Kruskal-Wallis test; $p > 0.05$, and Mann-Whitney test; $p > 0.05$) were found between the groups of patients, except regarding age of onset of disorder and score on the HARS and CGI-S scales (Mann-Whitney test; $p < 0.05$).

Table 6 Total score on the GAD Scale in the different visits and value of the changes (mean difference in crude value, standardized effect size SES and standardized response mean SRM)

	No.*	Mean	DT	Min.	Max.
Total score of GAD Scale**					
Baseline	145	10.6	1.6	5.0	12.0
Month 1	150	8.0	3.1	1.0	12.0
Month 3	150	5.6	4.0	0.0	12.0
Month 6	150	4.4	4.1	0.0	12.0
Decrease of score (mean difference in crude value)					
Baseline–month 1	145	–2.6	2.9	4.0	–9.0
Baseline–month 3	145	–5.0	4.0	1.0	–12.0
Baseline–month 6	145	–6.1	4.2	1.0	–12.0
Standardized effect size (SES)					
Baseline–month 1	145	–1.6	1.8	2.5	–5.6
Baseline–month 3	145	–3.1	2.5	0.6	–7.5
Baseline–month 6	145	–3.8	2.6	0.6	–7.5
Standardized response mean (SRM)					
Baseline–month 1	145	–0.9	1.0	1.4	–3.1
Baseline–month 3	145	–1.2	1.0	0.3	–3.0
Baseline–month 6	145	–1.5	1.0	0.2	–2.9

*There were some cases for which the datum was not specified. **Statistically significant differences were found during all the study (Friedman test; $p < 0.001$), and in each one of the follow-up visits regarding the baseline visit (Wilcoxon test; $p < 0.001$).

the DSM-III Diagnostic and Statistical Manual of Mental Disorders[1]. However, over the last two decades, it has been the object of many investigations that have contributed to clarify its conceptual limits with other anxiety disorders, such as panic attack, agoraphobia and social phobia as well as with other concomitant psychiatric disorders and to confirm its validity as a different diagnostic category[26,27]. Consequently, an especially relevant aspect to evaluate, but one also having intrinsic difficulty, is the capacity of the GAD Scale as a screening instrument of patients with GAD versus healthy subjects and versus patients with other types of disorders. In regards to its utility to detect GAD cases among healthy subjects, the GAD Scale was seen to be valid, presenting very satisfactory grades of sensitivity, specificity and positive and negative predictive value. Regarding the utility of this instrument to detect patients with GAD among the psychiatric population with other different disorders, it should be stated that a large part of the patients included in the study comorbidly had other anxiety disorders, related with concomitant substances or mood states other than GAD. This constitutes an unavoidable limitation of the study as it reflects the high rates of existing comorbidity in the reality with other psychiatric disorders in GAD patients[26,28]. In spite of this limitation, the GAD Scale not

only had an adequate convergent/divergent validity as can be inferred from the high correlations shown with the HARS and CGI-S Scales, but also an adequate discriminating validity on distinguishing between patients with different grades of anxiety and clinical severity. In addition, more specifically, the high correlation obtained between the items 1, 2, 3 and 5 of the GAD Scale and the following items of the HARS scale 1 («anxious mood») and 2 («tension»), all of which are focused on the presence of anxiety and excessive worry (apprehensive expectation) that characterize the accepted description of the GAD diagnosis according to DSM-IV[29], should also be stressed as a sign in favor of the validity of the GAD Scale.

Sensitivity to change

Although the GAD Scale is a screening instrument which should consequently be essentially discriminative, sensitivity to charge was also studied. This sensitivity is a fundamental psychometric property for the use of an instrument as evaluative[30]. In this sense, the GAD Scale also showed adequate sensitivity to change, as can be concluded from the differences observed between the baseline and post-treatment scores. On estimating the magnitude of the changes between before and after treatment, the changes observed also stand out. These changes observed, from month 1, indicated that the scale is sufficiently sensitive to record a mean standardized effect size ≥ 1.6 points and standardized response mean ≥ 0.9 points. This corroborates an elevated sensitivity to change of this scale, making it a valid indicator of the effectiveness of the GAD treatment. This elevated sensitivity to change was observed even in spite of the limitation that the scale had in serious patients that should be mentioned. This limitation consists in the presence of an elevated proportion of patient with maximum score (38.6 % of the patients in those who initiated or switched treatment of group A1) (ceiling effect).

Final conclusion

The recent interest to validate questionnaires that evaluate changes during GAD associated to treatment responds to the need to have instruments having valid and reliable measurement to detect and evaluate the impact that the new psychodrugs that have been appearing for its treatment exert on this disorder. In spite of this interest, there are few evaluation studies of psychometric properties of specific measurement instruments to evaluate this disorder. These are even more scare if only the generic studies on the measurement instruments of the anxiety adapted and validated to Spanish such as the HARS scale[31] are considered. These are totally non-existent if works on the specific instruments for the GAD are considered. Thus, the novelty of the results presented in this study stand out. They show that the GAD Scale has adequate psychometric properties for its use in investigation and usual clinical practice as a screening as well as evaluative instrument with patients diagnosed of GAD, in spite of the tendency of its measurement scale to have a ceiling effect in serious patients.

REFERENCES

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 3rd ed. Washington: American Psychiatric Association, 1980.
2. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed. rev. Washington: American Psychiatric Association, 2000.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 3rd ed. rev. Washington: American Psychiatric Association, 1987.
4. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed. Washington: American Psychiatric Association, 1994.
5. Kessler RC, Wittchen HU. Patterns and correlates of generalized anxiety disorder in community samples. J Clin Psychiatry 2002; 63(Suppl. 8):4-10.
6. Maier W, Gansicke M, Freyberger HJ, Linz M, Heun R, Lecrubier Y. Generalized anxiety disorder (ICD-10) in primary care from a cross-cultural perspective: a valid diagnostic entity? Acta Psychiatr Scand 2000;101:29-36.
7. Kupfer DJ, First MB, Regier DA. A research agenda for DSM-V. Washington: American Psychiatric Association, 2002.
8. Allgulander C. Generalized anxiety disorder: current challenges and future options. Estocolmo: comunicación presentada en el 14 Congreso ECNP, 2004.
9. Hamilton M. The assessment of anxiety states by rating. Br J Med Psychol 1959;32:50-5.
10. Echeburúa Odriozola E. Medición clínica en trastornos fóbicos y de ansiedad. En: Bulbena A, Berrios GE, Fernández de Larrinoa P, editores. Medición clínica en psiquiatría y psicología. Barcelona: Masson, 2002.
11. Rocca P, Fonza V, Scotta M, Sanadla E, Ravizza L. Paroxetine efficacy in the treatment of generalized anxiety disorder. Acta Psychiatr Scand 1997;95:444-50.
12. Davidson JRT, DuPont RL, Hedges D, Haskins JT. Efficacy, safety and tolerability of venlafaxine extended release and buspirone in outpatients with generalized anxiety disorder. J Clin Psychiatry 1999;60:528-35.
13. Rickels K, Pollack MH, Sheehan DV, Haskins JT. Efficacy of extended-release venlafaxine in nondepressed outpatients with generalized anxiety disorder. Am J Psychiatry 2000;157:969-74.
14. Gelenberg AJ, Lydiard RB, Rudolph RL, Aguiar L, Haskins JT, Salinas E. Efficacy of venlafaxine extended-release capsules in nondepressed outpatients with generalized anxiety disorder: a 6-month randomized controlled trial. JAMA 2000;283:3082-8.
15. Allgulander C, Hackett D, Salinas E. Venlafaxine extended release (ER) in the treatment of generalised anxiety disorder. Br J Psychiatr 2001;179:15-22.
16. Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the Penn State Worry Questionnaire. Behav Res Ther 1990;28:487-95.

17. Allgulander C, Nilsson B. A nationwide study in primary health care: one out of four patients suffers from anxiety and depression. Lakartidningen 2003;100:832-8.
18. Weiss MD, Gadow KD, Wasdell MB, Spraikin J, Bomben MM. Evaluating an new anxiety screening, severity and outcome measure: «What If?» Communication presented in the American Psychiatric Association 157th Annual Meeting (NR193). New York, 2004.
19. Wan GJ, Geissler EC, Zhang HF, Hackett D. Assessment of Symptom-Free Days in GAD. Communication presented in the American Psychiatric Association 156th Annual Meeting (NR623). San Francisco, 2003.
20. Wittchen HU, Boyer P. Screening for anxiety disorders. Sensitivity and specifity of the Anxiety Screening Questionnaire (ASQ-15). Br J Psychiatry 1998;173(Suppl. 34):10-7.
21. Carroll BJ, Davidson JRT. Screening Scale for DSM-IV GAD. Copyright, 2000.
22. Guy W. Early Clinical Drug Evaluation Unit (ECDEU) assessment manual for psychopharmacology. Revised. NIMH publication DHEW publ NO (Adm) 76-338. Bethesda: National Institute of mental Health, 1976; p. 217-22.
23. Hamilton M. The assessment of anxiety by rating. Br J Med Psychol 1959;32:50-5.
24. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989;27:S178-89.
25. Nunnally JC. Psychometric theory. New York: McGraw-Hill, 1978.
26. Stein DJ. Comorbidity in generalized anxiety disorder: impact and implications. J Clin Psychiatry 2001;62(Suppl. 11):29-34.
27. Wittchen HU, Hoyer J. Generalized anxiety disorder: nature and course. J Clin Psychiatry 2001;62(Suppl. 11):15-21.
28. Wittchen HU, Zhao S, Kessler RC, Eaton WW. DSM-III-R generalized anxiety disorder in the National Comorbidity Survey. Arch Gen Psychiatry 1994;51:355-64.
29. American Psychiatric Association. Manual diagnóstico y estadístico de los trastornos mentales. Versión española de la cuarta edición de la obra original en lengua inglesa. Barcelona: Masson, 1995.
30. Guyatt GH. A taxonomy of health status instruments. J Rheumatol 1995;22:1188-90.
31. Lobo A, Chamorro L, Luque A, Dal-Re R, Badia X, Baro E. Grupo de Validación en Español de Escalas Psicométricas (GVEEP). Validation of the Spanish versions of the Montgomery-Asberg depression and Hamilton anxiety rating scales. Med Clin (Barc) 2002;118:493-9.