

RESEARCH

Open Access



Psychometric properties and diagnostic accuracy of the short form of the geriatric anxiety scale (GAS-10)

Leonardo Carlucci^{1*}, Matteo Balestrieri², Elisa Maso³, Alessia Marini¹, Nadia Conte¹ and Michela Balsamo¹

Abstract

Background: Anxious symptoms have a negative impact on different aspects of the elderly's quality of life, ranging from the adoption of unhealthy lifestyle behaviours to an increased functional impairment and a greater physical disability. Different brief assessment instruments have been developed as efficacy measures of geriatric anxiety in order to overcome psychometric weaknesses of its long form. Among these, the 10-item Geriatric Anxiety Scale (GAS-10) showed strong psychometric properties in community-dwelling samples. However, its diagnostic accuracy is still unexplored, as well as its discriminative power in clinical samples.

Methods: In the present study, we explored the psychometric performance of the GAS-10 in the elderly through Item Response Theory in a sample of 1200 Italian community-dwelling middle-aged and elderly adults (53.8% males, mean age = 65.21 ± 9.19 years). Concurrent validity, as well as diagnostic accuracy, was examined in a non-clinical sample ($N = 229$; 46.72% males) and clinical sample composed of 35 elderly outpatients (74.28% females) with Generalized Anxiety Disorder (GAD).

Results: The GAS-10 displayed good internal construct validity, with unidimensional structure and no local dependency, good accuracy, and no signs of Differential Item Functioning (DIF) or measurement bias due to gender, but negligible due to the age. Differences in concurrent validity and diagnostic accuracy among the long form version of the GAS and the GAS-10 were not found significant. The GAS-10 may be more useful than the longer versions in many clinical and research applications, when time constraints or fatigue are issues.

Conclusion: Using the ROC curve, the GAS-10 showed good discriminant validity in categorizing outpatients with GAD disorder, and high anxiety symptoms as measured by the GAS-SF cut-off. The stable cut-off point provided could enhance the clinical usefulness of the GAS-10, which seems to be a promising valid and reliable tool for maximize diagnostic accuracy of geriatric anxiety symptoms.

Keywords: Anxiety, IRT, Geriatric, GAD, Elderly

* Correspondence: l.carlucci@unich.it

¹Department of Psychological Sciences, Health and Territory, "G. d'Annunzio" University of Chieti-Pescara, via dei Vestini, 31 66100 Chieti, Italy
Full list of author information is available at the end of the article



© The Author(s). 2021 **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/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

In Europe, anxiety disorders are one of the most common mental health problems among elderly, with lifetime prevalence estimates ranging from 20.1% in Italy to 32.6% in England [1, 2]. This prevalence is probably underestimated, considering that subsyndromal manifestations of anxious symptoms probably range from 15 to 52.3% in community-living older adults [3, 4]. With an increasingly ageing population, the socio-economic costs of mental health problems – such as anxiety disorders – are growing significantly [3], especially in countries like Italy with low fertility rates and a high life expectancy. As of January 1st 2020, 13.9 million Italians (23.1% of the population) were over 65 years of age, and this percentage continues to increase annually [5]. By 2050 this percentage is expected to reach nearly 30% in most of the western countries [6].

Despite its impact and the high prevalence among the elderly population, anxiety remains often undiagnosed [7, 8]. A possible explanation lies on the specific characteristics of anxious symptoms in later life. Indeed, older adults typically report an increased sensitivity to several types of somatic stimuli and more physical complaints than younger adults, which could be mistaken for psychosomatic symptoms [9]. Moreover, anxious symptoms are often comorbid with other medical conditions and depressive disorders, worsening their prognosis and making the differential diagnosis difficult [10, 11]. Anxiety has also a substantial socio-economic burden, due to an increased use of health services among those who report greater symptoms [12]. Thus, both clinicians and mental health practitioners need valid, reliable measures of anxious symptoms specifically tailored for this population and carefully developed to consider the clinical manifestations of late-life anxiety.

Self-report measures are commonly used to assess anxious symptoms in elderly thanks to their easy administration and to the reduced respondent burden compared to other assessment methods [13]. In their recent work, Balsamo and colleagues [13] reviewed 12 self-report instruments that are commonly used to assess anxiety in later life, some of which are specifically tailored for older population, while others have been validated on geriatric population only later. The first group includes the Geriatric Anxiety Scale (GAS) [14], the Adult Manifest Anxiety Scale (AMAS), [15] the Geriatric Anxiety Inventory (GAI), [16] and the Worry Scale (WS), [17]; while among the second group there are the State-Trait Anxiety Inventory (STAI), [18] the Beck Anxiety Inventory (BAI), [19] the Penn State Worry Questionnaire (PSWQ), [20], and the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA), [21]. They varied depending on the design format (Likert scales vs. dichotomous answers), the length (from 16 to

88 items), and the underlying factor structures (unidimensional vs. multidimensional). However, despite their popularity in contexts whereby assessing anxiety in older adults is required, some of these measures lack good psychometric properties, especially those not originally intended for older adults [13].

Among those instruments specifically tailored for older adults, the Geriatric Anxiety Scale (GAS) is among the most promising ones. Characterized by good to excellent psychometric properties, GAS was developed by Segal and colleagues in 2010. It is composed of 30 items (even though 5 items are intended for clinical purposes and therefore they are not included in the final score) rated on a 4-point Likert scale, ranging from 0 (not at all) to 3 (all of the time). The GAS investigates three dimensions of anxiety, namely Somatic, Cognitive, and Affective ones, and this factorial structure has been confirmed in recent research [22]. The original version of the instrument has been validated on a large sample of community dwelling older adults [14] and on a sample of medically ill older adults [23]. To date, the GAS is one of the most widely-adopted instrument for the measurement of anxiety in geriatric population, as evidenced by its translation in many languages: Arabic [24] Chinese [25], German [26], Italian [27], Persian [28], and Turkish [29]. Furthermore, psychometric properties of the GAS are good: it demonstrated excellent internal consistency ($\alpha = 0.88\text{--}0.93$) in non-clinical population and significant convergent and discriminant validity [13].

Due to the specificity of the target population, there is a need for development and validation of short-forms of geriatric anxiety questionnaires, that allow researchers and clinicians to collect necessary symptom data with less respondent burden [30]. Shorter versions may also help in reducing missing data, thereby improving data quality, and reduce the time required by mental health practitioners and clinicians for their administration and scoring [30]. In addition, the negative consequences of anxiety let emerge the need of short and easy-to-use instruments which could allow its detection at earlier stages, before that deleterious effects of chronic forms of anxiety occur [8].

A few short versions of GAS are currently available. The most recently developed was a German version of the GAS, named GAS-G-SF [31]. Validated both in the general population ($N = 242$) and medically ill sample ($N = 156$), this 9-item version showed good psychometric properties. However, its Somatic subscale exhibited lower reliability than the other two (Cognitive and Affective) subscales. Unlike the GAS-10, the GAS-G-SF was developed to be a multidimensional measure of geriatric anxiety. Specifically, three-factor model without any error covariance was provided an excellent fit by Gottschling et al. [31].

The 10-item GAS-10 version [32] was the most widespread short form of the GAS, developed and validated on a large US sample by using the Item Response Theory (IRT) approach. Excellent psychometric properties, including reliability, convergent validity, and unidimensionality, were reported by Mueller et al. [32]. More recently, Pifer and colleagues [33, 34] have developed an adapted 10-item version of the GAS (GAS-LTC), specifically tailored for long-term care settings. Starting from the already existing GAS-10, they slightly modified items according to specific cognitive needs of long-term care residents, as well as changed response format from an ordinal Likert-type scale to a dichotomous Yes/No response. Even though the initial validation was conducted in a small sample ($N=66$), it showed good internal consistency and convergent validity.

Overall, both the unidimensional and multidimensional short forms developed so far appear to retain the amount of information captured by the longer version of the GAS, and confirmed their utility in detecting anxiety in older adults.

It is well known that unidimensionality is a desirable requirement for calculating and interpreting a total score of a clinical self-report instrument [35]. However, health outcomes measures are rarely strictly unidimensional [36]. This is due to the heterogeneous items that represent the complexity of health constructs [37, 38]. Although multidimensional instruments potentially reflected the heterogeneity of psychiatric disorders [39], unidimensional health measures were crucial in measuring the performance of health systems, evaluating outcomes in clinical trials and in everyday clinical practice [40]. To this purpose, flexible methodological approaches, as well as statistical models have been proposed to assess unidimensionality of health measures (eg. a bi-factor model) [41].

A brief, specific and unidimensional method of assessment of the severity of anxiety symptoms in older adults seems to be the answer to the main challenges posed by the measurement of anxiety in this population [13].

Currently, an Italian shortened unidimensional version of the GAS has not been developed yet, despite the current trend of growing percentage of elderly among the Italian general population and the increasing burden of mental disorders, including anxiety.

Therefore, the aim of the present study was to validate the Italian short-form of the GAS-10 in clinical and non-clinical samples. We conducted a series of separate analyses. First, we analysed the properties of the Italian version of the GAS-10 items by assessing within the IRT framework (*Structural analyses*). Next, in two independent samples of elderly people we investigated the concurrent validity of the Italian GAS-10 through a comparison with other validated instruments (*Concurrent validity analysis*). Finally, we tested the diagnostic

performance of the GAS-10 with respect to Generalized Anxiety Disorder (GAD) (*Diagnostic performance analysis*).

Materials and methods

Samples

This study investigated the properties of GAS-10 administered to three independent samples: a community-dwelling older adults (samples 1 and 2) and a sample of geriatric outpatients with GAD diagnosis (sample 3). Samples characteristics were detailed in Table 1.

Sample 1

The first sample consisted of 1200 Italian community-dwelling middle-aged and elderly adults (53.8% male, $N=645$). Mean age of the total sample was 65.21 years ($SD=9.19$; range = 50 to 92 years). Mean age for males and females was 64.94 ± 9.03 and 65.45 ± 9.32 ($t(1, 192) = .954$; $p = .341$), respectively.

Sample 2

The second community sample included 229 community-dwelling elderly participants, of whom 46.72% were males. They were, on average, 65.01 ($SD = 9.0$) years.

In both the samples, subjects with cognitive and daily living skills impairment were excluded from the samples, by using the Mini Mental State Examination (MMSE); cut-off MMSE > 24 [42], the Activities of Daily Living (ADL) [43], and the Instrumental Activities of Daily Living (IADL) [44], respectively. Recruitment was through advertisements at several older adult associations. Study participants contributed voluntarily and anonymously; no honorarium was given for completing the assessments. All participants provided written informed consent, and filled out the questionnaires in-person using paper-and-pencil forms.

The methods inherent the assessment of the community sample were approved by the IRB of the Department of Psychological Sciences, Health and Territory, University of Chieti, Italy.

Sample 3

The clinical sample was composed of 35 elderly outpatients (74.28% were females; mean age of 74.89 ± 6.9 years) with Generalized Anxiety Disorder (GAD) according to DSM-TR criteria [40]. Cognitive functions and the ability to perform activities of daily living were preserved in the present clinical sample. They were recruited in the Psychiatric Clinic of the University Hospital of Udine, Italy, and diagnosed with the Structured Clinical Interview for Axis I DSM-IV disorders (SCID-I) [45]. Diagnoses were confirmed by the clinical consensus of two staff psychiatrists.

Table 1 Characteristics of the samples

<i>Samples</i>	Structural analysis	Concurrent and diagnostic performance analysis	
	Community (<i>N</i> = 1200)	Community (<i>N</i> = 229)	Clinical (<i>N</i> = 35)
Female	645	122	26
Male	555	107	9
Age - <i>M</i> ± <i>SD</i> (Range)	65.21 ± 9.19 (50–92)	65.01 ± 9.06 (53–91)	74.89 ± 6.99 (62–89)
Marital status			
Single	90	9	1
Married	859	173	19
Divorced/separated	77	11	3
Cohabiting	10	2	2
Widowed	151	33	10
<i>missing</i>	13		
Education			
None	23	6	–
Primary school	268	42	11
Lower secondary school	301	54	11
Uppersecondary school	405	92	9
Bachelor/Master/Doctorate	155	28	4
<i>missing</i>	48		
	<i>M</i> ± <i>SD</i>	<i>M</i> ± <i>SD</i>	
ADL (Range 0–6)	5.88 ± 0.52	5.90 ± 0.47	5.74 ± 0.51
IADL (Range 0–8)	7.27 ± 1.18	7.35 ± 1.04	6.94 ± 1.59
MMSE	27.63 ± 2.21	27.37 ± 2.52	28.69 ± 1.66

MMSE Mini Mental State Examination, SF12 – PCS Physical Component Score; SF12, MCS12 Mental Component Score, ADL Activities of Daily Living, IADL Instrumental Activities of Daily Living

***p* < .001; **p* < .01

The assessment of the clinical sample was approved by the Ethics Committee of Friuli Venezia-Giulia Region, as part of a general study on the assessment of geriatric patients and their caregivers accessing the Psychiatric Clinic of the University Hospital of Udine, Italy (CEUR-2018-Sper-67-ASUIUD).

All participants provided their written informed consent to participate in this study.

Instruments

Geriatric Anxiety Scale

The GAS long-form [14] Italian version [22, 27] is a 30-item self-report measure of anxiety symptoms among older adults. Participants were asked to indicate how often they have experienced each symptom during the immediately preceding week, including today. Response format uses a 4-point Likert scale ranging from 0 (not at all) to 3 (always), with higher scores indicating higher levels of anxiety. The GAS-long form covers three different domains of anxiety, which are common among older adults: somatic symptoms, cognitive symptoms, and affective symptoms. The GAS total score is based on the first 25 items. The additional 5 content items investigate

areas of anxiety often reported to be of concern for older adults (eg. health and financial concerns). Cronbach's alpha for the GAS-long form in the present study was good (.90).

A 10-item shortened version, the GAS-10, was developed, using Item Response Theory (IRT) approach, to reduce the burden of administration and scoring time, as well as to pose less burden on respondents [46]. The GAS-10 has strong evidence of reliability and validity for use with diverse samples of community-dwelling, medically ill, and treatment-seeking older adults [32, 33]. In the present study, Cronbach's alpha was .84.

Geriatric Anxiety Inventory– short form (GAI-SF)

The GAI-SF [47] is a five-item self-report measure used to assess anxiety symptom severity among older adults. Respondents were asked to rate their agreement or disagreement with each item (yes/no response format). The measure was developed for mild cognitively impaired older adults and was derived from the 20-item GAI long form [47]. In the present study, Cronbach's alpha was .70.

Geriatric Depressive Scale-15 (GDS-15)

Depressive symptoms were assessed by self-report with a visiting nurse assisted to read the questionnaire using the 15-item version of the GDS [48]. The GDS-15 is consisted of 15 items in a “yes/no” response format. Higher score indicates more severe depressive symptoms. In the current study, the internal consistency was high, with a Cronbach’s alpha coefficient of .78.

Short-Form 12-Item Health Survey (SF-12)

The Medical Outcomes Study Short-Form 12-Item Health Survey (SF-12) [49]; is a 12-item measure, assessed health related quality of life. The SF-12 is a shorter version of the SF-36 and has demonstrated similar psychometric properties [50]. It has also been validated to be used with older adults [51]. The measure includes two subscales, a Physical Component Summary score (PCS) and Mental Health Component (MCS) summary score, with higher scores indicating better physical and mental health, respectively.

Statistical analysis

In the precision analysis on sample 1, construct unidimensionality was assessed through confirmatory factor analysis (CFA). The one-factor CFA model of the GAS-10 using the mean and variance-adjusted robust weighted least squares (WLSMV) estimator was evaluated using the Mplus 7 software [52]. The goodness of fit of the model was evaluated based on several fit statistics: (a) robust WLSMV chi square (χ^2) statistic and its degrees of freedom; (b) Tucker Lewis Index (TLI); (c) comparative fit index (CFI); (d) root mean square error of approximation (RMSEA) and its 90% confidence interval (90% CI), and Weighted Root Mean Square Residual (WRMR). Due to the large sample size, interpretation of the robust WLSMV chi-square square as a measure of fit was eschewed. Values close to .06 for the RMSEA are indicative of a good fit, between .06 and .08 as moderate fit and values larger than .10 are indicative of a poor fit [53]. For the CFI and TLI, values of .95 or above indicate a good fit, whereas values between .90 and < .95 are taken as marginally acceptable fit [54].

According to the GAS-10 polytomous response format, and in line with the Mueller et al. [32] study, a Graded Response Model (GRM) [55]; was chosen. Although a number of different IRT models exists, the GRM logistic model assumes a single underlying trait (θ) and two item parameters: the difficulty parameter or threshold category (b) and the discrimination parameter (a). Under IRT, the difficulty of an item describes where the item functions along the trait, and it can be interpreted as a location index with regard to the trait being measured (the anxiety severity trait, in our case). Discrimination parameter, instead, describes how well an

item can differentiate among adults with different levels of anxiety severity (θ). TIF and the related standard errors of measurement (SE) indicating the precision of the whole test [56] were also estimated to determine at what level of geriatric anxiety the GAS-10 provides the most information. The parameters were estimated by employing the Marginal Maximum Likelihood Estimation (MMLE) method with the Expectation–Maximization algorithm (EM) [57]. All IRT analyses were conducted using the *mirt* package in R [58].

Local independence (LI) (the assumption that the only influence on an individual’s item response is that of the latent trait variable being measured and that no other variable is influencing individual item responses) was also assessed. A violation of the LI assumption can distort estimated item parameters, item standard errors, and model-fit statistics [59, 60]. To assess the tenability of LI, the standardized LD- χ^2 statistic, a local dependence matrix based on the χ^2 statistics [61], for each item was examined. Standardized LD residuals were expressed in the form of signed Cramers V coefficients. A value of |10| or greater is considered large and reflecting LI issues [52]. LD- χ^2 statistics were calculated using the residuals function in *mirt*.

Logistic regression was used to investigate DIF based on the latent anxiety estimates derived from IRT using the *lordif* package in R [57]. Effect size estimation for DIF was computed using the expected test score standardized difference (ETSSD) index [62], implemented in *mirt* package with the empirical_ES function. The ETSSD D can be interpreted using the guidelines on effect size given by Cohen [62, 63].

In order to test the adequacy of the model we computed the C_2 fit statistic [64], available using the $M2$ function in the *mirt* package. The C_2 is a limited-information goodness of fit test statistic for ordinal IRT models, and is mainly useful when the number of items is small and the number of categories is large. The goodness of fit of the GRM model was evaluated based on several fit statistics (such as CFI, TLI, RMSEA, and SRMR) that are strictly related to those in structural equation modeling (SEM and CFA) (*Model Diagnostic*). Information curves, ICCs, threshold parameters, and discrimination parameters were analyzed to examine the item properties of the measure and identify which items were more or less useful in reliably measuring trait levels of anxiety. Discrimination parameter (α) values theoretically range from $-\infty$ to ∞ , but from .50 to 3.0 represents a reasonably good range. Difficulty parameter (b) values ranging from -3 to $+3$ are in the typical range; and -2.8 to $+2.8$ are in the usual range [65].

The concurrent validity and diagnostic accuracy of the GAS-10 were conducted using two independent samples of elderly adults, a sample of community-dwelling

healthy adults and a outpatient sample of adults with GAD (sample 2 and 3). To investigate concurrent validity of test score interpretations, Pearson correlations were calculated between scores on the GAS-10 and scores on the GAS-long form, the GAI-SF, the GDS-15, and the SF-12. In order to provide an accurate estimation of the associations, the GRM based latent trait scores were estimated for the GAS-10 concurrent validity analysis. We also compared the GAS-10 and GAS-long form pairs of correlation coefficients in the analysis of concurrent and discriminant validity following Meng et al. [66]. This procedure involves performing a Fisher Z transformation on the correlation coefficients so that they can be compared via a t-test.

The diagnostic performance of the 10-item GAS was assessed using the Area Under (AUC) the Receiver Operating Characteristic curve (ROC). The Youden (J) method was employed in order to detect the cut-off score of the GAS-10. ROC curve analysis was done using the MedCalc software package [67]. Optimal values of AUC ranged from 0 “weak performance” to “1” perfect performance”, with a values of 0.7 to 0.8 are considered “fair,” 0.8 to 0.9 “good,” and ≥ 0.9 are considered “excellent” [68]. Belonging to the non-clinical and clinical sample (dichotomized as 0/1), and the GAI-SF cut-off score (> 2) [69] were employed to classify participants with low and high anxiety symptoms. We also computed a series of pairwise comparison of ROC curves to test whether the GAS long and short forms differed in performance across diagnoses.

Results

Structural analyses

As a preliminary step, unidimensionality of the geriatric anxiety construct was assessed on sample 1 with a CFA using the WLSMV. Results showed that the 10 GAS items measured one dimension, WLSMV χ^2 (35; $p < .001$) = 329.683, CFI = .95, TLI = .95, and RMSEA = .08 CI = [.076, .092], WRMR = 1.63. All factor loadings were significant ($p < .001$), ranging from .47 to .82. Having verified the assumptions that a single continuous construct accounted for the covariation between item responses, unidimensional GRM-IRT analyses were performed. The GRM was applied in order to estimate the item discrimination (*a*) and difficulty (*b*) parameters, in line with the Mueller et al. study [32]. The range of item discrimination (α) was between $0.941 \pm .09$ and $2.618 \pm .22$ logit (Table 2). The variation in discrimination parameter suggested that a GRM estimating unique slope parameters for each items was good for our data. Thus, the items can adequately distinguish among individuals with different levels anxiety. In line with the discrimination parameters found in Mueller et al. [32], items from somatic subscales (eg. item #4,

Table 2 Item discrimination (*a*) and difficulty parameter (*b*) estimates along with standard errors ($\pm SE$) for each item of the GAS-10

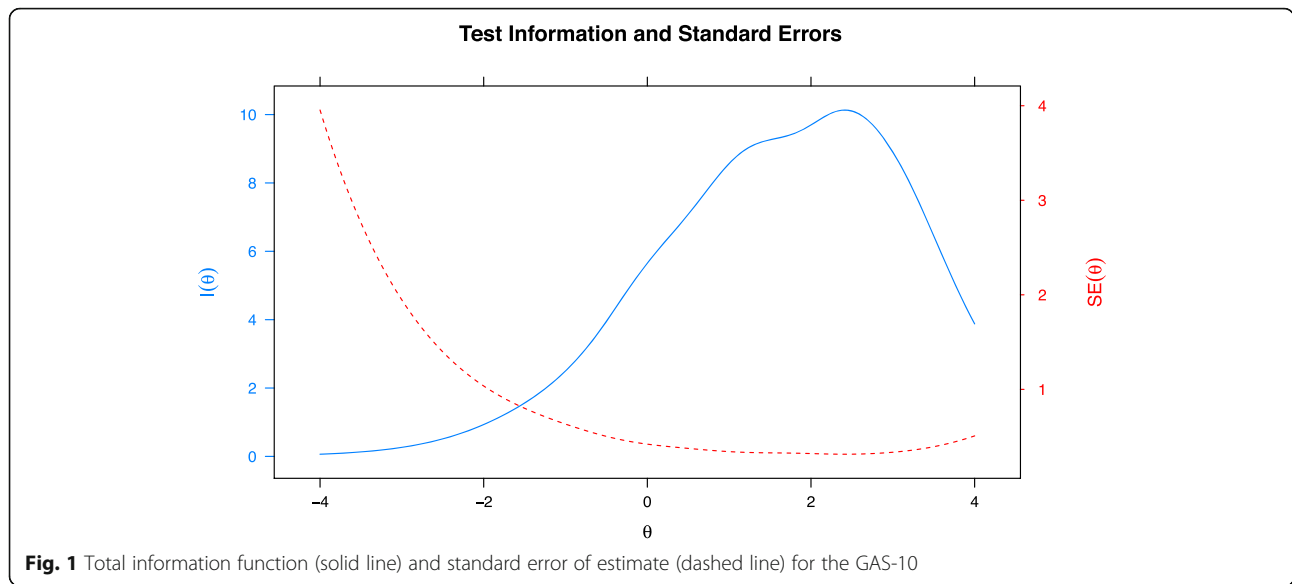
GAS-10 Item	<i>a</i> $\pm SE$	<i>b</i> ₁ $\pm SE$	<i>b</i> ₂ $\pm SE$	<i>b</i> ₃ $\pm SE$
Item #4	0.941 $\pm .09$	1.538 $\pm .14$	3.271 $\pm .30$	4.448 $\pm .43$
Item #1	1.446 $\pm .10$	-0.619 $\pm .06$	2.267 $\pm .14$	3.589 $\pm .25$
Item #8	1.496 $\pm .17$	-0.088 $\pm .04$	1.658 $\pm .08$	2.691 $\pm .13$
Item #7	1.527 $\pm .16$	-1.246 $\pm .06$	1.092 $\pm .12$	2.290 $\pm .24$
Item #2	1.528 $\pm .12$	0.836 $\pm .07$	2.660 $\pm .17$	3.843 $\pm .29$
Item #5	1.750 $\pm .10$	0.321 $\pm .05$	1.905 $\pm .10$	2.851 $\pm .16$
Item #10	2.163 $\pm .16$	0.992 $\pm .06$	2.157 $\pm .11$	2.786 $\pm .16$
Item #3	2.193 $\pm .10$	0.869 $\pm .08$	2.309 $\pm .07$	3.439 $\pm .13$
Item #6	2.508 $\pm .12$	-0.012 $\pm .05$	1.592 $\pm .10$	2.512 $\pm .17$
Item #9	2.618 $\pm .22$	1.230 $\pm .06$	2.401 $\pm .13$	3.068 $\pm .20$

$a = .940 \pm .09$) tended to discriminate less than affective and cognitive subdomains. The three *b* difficulty parameters indicated noticeable moderate to large increase in the level of the latent trait at each subsequent response category; that is, all items showed higher values in the level of the latent trait (anxiety) at each subsequent response category. For all items, the *b*₁, *b*₂, and *b*₃ values were quite evenly spaced, with *b*₁ the close to the mean trait (fixed at M = 0.00, SD = 1.00, by default), *b*₂ and *b*₃ above it. Specifically, the *b*₂ values were close to the 2SD above the mean trait level, the *b*₃ values were close to the 3SD above the mean trait level (Table 2). In sum, the GAS-10 items perform well in measuring a large spectrum of the underlying construct.

The goodness of fit for GRM model indicated a good-to-acceptable model fit for the total sample, C_2 (35; $p < .001$) = 247.1031, CFI = .97, TLI = .96, and RMSEA = .07 CI = [.062, .079], SRMSR = .04. Using guidelines developed in the context of linear factor analysis and structural equation modeling for continuous data, the SRMSR ($\leq .05$), TLI and CFI ($\geq .95$) indexes were above the threshold of the acceptability, and the RMSEA within the acceptability range ($.05 < RMSEA < .08$) [54].

Moreover, in order to identify the area of trait that is accurately assessed by the scale, the TIF was analyzed. Figure 1 displays the TIF for the 10-item GAS scale. The TIF shows the GAS-10 provided a relative uniform information (information $\cong 8.9$) for the range 0.5 to 3 logits, which has an associated empirical reliability of about .81 and standard error of estimate of about $\cong .33$.

Then, the GAS-10 was tested evaluating the presence of LI, to assess an excess of covariation among item responses that is not accounted for by a unidimensional IRT model. Results showed that a single factor model adequately represented the structure of the GAS-10 given that none of the LD statistics were greater than |.3|.



DIF was examined for each item across gender (males vs females) and age (under70 vs over70) groups, performing an ordinal (common odds-ratio) logistic regression DIF analysis using IRT theta (θ) estimates as the conditioning variable. Results showed no DIF for gender. Concerning the age groups, three items displayed DIF (#1: “I was irritable”, #3: “I felt like I was in a daze”, #6: “I felt restless, keyed up, or on edge”). However, an examination of the McFadden’s pseudo R^2 criterion showed that the magnitude of DIF for these items was negligible (close to .02) [64]. At the overall test level there is minimal difference in the total expected score at any anxiety level for under or over 70 aged participants, with a small effect size (ETSSD = .23).

Concurrent validity analysis

Differences in anxiety, and depression variables, as well as in health dimensions, were found among non-clinical (sample 2) and clinical samples (sample 3). Means and standard deviations, group *t*-tests were provided for each

independent variable in Table 3. Pairwise deletion techniques were applied to handling missing data. As expected, clinical adults exhibited significantly higher levels of anxiety, as measured by GAS-long form and the GAS-10 [$t(262) = -4.80/-4.28, p < .001$]; GAI-SF [$t(262) = -6.01, p < .001$], and depression [GDS-15; $t(262) = -6.56, p < .001$]. On the other hand, non-clinical sample showed high level of physical [PCS; $t(262) = 2.69, p < .01$] and mental [MCS; $t(262) = -6.01, p < .001$] health compared to the clinical sample. Both the GAS-long form and the GAS-10 scales were highly inter-correlated ($r = .806, p < .001$). Additionally, GAS-10 was moderately correlated with the GAI-SF ($r = .570, p < .001$), the GDS-15 ($r = .540, p < .001$) scales, and the Mental Component Score of the SF-12 scale ($r = -.418, p < .001$). No significant correlation was found between the GAS-10 and the Physical Component Score of the SF-12. Subsequently, correlation coefficients between the GAS-long form and GAS-10 with GAI-SF, and GDS-15 were statistically compared (Table 3) [66]. Comparisons

Table 3 Descriptive, differences and correlations among GAS-long form, GAS-10, GAI-SF, GDS, and SF-12

	Community (N = 229) M ± SD	Clinical (N = 35)	t-test	Correlations				
				2.	3.	4.	5.	6.
1. GAS-Long form	15.18 ± 9.26	23.60 ± 12.01	-4.80**	.806**	.558**	.523**	-.036	-.360**
2. GAS-10	5.04 ± 4.07	8.46 ± 6.07	-4.28** ^a		.540**	.570**	-.056	-.418**
3. GDS-15	3.59 ± 2.95	7.17 ± 3.38	-6.56**			.496**	.070	-.356**
4. GAI-SF	2.06 ± 1.56	3.71 ± 1.20	-6.01**				.004	-.440**
5. SF-12 PCS	41.12 ± 6.33	37.94 ± 6.37	2.69*					-.345**
6. SF-12 MCS	41.53 ± 6.52	36.24 ± 5.83	4.40**					

GAS Geriatric Anxiety Scale, GDS Geriatric Depression Scale, SF12 – PCS Physical Component Score, SF12 - MCS12 Mental Component Score

** $p < .001$; * $p < .01$

^at-test based on latent scores

revealed that the GAS-long form and the GAS-10 correlated equally with GAI-SF scores [$t(261) = 1.49$, $Z = 1.48$], as well as with the GDS-15 [$t(261) = .57$, $Z = .57$].

Diagnostic performance analysis

A first ROC curve analysis was performed to compare the elderly adults with high versus the low anxiety symptoms risk group based on the clinical/non-clinical sample (samples 2 and 3). The results indicated that the 10-item GAS scale was able to discriminate between the two groups (Table 4). The AUC for the GAS-10 total score was .702 (95%CI = .643–.757), suggesting fair discrimination between the groups. The *J* index of .35 for the GAS-10 total score was observed at a score of 6 points, corresponding to sensitivity of 62.86% and specificity of 72.93%. The positive and negative predictive powers were 26.2 and 92.8%, respectively.

Similarly, a second ROC curve was performed to compare the older adults with low versus high anxiety symptoms as assessed by the GAI-SF cut-off (> 2) [69]. The results indicated that the GAS-10 scale was able to discriminate the two groups with an AUC of .745 (95%CI of .688–.797). The *J* index of .37 for the GAS-10 total score was observed at a score of 6 point, corresponding to sensitivity of 53.64% and specificity of 83.77%. The positive and negative predictive powers were 70.2 and 71.7%, respectively. Thus, the GAS-10 cut-off score remained stable at 6 points despite the different classification criteria applied. This cut off point provided a balanced ratio between sensitivity and specificity, and correctly diagnosed elderly adults at high risk of anxiety symptoms compared to those with a low anxiety symptoms risk.

Far from what has been observed with the GAS-10, diagnostic accuracy of the GAS-long form were achieved with different cut-off scores in our sample (> 13 and > 15), displaying high sensitivity values. The high sensitivity of the GAS-long form remained high even if additional cut-off score in literature was applied (> 9) [70]. When we compared the predictive validity of the total scores of the GAS-long form and the GAS-10 model, the AUCs were quite identical ($\Delta AUC = .026$), and the differences between them were not found to be significant for those were diagnosed with GAD or for those were healthy. Subsequently, when we considered the GAI-SF as a diagnostic criterion, the predictive validity of the total scores of the GAS-long form and the GAS-10 model differed significantly at value of $p < .05$ ($\Delta AUC = .052$). Thus, results of the pairwise comparison revealed that GAS-10 did not differ in diagnostic accuracy from its long form, when classification criterion was based on sample characteristics (clinical vs non-clinical). Slightly differences were displayed between them, if GAI-SF measure was applied to classify elderly adults with high from low anxiety symptoms.

Discussion

Given the overlapping of symptoms of mental disorders (eg. depression and anxiety) in later life and the comorbidity with physical health problems, it is important to design unidimensional anxiety measures that have been specifically targeted for older adults that account for and measure this unique expression [13, 23]. This is also supported by the controversial findings derived from empirical studies on the composite ‘general distress score’, which combines depression and anxiety symptoms in a unidimensional domain [71]. This study aims

Table 4 Area Under the Curve (AUC) of the Receiver Operating Characteristic Curve (ROC) Analyses for the different versions of the GAS, and Comparison of independent ROC curves

Criteria	AUC	SE	(95%CI)	Sensitivity/ Specificity	Cut-off	PPV/ NPV
Clinical vs NonClinical						
GAS-long form	.729	.045	.672/.782	91.43/28.82	> 9 ^a	16.4/95.7
				88.57/51.53	> 13	21.8/96.7
GAS-10	.702	.046	.643/.757	62.86/72.93	> 6	26.2/92.8
ROC curve comparison	ΔAUC	SE	(95%CI)	<i>z</i>	<i>p</i>	
GAS-long form vs GAS-10	.026	.035	-.041/.095	.780	.435	
GAI-SF > 2						
GAS-long form	.797	.028	.743/.844	91.82/38.96	> 9 ^a	51.8/87.0
				71.82/75.97	> 15	68.1/79.1
GAS-10	.745	.030	.688/.797	53.64/83.77	> 6	70.2/71.7
ROC curve comparison	ΔAUC	SE	(95%CI)	<i>z</i>	<i>p</i>	
GAS-long form vs GAS-10	.052	.0233	.006/.097	2.226	.026	

^a cut-off identified by Gould et al., 2014 [70]

to filling this gap, adding evidence of the adequacy of the GAS-10 for the assessment of geriatric anxiety through IRT-based evidence.

The IRT analysis provided a clear framework of the good performance of the GAS-10 items in measuring the geriatric anxiety construct. In accordance with Muellet et al. [32] study, our study showed that the GRM model accurately explained the pattern of responses obtained by the GAS-10 scale. In deep, the item parameters showed that all the GAS-10 items were able to distinguish adequately among older adults with different levels of the trait being measured, and adequately covered the spectrum of the latent trait. Concerning the item difficulty values, our findings showed that the GAS-10 items had a medium to high level of difficulty, indicating a decreasing of the probability to endorse the response option “*all the time*” to self-statements that quantify the frequency that anxiety symptoms. Nonetheless, the discrimination parameter values in the present study ranged from “low” (item #4) to “moderately high” (item #9), suggesting that the GAS-10 items can discriminate older adults with different anxiety symptoms severity levels. Compared to the cognitive and affective items, the somatic items were found less informative. These findings had already been reported in previous studies employing both the GAS-10 [32] and other anxiety measures, such as the State-Trait Inventory for the Cognitive and Somatic Anxiety (STICSA) [21]. It is still open the debate among researchers about the inclusion of physiological symptoms items in anxiety measures, mainly in the elderly, since these could undermine the uniformity of the anxiety construct. Indeed, the content of somatic items in anxiety measures have been criticized due to the possible overlap with the depression symptomatology [30]. In fact, somatic symptoms (eg. fatigue, aches and pains) have been found to be a prominent part of the clinical presentation in depressed older adults [72]. In addition, the co-occurrence of physical health problems in older adults makes differentiating between somatic symptoms of anxiety, and depression difficult [73].

The GAS-10 adequately measured geriatric anxiety ranging from medium to high levels, whereas it was less precise for the lowest levels of the trait. To date, the highest informative point of the GAS-10, or the peak, was observed at 2.5 standard deviation above the mean level of anxiety, supporting its use as a clinical meaningful measure to assess elderly who experienced high levels of mental distress.

Far from previous study the DIF analysis for each item across gender confirmed the invariance of the GAS-10, providing evidence of its ability to scale males and females onto a common metric. Commonly, females tend to score highly in the GAS total score than males, since they are more comfortable in to express mental health

symptoms, as well as showed more risk factors [74, 75]. Potentially, this unbiased version of the GAS-10 could allow an easy and efficient assessment of anxiety among elderly, without the use of differentiated norms by gender. On the contrary, DIF has been found across age, despite its impact was negligible and close to the 2% of the items true score. Adults aged over 70 years scored higher on three items, two of these were drawn from the affective and one from the cognitive domain of the GAS-10. DIF of affective anxiety items in older adults, as well as in the cognitive item measuring daze/confusion, were likely to affect by the higher prevalence of depressed mood symptoms [76].

Additionally, the GAS-10 has been administered to test the validity and the diagnostic accuracy of the scale in two samples of healthy older adults and outpatients with GAD diagnosis. Similar to the GAS-long form, independent *t*-test analysis revealed that the GAS-10 was able to capture differences in anxiety symptoms between the two samples. As expected, outpatients with GAD diagnosis displayed statistically higher mean scores of the GAS-10, compared to the healthy sample. Results from correlation analysis, also, provided good support for the validity of the GAS-10. The brief and the long version of the GAS were found highly interrelated, and pairwise comparison analysis corroborated the high degree of overlap between them in correlating with concurrent validity measures.

In line with Gottschling et al. [31], both the GAS-10 and its long form were found to correlate on average with depression (as measured by the GDS), thus confirming the co-occurrence of anxiety and depression symptomatology among elderly. In addition, this finding was not surprising, since the GDS has been criticized for its lack of unidimensionality, and construct validity [13].

The diagnostic performance of the GAS-10 was assessed by means of ROC analysis to detect elderly with clinically significant GAD symptoms, as well as with significant anxiety symptoms, as measured by the GAI-SF cut-off scores. Differences in diagnostic accuracy between the long form version of the GAS and the GAS-10 was not found significant, when clinical/non-clinical classification variable has been taken into account. An optimal cut-off value score >6 for the GAS-10 was selected to maximize the sum of sensitivity and specificity. Therefore all the patients with a total score of 6 in the GAS-10 should be referred for further risk assessment and management. When GAI-SF cut-off has been implemented as classification variable in the ROC analysis, the GAS-10 cut-off remained unchanged, preserving a discrete specificity. Our findings are in line with the previous studies, in which a cut-off score of >9 optimized the balance of sensitivity and specificity (.60 and .75), for the GAS –long form [70]. Likewise, diagnostic

accuracy of the GAS- G-SF in clinical settings was low, showing a sensitivity of 70% and a specificity of 30% (cut-off > 5) [31].

Nevertheless, the Italian version of the GAS-10 could be considered a stable measure of geriatric anxiety, with a sufficient discriminant validity in categorizing outpatients diagnosed with GAD, and adults with clinical significant anxiety symptoms (as classified by the cut score of the GAI-SF). Although its use as a screening test might be limited, the identification of a clinical cut-off score of the GAS-10 could help clinicians and researchers to identify older adults with anxiety disorders in resource-constrained settings, where time constrain and fatigue issues affect clinical assessment validity. In this direction, future studies with larger clinical samples are needed to support the clinical utility of the GAS-10.

Some limitations of this study could be addressed in future research. First, participants in the present study were community-dwelling participants and outpatients' elderly with GAD diagnosis. There is also evidence that medically ill samples of older adults experienced severe levels of anxious distress [23, 77], and were most likely to overburden healthcare services [46, 78, 79]. Thus, further studies could address GAS-10 diagnostic accuracy in samples of older adults with GAD diagnosis in comorbidity with medically ill patients. A second limitation was the high prevalence of females in our clinical sample. This datum, however, was similar to real life, since being female gender was found one of the principal factor associated with the pure GAD [80, 81]. Another limitation of this study was its cross-sectional design which limits causal inferences, and make difficult to understand if the GAS-10 items are a stable anxiety symptoms measure. Further longitudinal studies are required in this direction.

Conclusion

Our study provided support for the psychometric functioning of the GAS-10, and its diagnostic performance in discriminate accurately older adults with GAD from elderly adults without significant clinically relevant symptoms of anxiety. In this sense, the present study improved upon previous studies by using a rigorous psychometric technique to validate the short measurement, the Item Response Theory (IRT) [82].

Indeed, differently from the Classical Test Theory (CTT), the IRT considers the characteristics of both the test and the sample, discarding items which are not endorsed by respondents and thus do not contribute significantly to the latent-attribute scores [37, 56, 83, 84].

Through the use of IRT, this study provided additional information about the psychometric functioning of the GAS-10 and some cues for improving its clinical utility.

Our results could increase the confidence in the assessment accuracy of the GAS-10 and give support to the huge amount of studies that employed it in the field of geriatric anxiety.

Acknowledgments

Not applicable.

Authors' contributions

LC, MB and MB contributed equally to the design of the study and supervision of data collection. LC, MB and MB all contributed to drafting of the manuscript. AM, NC, EM collected the data. LC and MB performed statistical analyses. All authors have contributed to and have approved the final manuscript.

Funding

This research did not receive any grants from public or private agencies.

Availability of data and materials

The datasets used during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was conducted in compliance with the Declaration of Helsinki and with Italian data protection regulations. Written informed consent was obtained from all participating GPs and patients prior to their inclusion. The assessment of the community sample was approved by the IRB of the Department of Psychological Sciences, Health and Territory, University of Chieti, Italy.

The assessment of the clinical sample was approved by the Ethics Committee of Friuli Venezia-Giulia Region, as part of a general study on the assessment of geriatric patients and their caregivers accessing the Psychiatric Clinic of the University Hospital of Udine, Italy (CEUR-2018-Sper-67-ASUIUD).

Consent for publication

Not applicable.

Competing interests

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

Author details

¹Department of Psychological Sciences, Health and Territory, "G. d'Annunzio" University of Chieti-Pescara, via dei Vestini, 31 66100 Chieti, Italy. ²Unit of Psychiatry, Department of Medicine (DAME), University of Udine, Udine, Italy. ³Psychiatric Clinic, Healthcare and University Integrated Trust, Udine, Italy.

Received: 23 March 2021 Accepted: 7 June 2021

Published online: 30 June 2021

References

- Balestrieri M, Isola M, Quartaroli M, Roncolato M, Bellantuono C. Assessing mixed anxiety-depressive disorder. A national primary care survey. *Psychiatry Res.* 2010;176(2–3):197–201. <https://doi.org/10.1016/j.psychres.2008.11.011>.
- Andreas S, Schulz H, Volkert J, Dehoust M, Sehner S, Suling A, et al. Prevalence of mental disorders in elderly people: the European MentDis_ ICF65+ study. *Br J Psychiatry.* 2017;210(2):125–31. <https://doi.org/10.1192/bjp.bp.115.180463>.
- Bryant C, Jackson H, Ames D. The prevalence of anxiety in older adults: methodological issues and a review of the literature. *J Affect Disord.* 2008; 109(3):233–50. <https://doi.org/10.1016/j.jad.2007.11.008>.
- Preti A, Demontis R, Cossu G, Kalcev G, Cabras F, Moro MF, et al. The lifetime prevalence and impact of generalized anxiety disorders in an epidemiologic Italian National Survey carried out by clinicians by means of semi-structured interviews. *BMC Psychiatry.* 2021;21(1):1–8.

5. ISTAT. Indicatori demografici 2019. Rome: Istituto Nazionale di Statistica; 2020. Retrieved from: <https://www.istat.it/it/archivio/238447>.
6. Organization WH. World report on ageing and health. Geneva: World Health Organization; 2015.
7. Balestrieri M, Baldacci S, Bellomo A, Bellantuono C, Conti L, Perugi G, et al. Clinical vs. structured interview on anxiety and affective disorders by primary care physicians. Understanding diagnostic discordance. *Epidemiol Psychiatr Sci*. 2007;16(2):144–51. <https://doi.org/10.1017/S1121189X00004772>.
8. Koychev I, Ebmeier KP. Anxiety in older adults often goes undiagnosed. *Practitioner*. 2016;260(1789):17–20.
9. Schuurmans J, van Balkom A. Late-life anxiety disorders: a review. *Curr Psychiatry Rep*. 2011;13(4):267–73. <https://doi.org/10.1007/s11920-011-0204-4>.
10. An MH, Park SS, You SC, Park RW, Park B, Woo HK, et al. Depressive symptom network associated with comorbid anxiety in late-life depression. *Front Psychiatry*. 2019;10:856.
11. El-Gabalawy R, Mackenzie CS, Shoostari S, Sareen J. Comorbid physical health conditions and anxiety disorders: a population-based exploration of prevalence and health outcomes among older adults. *Gen Hosp Psychiatry*. 2011;33(6):556–64. <https://doi.org/10.1016/j.genhosppsych.2011.07.005>.
12. Ramos K, Stanley MA. Anxiety disorders in late life. *Psychiatr Clin*. 2018;41(1):55–64.
13. Balsamo M, Cataldi F, Carlucci L, Fairfield B. Assessment of anxiety in older adults: a review of self-report measures. *Clin Interv Aging*. 2018;13:573–93. <https://doi.org/10.2147/CIA.S114100>.
14. Segal DL, June A, Payne M, Coolidge FL, Yochim B. Development and initial validation of a self-report assessment tool for anxiety among older adults: the geriatric anxiety scale. *J Anxiety Disord*. 2010;24(7):709–14. <https://doi.org/10.1016/j.janxdis.2010.05.002>.
15. Reynolds CR, Richmond BO, Lowe PA. The adult manifest anxiety scale (AMAS): manual. Torrance: Western Psychological Services (WPS); 2003.
16. Pachana NA, Byrne GJ, Siddle H, Koloski N, Harley E, Arnold E. Development and validation of the geriatric anxiety inventory. *Int Psychogeriatr*. 2007;19(1):103–14. <https://doi.org/10.1017/S1041610206003504>.
17. Wisocki PA, Handen B, Morse CK. The worry scale as a measure of anxiety among homebound and community active elderly. *Behav Ther*. 1986;9(5):91–5.
18. Spielberger CD. Manual for the state-trait anxiety inventory STAI (form Y): mind garden; 1983.
19. Beck AT, Steer RA. Beck anxiety inventory: BAI. San Antonio: Psychological Corporation; 1993.
20. Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Therapy: development and validation of the penn state worry questionnaire. *Behav Res*. 1990;28(6):487–95. [https://doi.org/10.1016/0005-7967\(90\)90135-6](https://doi.org/10.1016/0005-7967(90)90135-6).
21. Ree MJ, French D, MacLeod C, Locke V. Distinguishing cognitive and somatic dimensions of state and trait anxiety: development and validation of the state-trait inventory for cognitive and somatic anxiety (STICSA). *Behav Cogn Psychother*. 2008;36(3):313–32.
22. Picconi L, Balsamo M, Palumbo R, Fairfield B. Testing factor structure and measurement invariance across gender with Italian geriatric anxiety scale. *Front Psychol*. 2018;9:1164. <https://doi.org/10.3389/fpsyg.2018.01164>.
23. Segal DL, Mueller AE. Evidence of validity of the geriatric anxiety scale for use among medically ill older adults. *J Depress Anxiety Forecast*. 2019;2(1):1–3, 1010.
24. Hallit S, Hallit R, Hachem D, Nasra M, Kheir N, Salameh P. Validation of the Arabic version of the geriatric anxiety scale among Lebanese population of older adults. *J Psychopathol*. 2017;1(1):26–34.
25. Lin X-L, Lu D-L, Gottschling J, Segal DL, Tang S-Y. Validation of a Chinese version of the geriatric anxiety scale among community-dwelling older adults in mainland China. *J Cross Cult Gerontol*. 2017;32(1):57–70. <https://doi.org/10.1007/s10823-016-9302-4>.
26. Gottschling J, Segal DL, Häusele C, Spinath FM, Stoll G. Assessment of anxiety in older adults: translation and psychometric evaluation of the German version of the geriatric anxiety scale (GAS). *J Psychopathol Behav Assessment*. 2016;38(1):136–48. <https://doi.org/10.1007/s10862-015-9504-z>.
27. Gatti A, Gottschling J, Brugnera A, Adorni R, Zarbo C, Compare A, et al. An investigation of the psychometric properties of the geriatric anxiety scale (GAS) in an Italian sample of community-dwelling older adults. *Aging Mental Health*. 2018;22(9):1176–84. <https://doi.org/10.1080/13607863.2017.1347141>.
28. Bolghan-Abadi M, Segal DL, Coolidge FL, Gottschling J. Persian version of the geriatric anxiety scale: translation and preliminary psychometric properties among Iranian older adults. *Aging Mental Health*. 2013;17(7):896–900. <https://doi.org/10.1080/13607863.2013.788999>.
29. Karahan FS, Hamarta E, Karahan AY. The Turkish adaptation and psychometric properties of the geriatric anxiety scale. *Ment Illn*. 2018;10(1):1–5. <https://doi.org/10.1108/mi.2018.7580>.
30. McHorney CA. Measuring and monitoring general health status in elderly persons: practical and methodological issues in using the SF-36 health survey. *Gerontologist*. 1996;36(5):571–83. <https://doi.org/10.1093/geront/36.5.571>.
31. Gottschling J, Dörendahl J, Prell T, Karbach J, Bunz M, Kindermann I, et al. Measuring anxiety in older adults: development, diagnostic accuracy, and preliminary validation of a short-form of the German geriatric anxiety scale (GAS-G-SF). *J Pers Assess*. 2020;102(2):196–204. <https://doi.org/10.1080/00223891.2019.1677679>.
32. Mueller AE, Segal DL, Gavett B, Marty MA, Yochim B, June A, et al. Geriatric anxiety scale: item response theory analysis, differential item functioning, and creation of a ten-item short form (GAS-10). *Int Psychogeriatr*. 2015;27(7):1099–111. <https://doi.org/10.1017/S1041610214000210>.
33. Pifer MA, Segal DL. Geriatric anxiety scale: development and preliminary validation of a long-term care anxiety assessment measure. *Clin Gerontol*. 2020;43(3):295–307. <https://doi.org/10.1080/07317115.2020.1725793>.
34. Pifer MA, Segal DL, Jester DJ, Molinari V. Assessment of anxiety in long-term care residents: issues and strategies. *Int J Aging Hum Dev*. 2020;0091415020943321.
35. Balsamo M, Saggino A, Carlucci L. Tailored screening for late-life depression: A short version of the Teate Depression Inventory (TDI-E). *Front Psychol*. 2019;10:2693.
36. Reise SP, Morizot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Qual Life Res*. 2007;16(1):19–31. <https://doi.org/10.1007/s11136-007-9183-7>.
37. Reise SP, Waller NG. Item response theory and clinical measurement. *Annu Rev Clin Psychol*. 2009;5(1):27–48. <https://doi.org/10.1146/annurev.clinpsy.032408.153553>.
38. Smith GT, McCarthy DM, Zapolski TC. On the value of homogeneous constructs for construct validation, theory testing, and the description of psychopathology. *Psychol Assess*. 2009;21(3):272–84. <https://doi.org/10.1037/a0016699>.
39. Falissard B. The unidimensionality of a psychiatric scale: a statistical point of view. *Int J Methods Psychiatr Res*. 1999;8(3):162–7. <https://doi.org/10.1002/mpr.66>.
40. McKenna SP, Heaney A. Composite outcome measurement in clinical research: the triumph of illusion over reality? *J Med Econ*. 2020;23(10):1196–204. <https://doi.org/10.1080/13696998.2020.1797755>.
41. Rodriguez A, Reise SP, Haviland MG. Applying bifactor statistical indices in the evaluation of psychological measures. *J Pers Assess*. 2016;98(3):223–37. <https://doi.org/10.1080/00223891.2015.1089249>.
42. Folstein MF, Folstein SE, McHugh PR. “mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189–98. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6).
43. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index: a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. *Maryland State Med J*. 1965;14:61–5.
44. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*. 1969;9(3_Part_1):179–86.
45. First M, Spitzer R, Gibbon M, Williams J. Structured clinical interview for DSM-IV Axis I disorders (SCID). New York: New York State Psychiatric Institute; 1995.
46. Segal DL, Granier KL, Pifer MA, Stone LE. Mental health in older adults: an introduction for integrated care professionals. *Clin Integr Care*. 2020;2:100015. <https://doi.org/10.1016/j.jintcar.2020.100015>.
47. Byrne GJ, Pachana NA. Development and validation of a short form of the geriatric anxiety inventory-the GAI-SF. *Int Psychogeriatr*. 2011;23(1):125–31. <https://doi.org/10.1017/S1041610210001237>.
48. Yesavage JA, Sheikh JJ. Geriatric depression scale (GDS) recent evidence and development of a shorter version. *Clin Gerontol*. 1986;5(1–2):165–73. https://doi.org/10.1300/J018v05n01_09.
49. Ware JE Jr, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220–33. <https://doi.org/10.1097/00005650-199603000-00003>.

50. Jenkinson C, Layte R, Jenkinson D, Lawrence K, Petersen S, Paice C, et al. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Public Health*. 1997;19(2):179–86. <https://doi.org/10.1093/oxfordjournals.pubmed.a024606>.
51. Resnick B, Nahm ES. Reliability and validity testing of the revised 12-item short-form health survey in older adults. *J Nurs Meas*. 2001;9(2):151–61. <https://doi.org/10.1891/1061-3749.9.2.151>.
52. Muthén LK, Muthén BO. *Mplus User's Guide*. 7th ed. Los Angeles: Muthén & Muthén; 1998-2015.
53. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J*. 1999;6(1):1–55. <https://doi.org/10.1080/10705519909540118>.
54. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol Res Online*. 2003;8(2):23–74.
55. Samejima F. Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph*. 1969;17:1–100.
56. Reise S, Rodriguez A. Item response theory and the measurement of psychiatric constructs: some empirical and conceptual issues and challenges. *Psychol Med*. 2016;46(10):2025–39. <https://doi.org/10.1017/S0033291716000520>.
57. Bock RD, Aitkin M. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. *Psychometrika*. 1981;46(4):443–59. <https://doi.org/10.1007/BF02293801>.
58. Chalmers RP. Mirt: a multidimensional item response theory package for the R environment. *J Stat Softw*. 2012;48(6):1–29.
59. De Ayala RJ. *The theory and practice of item response theory*. New York: Guilford Press; 2009.
60. Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Qual Life Res*. 2007;16(1):5–18. <https://doi.org/10.1007/s11136-007-9198-0>.
61. Chen W-H, Thissen D. Local dependence indexes for item pairs using item response theory. *J Educ Behav Stat*. 1997;22(3):265–89. <https://doi.org/10.3102/10769986022003265>.
62. Meade A. A taxonomy of effect size measures for the differential functioning of items and scales. *J Appl Psychol*. 2010;95(4):728–43. <https://doi.org/10.1037/a0018966>.
63. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum; 1988.
64. Cai L, Monroe S. A New Statistic for Evaluating Item Response Theory Models for Ordinal Data. In: National Center for Research on Evaluation, Standards, Student Testing; 2014.
65. Toland MD. Practical guide to conducting an item response theory analysis. *J Early Adolesc*. 2014;34(1):120–51. <https://doi.org/10.1177/0272431613511332>.
66. Meng X-L, Rosenthal R, Rubin DB. Comparing correlated correlation coefficients. *Psychol Bull*. 1992;111(1):172–5. <https://doi.org/10.1037/0033-2909.111.1.172>.
67. Schoonjans F, Zalata A, Depuydt C, Comhaire F. MedCalc: a new computer program for medical statistics. *Comput Methods Programs Biomed*. 1995;48(3):257–62. [https://doi.org/10.1016/0169-2607\(95\)01703-8](https://doi.org/10.1016/0169-2607(95)01703-8).
68. Youngstrom EA. A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: we are ready to ROC. *J Pediatr Psychol*. 2014;39(2):204–21. <https://doi.org/10.1093/jpepsy/jst062>.
69. Pachana NA, Byrne GJ. The geriatric anxiety inventory: international use and future directions. *Aust Psychol*. 2012;47(1):33–8. <https://doi.org/10.1111/j.1742-9544.2011.00052.x>.
70. Gould CE, Segal DL, Yochim BP, Pachana NA, Byrne GJ, Beaudreau SA. Measuring anxiety in late life: a psychometric examination of the geriatric anxiety inventory and geriatric anxiety scale. *J Anxiety Disord*. 2014;28(8):804–11. <https://doi.org/10.1016/j.janxdis.2014.08.001>.
71. Schouten RW, Nadort E, van Ballegoijen W, Loosman WL, Honig A, Siebert CE, et al. General distress and symptoms of anxiety and depression: a factor analysis in two cohorts of dialysis patients. *Gen Hosp Psychiatry*. 2020;65:91–9. <https://doi.org/10.1016/j.genhospsych.2020.04.004>.
72. Morin RT, Nelson C, Bickford D, Insel PS, Mackin RS. Somatic and anxiety symptoms of depression are associated with disability in late life depression. *Aging Mental Health*. 2020;24(8):1225–8. <https://doi.org/10.1080/13607863.2019.1597013>.
73. Morin RT, Galatzer-Levy IR, Maccallum F, Bonanno GA. Do multiple health events reduce resilience when compared with single events? *Health Psychol*. 2017;36(8):721–8. <https://doi.org/10.1037/hea0000481>.
74. Leach LS, Poyser C, Cooklin AR, Giallo R. Prevalence and course of anxiety disorders (and symptom levels) in men across the perinatal period: a systematic review. *J Affect Disord*. 2016;190:675–86. <https://doi.org/10.1016/j.jad.2015.09.063>.
75. Jagtap N, Muliya KP, Chaturvedi SK. Depression, anxiety, and physical morbidity in women. In: Chandra P, Herrman H, Fisher J, Richer-Rössler A, editors. *Mental health illness of women Mental Health and Illness Worldwide*. Singapore: Springer; 2020. p. 259–79.
76. Byers AL, Yaffe K, Covinsky KE, Friedman MB, Bruce ML. High occurrence of mood and anxiety disorders among older adults: the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2010;67(5):489–96. <https://doi.org/10.1001/archgenpsychiatry.2010.35>.
77. Hajure M, Tariku M, Mohammedhussein M, Dule AJ. Depression, anxiety and associated factors among chronic medical patients amid COVID-19 pandemic in Mettu Karl referral hospital, Mettu, Ethiopia, 2020. *Neuropsychiatr Dis Treat*. 2020;16:2511–8. <https://doi.org/10.2147/NDT.S281995>.
78. Conlon M, Tew J, Solai LK, Gopalan P, Azzam P, Karp JF. Care transitions in the psychiatric hospital: focus on older adults. *Am J Geriatr Psychiatry*. 2020;28(3):368–77. <https://doi.org/10.1016/j.jagp.2019.07.006>.
79. Horenstein A, Heimberg RG. Anxiety disorders and healthcare utilization: a systematic review. *Clin Psychol Rev*. 2020;81(2021):101894. <https://doi.org/10.1016/j.cpr.2020.101894>.
80. Zhang X, Norton J, Carrière I, Ritchie K, Chaudieu I, Ancelin M-L. Generalized anxiety in community-dwelling elderly: prevalence and clinical characteristics. *J Affect Disord*. 2015;172:24–9. <https://doi.org/10.1016/j.jad.2014.09.036>.
81. Curran E, Rosato M, Ferry F, Leavey G. Prevalence and factors associated with anxiety and depression in older adults: gender differences in psychosocial indicators. *J Affect Disord*. 2020;267:114–22. <https://doi.org/10.1016/j.jad.2020.02.018>.
82. Smith G, McCarthy D, Anderson K. On the sins of short-form development. *Psychol Assess*. 2000;12(1):102–11. <https://doi.org/10.1037/1040-3590.12.1.102>.
83. Yang FM. Item response theory for measurement validity. *Shanghai Arch Psychiatry*. 2014;26(3):171–7. <https://doi.org/10.3969/j.issn.1002-0829.2014.03.010>.
84. Balsamo M, Carlucci L, Innamorati M, Lester D, Pompili M. Further insights into the Beck hopelessness scale (BHS): unidimensionality among psychiatric inpatients. *Front Psychiatry*. 2020;11. <https://doi.org/10.3389/fpsy.2020.00727>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

