Assessment of anxiety in older adults: A systematic review of commonly used measures

Zoe Therrien* and John Hunsley

School of Psychology, University of Ottawa, Ottawa, Canada

(Received 22 March 2011; final version received 17 June 2011)

Objectives: The authors set out to systematically review the research literature in order to identify the anxiety measures most commonly used in the assessment of older adults. Once identified, the literature was reviewed to determine the extent to which these instruments had age-relevant norms and psychometric data supporting their use with older adults.

Method: Literature searches were conducted in PsycINFO and PubMed to identify research articles in which anxiety measures were completed by older adults. After screening for suitability, a total of 213 articles were reviewed to determine the most commonly used anxiety measures with older adults to examine the psychometric properties of these instruments and to evaluate whether the instruments are appropriate for use with older adults.

Results: A total of 91 different anxiety measures were used in the 213 included articles. Twelve anxiety measures were most commonly used in the literature and of those three were specifically developed for older adults.

Conclusions: Of the most commonly used measures, the majority lacked sufficient evidence to warrant their use with older adults. Based on psychometric evidence, three measures (Beck Anxiety Inventory, Penn State Worry Questionnaire, and Geriatric Mental Status Examination) showed psychometric properties sufficient to justify the use of these instruments when assessing anxiety in older adults. In addition, two measures developed specifically for older adults (Worry Scale and Geriatric Anxiety Inventory) were also found to be appropriate for use with older adults.

Keywords: older adults; anxiety; assessment; psychometric properties; systematic review

Introduction

Over the past decade, there has been a growing emphasis on the importance of evidence-based mental health practice (American Psychological Association Presidential Task Force on Evidence-Based Practice, 2006; Institute of Medicine, 2001). According to the Institute of Medicine (2001), evidence-based practice in health care services integrates information derived from the best research evidence, clinical expertise, and the patient’s values when contemplating health care services for a patient. This framework is based on the acknowledgment that positive outcomes are less likely if the chosen service does not have a research base that shows the potential to improve the client’s functioning (Forman, Fagley, Steiner, & Schneider, 2009). The evidence-based practice movement has been endorsed through reports released by several scientific organizations in the field of health and mental health.

Much of the efforts to evaluate and disseminate evidence-based mental health practices have focused on intervention services. For example, a growing body of research has concentrated on the identification of efficacious psychological interventions (Ayers, Sorrell, Thorp, & Wetherell, 2007) and the establishment of guidelines for determining whether an intervention can be considered evidence-based (Chambless & Hollon, 1998). As assessment is a necessary component of health care services that should inform treatment selection and implementation, it seems incongruous that so little effort has been put toward identifying and promoting evidence-based assessment (Hunsley & Mash, 2007). Indeed, without scientifically sound assessment, clinicians cannot clearly evaluate a patient’s level of functioning; without this information, the development of solid case formulations (including diagnostic considerations) is not possible, and without high quality case formulations, it becomes difficult to make informed treatment choices. Moreover, in research contexts, assessment measures are used to select participants and evaluate treatment; thus, they are critical for developing evidence-based treatment (Cohen et al., 2008). In sum, whether assessment measures are used for diagnostic purposes, to select research participants, to establish case conceptualization, to inform the choice of a treatment plan, or to monitor treatment outcome, the choice of measures is central to the quality of services provided. Fortunately, some recent efforts have been made toward the development of general guidelines to be used in selecting the best instruments. These efforts emphasize, in particular, the importance of solid psychometric properties, appropriate norms, and evidence of clinical utility (Holmbeck & Devine, 2009; Hunsley & Mash, 2008).

Although research on evidence-based assessment is increasing, research on the assessment of older adults is much less developed (Ayers et al., 2007). With the growing number of older adults in the general population, there is also a concomitant rise in the number of
older adults who require mental health services. As diagnosis and treatment selection is informed by assessment data, it is necessary to have measures that are appropriate for an older population, but the lack of research evidence for the psychometric quality of many of these measures makes it challenging to choose an appropriate measure for use with older adults (Edelstein et al., 2001).

The empirical literature on anxiety prevalence suggests that it has become a widespread problem in late life. With prevalence estimates ranging from 1.2% to 15% in community samples of older adults (Bryant, Jackson, & Ames, 2009; Wolitsky-Taylor, Castriotta, Lenze, Stanley, & Craske, 2010), and from 1% to 28% in clinical samples of older adults (Bryant et al., 2009), it is more common than depression. Furthermore, among hospitalized geriatric patients, the prevalence of anxiety disorders is estimated to be as high as 43%, and known to remain high after discharge (Kvaal, Macijauskiene, Engedal, & Laake, 2001). Moreover, the prevalence of older adults with anxiety symptoms that do not meet criteria for an anxiety disorder has been found to range from 15% to 52.3% in community samples and 15–56% in clinical samples (Bryant et al., 2009). There is significant impairment and a lower level of quality of life among anxious individuals (Mendlowicz & Stein, 2000) and, according to de Beurs et al. (1999), older adults who present some anxiety symptoms are as negatively affected in their quality of life as those who meet criteria for an anxiety disorder.

Unfortunately, research on anxiety disorders in older adults has not grown at the same rate as research on anxiety disorders in younger populations (Dennis, Boddington, & Funnell, 2007). One consequence of this is that relatively little is known about the evidence-based assessment of anxiety in older adults. In both clinical and research contexts, self-report appears to be the dominant assessment method for gathering information on the experience of anxiety (Alwahhabi, 2003; Dennis et al., 2007). For self-report measures to provide valid information, information on additional factors such as the frequent comorbidity of mental and physical health problems (Cully et al., 2006; Wolitsky-Taylor et al., 2010) and the use of multiple medications in older populations must be obtained and incorporated into assessment decisions.

The assessment and diagnosis of late life anxiety is especially challenging, as symptoms of anxiety can be confused with some aspects of the normal aging process (Lenze & Wetherell, 2009) as well as with medical conditions and comorbid mental disorders (Kogan, Edelstein, & McKee, 2000). It is well-established that anxiety and depression are frequently comorbid in younger adults (de Graaf, Bijl, Spijker, Beekman, & Vollebergh, 2003; Kessler et al., 2005). In the Epidemiological Catchment Area study of people aged 18–54, 20% of individuals who received a diagnosis of any anxiety disorder in the past 6 months also received a diagnosis of some type of affective disorder (Regier et al., 1988). With regard to the older adult population, one study using a community-based sample of Canadian adults aged 55 and older found that depression was the most common comorbid disorder among those with anxiety disorders, with 23% of those with an anxiety disorder also meeting criteria for major depressive disorder (Cairney, Corna, Veldhuizen, Herrmann, & Streiner, 2008). Studies of depressed older adults also indicate that approximately half of these individuals meet the criteria for an anxiety disorder (Beekman et al., 2000).

The detection of anxiety in older adults is also complicated by the high frequency of medical disorders in this age group. With studies suggesting that between 80% and 86% of adults aged 65 and older have at least one significant medical condition (Dawson, Kline, Wiancko, & Wells, 1986; Haley, 1996; Naughton, Feely, & Bennett, 2007), many symptoms of anxiety may be overlooked or wrongly attributed to a medical illness. Older adults may be more likely to attribute physical symptoms related to anxiety (including muscle tension, hypervigilance, and difficulties related to sleep) to a medical problem than to anxiety (Gurian & Miner, 1991). In turn, many physical conditions such as cardiovascular disease, respiratory disease, hyperthyroidism, and pulmonary difficulties can involve anxiety symptoms, making it difficult to establish the underlying cause of these symptoms (Alwahhabi, 2003; Kogan et al., 2000). Additionally, anxiety symptoms can occur as side effects of medication that is being used to treat a medical condition.

In assessing elderly patients with anxiety complaints or unexplained physical symptoms, clinicians need to obtain enough information to make an appropriate diagnosis or, at the least, to describe possible differential diagnoses. Moreover, current anxiety diagnostic criteria and measures, most of which were developed originally for use with much younger persons, are weighted heavily with somatic items, making it difficult to distinguish between medical and psychological causes of anxiety in this population. Methods for assessing anxiety in older adults can be enhanced through consideration of unique aspects of anxiety in this population, including age-relevant aspects of physical and mental health status.

Anxiety disorders are also frequently comorbid with cognitive decline and dementia among the elderly (Seignourel, Kunik, Snow, Wilson, & Stanley, 2008). This comorbidity is partially due to the fact that there may be a specific relation with anxiety and several types of cognitive impairments (Wolizky-Taylor et al., 2010). Cognitive decline is an important factor to consider when assessing anxiety in this age group because it may affect the presentation of the symptoms as well as the ability to communicate them to a clinician. What seems to be symptoms of anxiety in older adults (e.g., agitation) may in reality be the result of the challenges associated with memory impairment. Finally, older adults or clinicians may misattribute
anxiety symptoms (e.g., fatigue, difficulty concentrating) to normal aging processes. It has also been suggested that the previous cohort of older adults might be less comfortable than current generations in discussing emotions and, therefore, more likely to minimize their symptoms (Pachana, 2008).

In sum, as older adults experience anxious symptoms in a different way than younger adults do, and as there are many factors that can complicate the use of self-report measures with older adults, it is critical that the anxiety measures used with older adults be demonstrated to be scientifically sound. In recent years, researchers have reviewed anxiety measures often used with older adults (Edelstein et al., 2008; Kogan et al., 2000). Although such reviews provide an important first step in evaluating the scientific merits of available instruments, they do have significant limitations. Most notably, neither Edelstein et al. (2008) nor Kogan et al. (2000) indicated the bases for selecting the instruments they chose to review. Furthermore, only limited psychometric information on selected instruments was presented in these reviews, thus making it difficult for readers to determine the scientific adequacy of these instruments for clinical or research purposes.

With this background in mind, we set out to systematically review the research literature in order to identify the anxiety measures most commonly used in the assessment of older adults. Once identified, our intent was to review the literature to determine the extent to which these instruments had age-relevant norms and psychometric data (reliability, convergent validity, discriminant validity, and treatment sensitivity) supporting their use with older adults. By summarizing this research, we provide critical up-to-date information regarding the status of evidence-based instruments for assessing anxiety in older adults. As indicated, this information is crucial in developing and providing evidence-based mental health services to older adults.

### Method

#### Study eligibility and search strategy

The electronic databases PubMed and PsycINFO were searched for published journal articles that included some type of anxiety measure used in a population of older adults. Studies meeting the following criteria were selected: (1) an empirical study, (2) used an anxiety measure with at least one participant, and (3) only included adults aged 65 years and above in the sample.

To ensure the broadest possible search of the databases, two separate searches combining different key words and study criteria were carried out in each database. The first search included the following terms: anxiety, anxiety disorder, generalized anxiety disorder AND assessment, measurement. The second search combined the following terms anxiety, anxiety disorders, generalized anxiety disorder AND geriatric assessment, geriatric patients, geriatric psychiatry, geriatric gerontology. The key words were searched as article keywords, titles, and abstracts. The search was limited to older adults by selecting ‘Age 65 and older’ in the Age Group section of the search. The search was restricted to articles published between January 1960 and August 2009, in English or French (as these were the languages understood by the authors). Articles were obtained through our university library network or, for the articles that were unavailable in print or electronically at our university, through inter-library loans. If an article was unavailable through both these methods, a search was conducted at the National Archives of Canada.

#### Search results

The search yielded a total of 1427 articles (785 in PubMed and 642 in PsycINFO). After reviewing the titles and abstracts, 592 unique articles were retained at this point for further examination to determine if they fully met our inclusion criteria. As part of this review, only studies presenting original data were included in our final set of retained articles. Articles were excluded because they did not have a sample that included only older adults (49.2%), did not use an anxiety measure (8.5%), did not have a sample of older adults and did not have an anxiety measure (7.1%), were not empirical studies (20.0%), did not use original data (6.3%), were not in French or English (5.2%), or were unavailable in print or electronic form within our university library network, through inter-library loans or through the National Archives of Canada (3.6%).

#### Data coding

For all the 592 selected articles, the full article was obtained and reviewed to assess the fit with our inclusion criteria. After this detailed review, a total of 213 articles were retained. Articles were excluded because, after thorough review, they were found to not meet the inclusion criteria which required the articles to be empirical studies that employed an anxiety measure with adults aged 65 and older. The details of the 213 retained articles were coded to summarize key aspects of the study and the anxiety measures used in the study. For each anxiety measure, the rater noted if the authors indicated that the measure was appropriate for older adults, if it was designed for older adults, and if relevant norms are available. Participant variables (age range of the sample, the mean age and standard deviation of the sample, proportion of men and women included in the sample), and sample information (selection of participants, research setting) were also coded.
Results

General characteristics of the studies

Of the retained articles, close to half (45%) were in journals that specialize in research on older adults. The majority of articles meeting the inclusion criteria (83%) were published after 1997. The age of the participants in the retained articles ranged from 65 to 102 years. The samples reported in the majority of studies consisted of older adults recruited from within their communities (56%). With respect to study recruitment procedures, participants were also recruited from medical settings (38%), mental health settings (9%), and residential settings (11%). It is worth noting that the recruitment total does not add to 100% as some studies recruited participants in more than one setting (e.g., use of both a clinical and community sample in a study). In approximately half of the studies, the intent was to conduct research on normal, community-dwelling older adults; samples were also recruited because of having a specific medical disorder (26%), mental disorder (15%), or both (1%).

Most commonly used measures of anxiety

A total of 91 different anxiety measures were used in the 213 included articles. However, the majority of these measures were used in only one or two studies. Most (89%) of the anxiety measures were developed for use with younger adults but used with older aged samples; few measures (16%) were created with older adults’ specific experiences or needs in mind. Of the measures created specifically for older adults, only one (Geriatric Mental State Examination, GMSE; Copeland et al., 1976) was commonly used in the studies we examined (14 times). Study authors rarely mentioned whether measures were appropriate for older adults (24%), and even fewer reported whether specific age-relevant norms were available to aid in the interpretation of the anxiety data (21%).

In our review of the scientific status of anxiety measures used with older adults, we focused on measures that had been used frequently in the literature, thus allowing for independent replication of findings and the availability of data from multiple types of samples. With this in mind, and in consideration of the measure frequency distribution we found in the 213 studies, we decided to concentrate on instruments that had been used in six or more studies. Using this criterion, the most commonly employed measures to evaluate anxiety in older adults were the State Trait Anxiety Inventory (STAI; Spielberger et al., 1983) is a 40-item self-report questionnaire, derived from the Minnesota Multiphasic Personality Inventory, that is designed to measure and differentiate between anxiety as a trait and as a state. The two scales consist of 20 items answered on a four-item scale and provide a score ranging from 20 to 80, with higher scores indicating higher levels of anxiety. A cut-off score of 39/40 for both single scales is normally used to identify clinically significant symptoms of anxiety. However, optimal cut-off scores for older adults have been found to range between 44 and 55 (Himmelfarb & Murrell, 1984;
The scale was developed for young and middle-aged adults, but further research has examined its psychometric properties with older adults. Good internal consistency has been shown for both the trait and state scales in older psychiatric outpatients ($\alpha = 0.92 - 0.94$ for the state version and $\alpha = 0.88 - 0.90$ for the trait version; Kabacoff, Segal, Hersen, & Van Hasselt, 1997; Stanley, Beck, & Zebb, 1996; Stanley, Novy, Bourland, Beck, & Averill, 2001) and in community-dwelling older adults ($\alpha = 0.79 - 0.90$ for the trait version and 0.85 for the state version; Himmelfarb & Murrell, 1984; Stanley et al., 1996). Test–retest reliability has been found to be good for the trait scale ($r = 0.58 - 0.84$) and appropriately lower for the state scale ($r = 0.51 - 0.62$; Stanley et al., 1996, 2001). Unfortunately, there is only limited evidence of concurrent validity, as demonstrated by moderate correlations between the STAI-T and other measures of anxiety ($r = 0.33 - 0.57$; Kabacoff et al., 1997; Stanley et al., 1996, 2001); slightly lower correlations were found between the STAI-S and other measures of anxiety ($r = 0.15 - 0.52$; Kabacoff et al., 1996; Stanley et al., 1996, 2001). Both scales are substantially correlated with measures of depression ($r = 0.41 - 0.70$; Stanley et al., 1996, 2001), indicating only limited discriminant validity. There is also a concern that the STAI is lengthy and easily misinterpreted by older adults (Dennis et al., 2007).

Few studies identified in our review used the STAI to assess the effects of mental health treatments for older adults. In a sample of older adults receiving cognitive-behavior therapy (CBT) for GAD, the STAI was used before and after initiating treatment (5–20 weeks after the pretreatment assessment; Stanley et al., 2001). Post hoc paired comparisons of scores on the STAI revealed a significant reduction in anxiety symptoms from pre- to post-treatment. Another study (Stanley et al., 2003) examined the efficacy of a CBT treatment relative to a minimal contact control in a sample of older adults with GAD. All participants completed measures of anxiety (STAI and HARS) and worry (WS and PSWQ) before and after initiating treatment. An analysis of simple effects demonstrated significant improvements on the STAI as well as on the HARS and PSWQ for the CBT group but not for the control group. Thus, evidence from two studies suggests that the STAI can be sensitive to treatment effects. Although promising, evidence of the instrument’s treatment sensitivity with other disorders and other treatments is needed. Overall, the studies examining the psychometric properties of the STAI have yielded mixed results and, therefore, it should be used with some caution when assessing anxiety in older adults.

**Hospital Anxiety and Depression Scale**

The HADS (Zigmond & Snaith, 1983) is a 14-item self-report questionnaire developed to evaluate the presence and the severity of anxiety and depression in non-psychiatric outpatients. Items referring to symptoms that may have physical causes such as insomnia and dizziness were excluded from the scale during its development. It is, therefore, unbiased by comorbid medical conditions. The measure is divided into a seven-item anxiety subscale (HADS-A) and a seven-item depression subscale (HADS-D). The items are rated on a four-point scale and summed to provide a score ranging from 0 to 21 for anxiety and for depression. There are no fixed cut-off scores for the HADS but, in their original study, Zigmond and Snaith (1983) recommended two cut-off scores: 7/8 for possible anxiety and depression and 10/11 for probable anxiety or depression. A cut-off score of 14/15 was later added for severe anxiety or depression, but no empirical data support this score (Snaith & Zigmond, 1994). The cut-off scores have not been validated with older adults but there is some evidence that use of the 7/8 cut-off score will correctly identify the majority of anxious older adults (Dennis et al., 2007; Haworth, Moniz-Cook, Clark, & Wang, 2007). No norms are available for older adults.

The HADS was originally developed for medical outpatients aged between 16 and 65, but further research has since been conducted to validate its use with older adults. Even though not originally designed for use with psychiatric outpatients, the internal consistency of the HADS is high in samples of such individuals ($\alpha = 0.73 - 0.80$; Flint & Rifat, 2002; Wetherell, Birchler, Ramsdell, & Unutzer, 2007). It has also been found to be high in samples of older medical inpatients ($\alpha = 0.75 - 0.84$; Bryant et al., 2009; Johnston, Pollard, & Hennessey, 2000; Yu, Lee, Woo, & Hui, 2007) and in community samples ($\alpha = 0.84 - 0.85$; Spinhoven et al., 1997). Although we found no study that reported evidence of test–retest reliability, the anxiety scale of the HADS has been found to show no significant changes when administered at a two month interval in a population of inpatients in a geriatric hospital (Bryant et al., 2009). The HADS-A has demonstrated only limited evidence of concurrent validity, with moderate correlations with other psychological distress measures such as the BSI ($r = 0.54$; Wetherell & Aréan, 2007), the HARS ($r = 0.57$; Dennis et al., 2007), and observer ratings of global anxiety ($r = 0.28$; Kenn, Wood, Kucyj, Wattis, & Conane, 1987). It has shown moderate evidence of discriminant validity with depression measures ($r = 0.26 - 0.47$), but moderate correlations between the anxiety and depression subscale ($r = 0.43 - 0.73$) could result in high rates of misclassification (Davies, Burn, Mckenzie, Brotherwell, & Wattis, 1993; Dennis et al., 2007; Flint & Rifat, 2002; Johnson et al., 1995; Spinhoven et al., 1997).

A study conducted by Yu et al. (2007) examined the effects of relaxation and exercise training on...
psychological outcomes in older patients with heart failure. Participants completed the HADS at baseline and at the 12th week of treatment. Analyses revealed that older adults who participated in the relaxation or exercise therapy showed lower levels of anxiety on the HADS compared to the control sample. Although supportive of the instrument’s treatment sensitivity, additional research studies assessing the effects of mental health treatments with the HADS are needed in the geriatric population.

Overall, although the HADS excludes somatic symptoms and shows high internal consistency, the correlations between the anxiety and depression scale suggests that it may be most useful as an overall indicator of distress. These factors, when combined with the lack of evidence for clinical cut scores or norms relevant to older adults, suggest that this frequently used measure is not a good option for assessing anxiety in older adults (Bryant et al., 2009).

**Geriatric Mental State Examination**

The GMSE (Copeland et al., 1976) is a semi-structured clinical interview designed as a mental health assessment for older adults. The original version of the GMSE had 541 items, but it has since been shortened so that it can be administered in 20–50 min. The data can be analyzed by the computerized system AGECAT (Automated Geriatric Examination or Computer Assisted Taxonomy) to obtain a suggested psychiatric diagnosis. Symptoms are grouped into eight syndrome clusters: organicity, schizophrenia and related paranoia, mania, depression, hypochondrias, phobias, and obsessional and anxiety neurosis. A diagnostic confidence level is provided for each syndrome, ranging from 0 (no symptoms) to 5 (very severely affected). A level of three or more in a cluster represents a diagnostic case (Copeland, Dewey, & Griffiths-Jones, 1986).

The GMSE was developed and normed for older adults, and has become one of the most widely used comprehensive structured mental health assessments for older adults (Copeland et al., 2002). High inter-rater reliability for the general scale has been found in samples of community-dwelling older adults, older medical patients, and older psychiatric inpatients due to its well-established procedures and structured approach to the administration of questionnaire (κ = 0.73–0.80; Ames, Flynn, Tuckwell, & Harrigan, 1994; Copeland et al., 1975; Turrina et al., 1991). Test–retest reliability has been somewhat inconsistent, with r values ranging from 0.49 to 0.75 (Copeland et al., 1976; Henderson, Duncan-Jones, & Finlay-Jones, 1983). Evidence of concurrent validity in a variety of cultures has been shown with high correlations between the GMSE and DSM diagnostic criteria (r = 0.76–0.78, Ames et al., 1994; Copeland et al., 1999). We were unable to find any study in which discriminant validity of the anxiety-related clusters was examined. There were no studies identified in our review in which the GMSE was used to assess the effects of mental health treatments for older adults. In sum, the GMSE is a useful tool to assess the mental health of older adults in medical settings as it excludes the effects of physical illness. It has been subjected to many reliability and validation studies and is often used when examining the validity of other instruments used with older adults (Mottram, Wilson, & Copeland, 2000). Although evidence regarding the validity of the anxiety-related clusters and the treatment sensitivity of the impairment is required, data on the correspondence with DSM diagnostic criteria suggest that it is likely to be useful in assessing clinically significant anxiety in older adults.

**Hamilton Anxiety Rating Scale**

The HARS (Hamilton, 1959) is a 14-item clinician-administered rating scale developed to assess the severity of anxiety symptoms in adults. Seven of the items address psychic/cognitive anxiety and the remaining seven items somatic anxiety. The items are rated on a five-point scale and summed to provide a score ranging from 0 to 56. A score of 17 or less represents mild anxiety, a score between 18 to 24 mild to moderate anxiety, and a score of 25 and above moderate to severe anxiety. The cut-off scores have not been validated with older adults and there are no published norms for older adults (Kogan, Edelstein, & McKee, 2000; Scheik, 1991).

Although the HARS was developed for young and middle-aged adults, there is some recent support for its use with older adults. Adequate internal consistency has been shown in samples of older adults diagnosed with generalized anxiety disorder (α = 0.77–0.86; Beck, Stanley, & Zebb, 1999; Diefenbach et al., 2001; Schuurmans et al., 2009). High inter-rater reliability has been shown with community samples and with older adults diagnosed with generalized anxiety disorder (r = 0.81–0.95; Lenze et al., 2009; Stanley et al., 2009; Wetherell, Gatz, & Craske, 2003). The HARS showed limited concurrent validity with the STAI-T (r = 0.23; Diefenbach et al., 2001) and the BAI (r = 0.47; Morin et al., 1999). The scale has been shown to differentiate older adults with generalized anxiety disorder from those with no anxiety disorders (Beck, Stanley, & Zebb, 1996; Edelstein et al., 2008). However, the scale correlates considerably with the Hamilton Depression Rating Scale in samples of older adults, raising concerns about its discriminant validity (r = 0.72–0.92; Beck et al., 1996, 1999; Diefenbach et al., 2001). Importantly, the usefulness of the HARS with older adults has been questioned due to the heavy emphasis placed on somatic symptoms (e.g., tension) that are common experiences in aging individuals (Kogan et al., 2000; Skopp et al., 2006).

Several studies identified in our review used the HARS to assess the effects of psychological and pharmacological treatments for anxiety in older adults. In a randomized controlled trial comparing sertraline and CBT for the treatment of late-life anxiety, the HARS showed moderate to large effect...
sizes for the sertraline group and small to moderate effect sizes for the CBT group at both post-treatment and one year follow-up (Schuurmans et al., 2009). Similar results were found in a recent randomized controlled trial comparing the effect of escitalopram with a placebo (Lenze et al., 2009). However, studies examining the effect of CBT in samples of older adults with GAD have yielded mixed results. Two studies (Stanley et al., 2003; Wetherell et al., 2003) found that CBT participants had improved significantly on the HARS as well as on other anxiety measures, whereas another study (Stanley et al., 2009) found that changes on the HARS were not significantly different between the CBT and control group.

Overall, the HARS has shown high reliability in samples of older adults and some evidence of treatment sensitivity with psychopharmacological interventions. However, more studies are needed to establish its validity and treatment sensitivity across treatments in diverse groups of older adults. Given the lack of age-relevant norms and cut-off scores, and concerns about discriminant validity, this rating scale is not an optimal choice for use with older adults.

**Goldberg Anxiety and Depression Scale**

The GADS (Goldberg et al., 1988) is an 18-item self-report questionnaire that measures symptoms of depression and anxiety experienced in the past month. The items are rated on yes (1) or no (0) answers and are summed to provide a score ranging from 0 to 9 for the depression subscale and for the anxiety subscale. According to Goldberg (1988), patients with anxiety scores of 5 or more or with depression scores of 2 or more have a 50% chance of a clinically important disturbance, and the probability of a significant disturbance increases substantially with higher scores. The anxiety cut-off score has not been validated with older adults and there are no normative data available for older populations.

The GADS was not developed for use with older adults and only a few studies have examined its psychometric properties with this population. The GADS showed good internal consistency in a sample of adults aged 18–79 where no substantial differences were noted when the alpha was calculated separately for different age groups (α = 0.82; Christensen et al., 1999). Another study, conducted with older medical inpatients, reported good internal consistency (α = 0.82; Huber, Mulligan, Mackinnon, Nebuloni-French, & Michel, 1999). Although this shows initial evidence of reliability, no other published studies have examined reliability indices specifically in samples of older adults. Furthermore, we were unable to find any reports of test–retest reliability among samples of older adults. In the general population, the GADS correctly identified over 80% of adult patients with anxiety disorders (Goldberg et al., 1988). However, when it comes to older adults, poor agreement has been found between the anxiety subscale of the GADS and other

anxiety measures, with κ values of −0.13–0.28 indicating poor concurrent validity (Koloski, Smith, Pachana, & Dobson, 2008). Moderate correlations have been found between the GADS anxiety subscale and the GAI in a sample of older adults (Pachana et al., 2007). Finally, there is also evidence that the anxiety and depression subscales of the measure are highly correlated in both the general population and in samples of older adults, which suggests poor discriminant validity (Christensen et al., 1999; Huber et al., 1999; Koloski et al., 2008). There were no studies identified in our review in which the GADS was used to assess the effects of mental health treatments for older adults. As a result, there is no evidence with respect to the treatment sensitivity of the GADS. The inclusion of somatic symptoms (e.g., waking early and headaches) can overestimate the prevalence of anxiety and depression and result in classification errors. Overall, the psychometric evidence available thus far on the GADS is limited and provides little support for its use with older adults.

**Beck Anxiety Inventory**

The BAI (Beck et al., 1988) is a 21-item self-report questionnaire designed to measure the severity of anxiety and to distinguish anxiety from depression. The items are rated on a four-point scale and summed to provide a score ranging from 0 to 63, with higher scores representing higher levels of anxiety. According to the manual (Beck & Steer, 1990), the score can be interpreted as follows: 0–9 (normal anxiety), 10–18 (mild to moderate anxiety), 19–29 (moderate to severe anxiety), and 30–63 (severe anxiety). Based on the information reported in the manual, it is unclear how those cut-offs were derived and there is no mention of whether different cut scores should be used with older populations. In subsequent research conducted with older adults, no single BAI cutoff proved to be optimal due to the tradeoffs between sensitivity and specificity (Kabacoff et al., 1997).

The BAI was developed and normed with samples of psychiatric adult outpatients. However, since its development, there have been several studies that evaluated its use with older populations. The internal consistency of data collected with the BAI is high in samples of older adult medical outpatients (α = 0.91–0.92; Diefenbach, Tolin, Meunier, & Gilliam, 2009; Wetherell & Aréan, 1997), older adult psychiatric outpatients (α = 0.81–0.93; Kabacoff et al., 1997; Wetherell & Gatz, 2005), and in community samples (α = 0.87–0.89; Morin et al., 1999; Wetherell & Gatz, 2005). The BAI showed adequate test–retest reliability (r = 0.64–0.75) in samples of older adults (Beck et al., 1988; Diefenbach et al., 2009). Moderate correlations between the BAI and other anxiety measures show evidence of concurrent validity, with correlations ranging from 0.29 to 0.63 (Dennis et al., 2007; Diefenbach et al., 2009; Kabacoff et al., 1997; Wetherell & Gatz, 2005). However, despite efforts in
the development of the BAI to disentangle symptoms of anxiety and depression, relatively high correlations ($r = 0.56–0.65$) between the BAI and depression measures show only limited evidence of discriminant validity (Wetherell & Aréan, 1997).

In a study of older adults with GAD, participants were randomly assigned to a CBT group, a discussion group, or a waiting list, and were assessed before and after treatment (Wetherell et al., 2003). At post-treatment, mean effects for time among GAD participants were significant for other psychological distress measures, but not for the BAI and the HDRS, suggesting that the BAI may not be sensitive to treatment change. Overall, the initial evidence of psychometric properties as well as the simplicity of the BAI make it a useful tool to detect the presence of anxiety in older adults. However, because of (1) potential confounds with depressive symptoms and (2) high somatic item content (13 of the 21 items are related to somatic symptoms), the BAI should be used with caution, especially with samples recruited from medical settings. Additionally, more evidence is required before it is deemed acceptable for use in evaluating treatment effects.

**General Health Questionnaire**

The GHQ (Goldberg, 1978) is a self-report questionnaire designed to evaluate the presence of minor, non-psychotic psychiatric disorders in community setting. The original questionnaire consists of 60 items, but shorter versions of 30, 28, 20, and 12 items have also been developed. The 28-item version is most commonly used in the general population but the shorter 12-item version has been found to be more appropriate with older populations (Clarke & Clarkson, 2009). The GHQ incorporates four scales: somatic symptoms, anxiety and insomnia, social dysfunction, and severe depression. The items are rated on a four-point scale (‘not at all,’ ‘no more than usual,’ ‘rather more than usual,’ and ‘much more than usual’) and are summed to provide a score ranging from 0 to 84, with higher scores representing higher levels of distress. When using the full 60-item scale, a cutoff of 23/24 for the total scale is suggested. However, Goldberg and Hillier (1979) suggest using an alternative binary scoring method in which the two least symptomatic answers (‘not at all’ and ‘no more than usual’) are given a score of 0 and the two most symptomatic answers (‘rather more than usual’ and ‘much more than usual’) a score of 1. A total score of 4 or more on any subscale suggests caseness. There is no specific cutoff suggested, nor are norms available, for older adults. However, some evidence has shown that when using the binary scoring method with a population of older adults, the best cut-off score for each subscale is 3/4 (Pappasotirooulos, Heun, & Maier, 1997).

The GHQ was developed to be used with adults and adolescents. Although it is one of the most commonly used measures to detect the presence of mental disorders, very little is known about its use with older adults. The internal consistency of the measure is high in samples of older community-dwelling adults when using either the 12- or 28-item version ($\alpha = 0.75–0.90$; Boey & Chiu, 1998; Cheung, 2002; Clarke & Clarkson, 2009; Costa et al., 2006; Malakouti, Fatollahi, Mirabzadeh, & Zandi, 2007), in cognitively impaired older adults when using the 12-item version ($\alpha = 0.81$, Costa et al., 2006) and in older medical outpatients when using either the 12- or 30-item scale ($\alpha = 0.82–0.92$; Dale, Saevareid, & Soderhann, 2009; Thygesen, Saevareid, Lindstrom, Nygaard, & Engedal, 2008). We were unable to locate any study reporting evidence of test–retest reliability for older adults. Few studies have examined the validity of the GHQ with older adults. There were no studies identified in our review in which the GHQ was used to assess the effects of mental health treatments for older adults. As a result, there is no evidence with respect to the treatment sensitivity of the GHQ. Although most studies found that the GHQ could differentiate between older adults with and without mental disorder (Costa et al., 2006; Malakouti et al., 2007; Mowry & Burvill, 1990; Seva, Sarasola, Merino, & Magallon, 1991), one found that it does not differentiate mentally ill patients from those with somatic illness (Malakouti et al., 2007). In light of the limited psychometric information available on the GHQ when used with older adults, the GHQ should be used with considerable caution when assessing anxiety symptoms in older adults.

**Brief Symptom Inventory**

The BSI (Derogatis & Spencer, 1982) is a 53-item self-report questionnaire designed to assess the psychological distress of medical and psychiatric patients. It is a brief form of the SCL-90-R and covers nine symptom dimensions (somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychotism) and three global indices of distress (Global Severity Index (GSI), Positive Symptom Distress Index, and Positive Symptom Total). The global indices measure the level of symptomatology, intensity of symptoms, and number of reported symptoms. The items are rated on a five-point scale ranging from 0 (not at all) to 4 (extremely). The overall score of the BSI is referred to as the GSI. Scores on the GSI ranges from 0 to 72 and each subscale has a score ranging between 0 and 24. The scores are interpreted by comparing them to age appropriate norms. The manual provides norms for adult non-patients, psychiatric outpatients, inpatients, and adolescent non-patients, but not for older adults. However, Hale, Cochran, and Hedgepheth (1984) have provided age-relevant norms for community-dwelling older adults.

Although the BSI was developed for adolescents and adults, some research has been done on its use with older adults. High internal consistency has been found
in samples of cognitively impaired older adults ($\alpha = 0.82$; Fisher, Segal, & Coolidge, 2003), older medical outpatients ($\alpha = 0.89$; Petkus, Gum, King-Kallimanis, & Wetherell, 2009; Petkus et al., 2010), community-dwelling older adults ($\alpha = 0.72–0.79$; Fisher et al., 2003; Petkus et al., 2009), and older adult psychiatric outpatients ($\alpha = 0.72–0.90$; Wetherell et al., 2007, 2010). We found no study reporting test–retest reliability data for the BSI in samples of older adults. There has been little effort to validate the BSI with older adults. One study found that the BSI did not distinguish between medically ill older patients with and without an anxiety disorder (Wetherell et al., 2007), whereas another found that it could discriminate between homebound older adults with and without anxiety disorders (Petkus et al., 2010). There were no studies identified in our review in which the BSI was used to assess the effects of mental health treatments for older adults. Somatic symptoms of the BSI are grouped in a somatization subscale and therefore provide some assurance scores on anxiety-related subscales; they are not inflated by the presence of symptoms better explained by a medical condition. Although the BSI is simple and covers a wide range of symptoms, the lack of supporting psychometric evidence for older adults, especially validity evidence, severely limits its usefulness with older adults.

**Penn State Worry Questionnaire**

The PSWQ (Meyer et al., 1990) is a 16-item self-report questionnaire designed to evaluate pathological worry. The items are rated on a five-point scale and summed to provide a score ranging from 16 to 80, with higher scores reflecting higher levels of worry. Eleven items are worded in the direction of pathological worry, whereas the remaining items are worded to indicate the absence of worry. Although there are no specific cut-off scores, the mean score for individuals with generalized anxiety disorder is between 60 and 68. There are no norms for older adults, but a cut-off score of 50 has been suggested for use with samples of older medical patients (Stanley et al., 2003).

The PSWQ was developed and normed for younger adults, but several studies have examined its psychometric properties in samples of older adults. High internal consistency has been found in samples of older adults diagnosed with generalized anxiety disorder ($\alpha = 0.81–0.89$; Beck, Stanley, & Zebb, 1995; Stanley et al., 2001; Wetherell et al., 2003), in home care residents ($\alpha = 0.79$; Diefenbach et al., 2009), and in community-dwelling older adults ($\alpha = 0.80–0.91$; Beck et al., 1995; Hunt, Wisocki, & Yanko, 2003; Senior et al., 2007). Moderate to high test–retest reliability has been found in samples of older adults ($r = 0.54–0.78$; Hopko et al., 2003; Stanley et al., 2001). The PSWQ showed adequate concurrent validity by virtue of significant correlations with other self-report measures of anxiety ($r = 0.29–0.79$; Andreescu et al., 2008; Diefenbach et al., 2009; Hopko et al., 2003; Kogan et al., 2000; Stanley et al., 2001; Wetherell et al., 2003). Correlations with self-report measures of depression were lower and showed some evidence of discriminant validity ($r = 0.12–0.51$; Diefenbach et al., 2009; Hopko et al., 2003; Kogan et al., 2000; Senior et al., 2007; Stanley et al., 2001; Wetherell et al., 2003). There is concern that some older adults have difficulty completing and interpreting the content of the reversed items of the PSWQ (Stanley et al., 2003; Wetherell et al., 2003). In order to respond to this problem, Hopko et al. (2003) eliminated eight of the original items to create an abbreviated scale (PSWQ-A). This scale has shown good psychometric properties in samples of older adults (Critendon & Hopko, 2006; Hopko et al., 2003; Nuevo, Mackintosh, Gatz, Montorie, & Wetherell, 2007).

Two studies identified in our review used the PSWQ to evaluate the effects of CBT for treating late-life anxiety. In a study examining the efficacy of a CBT treatment relative to a minimal contact control group in older adults with GAD, an analysis of simple effects demonstrated significant improvements on the PSWQ as well as in other anxiety measures for the CBT group (Stanley et al., 2003). In the aforementioned study (Wetherell et al., 2003), the main effects for time were significant for the PSWQ as well as most of the other anxiety and depression measures. These two studies are supportive of the treatment sensitivity of the PSWQ, but additional studies assessing the effects of mental health treatments for late-life anxiety with the PSWQ are needed. Overall then, although the lack of norms for older adults is problematic, initial evidence suggests that the PSWQ and its abbreviated form may be useful in assessing worry in older adults.

**Symptom Checklist 90-R**

The SCL-90-R (Derogatis, 1994) is a 90-item self-report questionnaire designed to evaluate a wide range of psychological problems and symptoms of psychopathology. It covers nine primary symptom dimensions (somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychotism) and has three global indices of distress which give an overall sense of psychological distress. The items are rated on a five-point scale. A score of 50 or lower on any scale is considered to be in the normal range and a score of 65 or above is considered to be a clinical case. The cut-off scores have not been validated for use with older adults. The manual provides norms for adult and adolescents non-patients, adult psychiatric outpatients, and adult psychiatric inpatients; no norms are provided for older adults. Furthermore, we were unable to find any studies examining its psychometric properties with older adults and the reviewed articles that used the SCL-90-R did not report any psychometric information. There were no studies identified in our review in which the SCL-90-R was used to assess the effects of mental health treatments for older adults. Overall, the
lack of evidence for the psychometric soundness of the SCL-90-R with older adults means that it is currently not an appropriate measure for use with this population.

**Worry Scale**

The WS (Wisocki et al., 1986) is a 35-item self-report questionnaire developed to measure worries in the areas of health (17 items), finances (five items), and social conditions (13 items). The items are rated on a five-point scale ranging from 0 (never) to 4 (much of the time) and summed to provide a score ranging from 0 to 140, with higher scores reflecting higher levels of worry. No norms or cut-off scores are provided in the original study. However, a mean total score of 35.4 for individual with GAD and of 11 for non-anxious control have been reported in the literature (Stanley, et al., 1996). Community samples of active older adults have yielded total mean scores ranging from 10.4 to 17.4 and samples of homebound older adults have found mean scores ranging from 17.3 to 23.7 (Wisocki, 1994).

The WS has been created specifically for use with older adults. The WS total score is associated with excellent overall internal consistency in both GAD ($\alpha=0.81-0.93$, Stanley et al., 1996, 2001) and non-anxious samples ($\alpha=0.93-0.94$; Hunt et al., 2003; Stanley et al., 1996) and good internal consistency for its subscales ($\alpha=0.76-0.95$; Hunt et al., 2003; Stanley et al., 1996). Adequate test–retest reliabilities have been found in samples of older adults for the total scale ($r=0.69-0.70$, Stanley et al., 1996, 2001) and the subscales ($r=0.58-0.80$; Stanley et al., 1996, 2001). Concurrent validity for the WS has been shown by significant correlations between the scale and measures of anxiety ($r=0.54-0.63$, Stanley et al., 2001; Wisocki, 1988; Wisocki et al., 1986). However, the scale also showed high correlations with measures of depression ($r=0.500.78$ Wisocki, 1988; Wisocki et al., 1986). As noted previously, examining a CBT intervention for GAD, Stanley et al. (2003) found that improvements were noted for the CBT group on the WS and other anxiety measures.

The original scale has been revised and expanded to include 88 items in six dimensions (finances, health, social conditions, personal concerns, family concerns, and world issues). Its psychometric properties are currently under investigation, with initial evidence indicating that they are similar to the original scale (Hunt et al., 2003; Watari & Brodbeck, 2000). Overall then, although more research is needed on discriminant validity and treatment sensitivity, initial evidence suggests that the WS may be useful in assessing worry in older adults.

**Geriatric Anxiety Inventory**

The GAI (Pachana et al., 2007) is a 20-item self-report questionnaire designed to measure anxiety symptoms in older adults. The questionnaire uses an agree/disagree response choice format, with the number of ‘agree’ responses added for the total score. The maximum score is 20, with higher scores indicating higher anxiety. According to the original study, the optimal cut-off score to identify GAD in older adults is 10/11, and 8/9 to identify other anxiety disorders (Pachana et al., 2007). Similar results have been found in other studies of psychogeriatric patients (Byrne et al., 2010; Pachana et al., 2007).

The GAI was developed and normed with samples of community-dwelling older adults and older adults receiving psychiatric services. Excellent internal consistency has been shown in the original study of community-dwelling older adults ($\alpha=0.92$; Pachana et al., 2007) and older adults receiving psychiatric services ($\alpha=0.93$; Pachana et al., 2007). High internal consistency has also been found in other samples of community-dwelling older adults ($\alpha=0.90-0.92$; Andrew & Dulin, 2007; Byrne et al., 2010; Pachana et al., 2007), in psychogeriatric samples ($\alpha=0.93$; Pachana et al., 2007), and in an older adult sample receiving home care ($\alpha=0.93$; Diefenbach et al., 2009). Lower, but still good, internal consistency was found in older adults with mild cognitive impairment ($\alpha=0.76$; Rozzini et al., 2009). Researchers have described the data obtained from the instrument as having sound test–retest reliability (Pachana et al., 2007), but not report precise reliability information. One study conducted with older adults with mild cognitive impairment showed good test–retest reliability ($r=0.86$; Rozzini et al., 2009). Moderate to strong correlations between the GAI and other anxiety measures show evidence of concurrent validity, with correlations ranging from 0.58 to 0.86 (Byrne et al., 2010; Diefenbach et al., 2009; Pachana et al., 2007). However, relatively high correlations ($r=0.65-0.79$) between the GAI and depression measures provide limited evidence of discriminant validity (Andrew et al., 2006; Diefenbach et al., 2009; Paukert, 2009). There were no studies identified in our review in which the GAI was used to assess the effects of mental health treatments for older adults. Taken together, although much more research evidence is needed, initial evidence suggests that the GAI is likely to be a useful tool for detecting anxiety in older adults.

**Discussion**

The aims of this review were to evaluate the mental health research literature on older adults in order to (1) identify the most commonly used anxiety measures and (2) determine how appropriate the measures are for clinical or research use with older adults by examining their psychometric properties. Results of our literature search indicate that, with more than 90 measures used to evaluate anxiety in older adults, no clear consensus exists among researchers on which anxiety measure is more appropriate to use when evaluating a
geriatric population. Out of the most commonly used measures, only one (GMSE) was developed specifically for older adults. Therefore, it seems that, with the lack of anxiety measures created and validated for older adults, researchers and clinicians often need to rely on anxiety measures created for younger adults when assessing older populations. However, because many differences exist between younger and older adults, it is unlikely that a single measure can adequately assess anxiety across the entire adult life span. For example, compared to younger adults, older adults report more somatic symptoms, which means that anxiety measures that were not specifically developed for older adults may not provide an accurate assessment of anxiety (Fuentes & Cox, 1997; Wetherell & Gatz, 2005). As a result, it is extremely important for clinicians and researchers to have access to valid and reliable measures with established age-appropriate norms.

Most of the assessment measures we reviewed lack sufficient evidence for their psychometric soundness when used with older adults. Several critical considerations limit their use with geriatric populations. First, although each reviewed measure showed adequate internal consistency ($\alpha \geq 0.70$), the existing data often come from a single published study and mostly from samples of older psychiatric outpatients. Both replication and extension of previous findings are necessary in order to determine if the measures are good general screening measures for anxiety in older adults or whether they may have relevance only for a specific subgroup of older adults. Second, only four measures (BAI, GMSE, PSWQ, and STAI) showed evidence of adequate test–retest reliability in older adult samples. The lack of test–retest reliability needs to be taken into consideration if the measure is to be used numerous times, such as when evaluating treatment effects. The third limitation resides in the lack of evidence for both discriminant and concurrent validities in most measures. Three measures showed adequate concurrent validity with older adults (GMSE, BAI, and PSWQ), four showed limited concurrent validity (GADS, HARS, HADS, and STAI), whereas no information was found for the other three measures. Every measure, except the PSWQ, either showed limited or low discriminant validity. Also, very few instruments have adequate normative data for older adults, which severely limit their clinical value. Only the GMSE has been normed for older adults and provides cut-off scores validated for this population. Finally, few measures had evidence of treatment sensitivity, suggesting the urgent need for studies assessing the sensitivity of commonly used anxiety instruments to the effects of mental health treatments for older adults.

Taken together, three of the most commonly used measures showed sufficient psychometric evidence to warrant their use in assessing anxiety in older adults. The BAI, a measure of general anxiety, and the PSWQ, a measure of worry, have both demonstrated high internal consistency in older psychiatric outpatients as well as in community-dwelling older adults. Test–retest reliability of the two measures has been shown to be adequate. Convergent validity for these measures has been suggested by moderate correlations with other self-report measures of anxiety. The PSWQ showed good discriminant validity with low correlations with self-report measures of depression but the BAI show only limited discriminant validity. Although both measures were initially developed for younger populations, subsequent evidence suggests that they might be good choices when selecting an anxiety measure for a geriatric population. The GMSE, a semi-structured interview created for older adults that includes anxiety clusters, is the only measure that provides norms and cut-off scores validated for older adults. That, combined with its appropriate psychometric properties, suggests that it is a good option for assessing anxiety in older adults. However, the GMSE has not been used in research published in the past decade, which may suggest that more recent measures may be more appropriate to measure anxiety in older adults.

Indeed there is a growing number of anxiety instruments designed for use with older adults. Both the WS and the GAI were specifically developed for older adults and showed sufficient psychometric evidence to warrant their use in assessing anxiety in older adults. In many instances, these instruments might be preferred over the more commonly used measures that were not developed for the geriatric population. Both measures demonstrated high internal consistency in older psychiatric outpatients as well as in community-dwelling older adults. Test–retest reliability of the two measures has been shown to be good. Convergent validity for these measures has been provided by significant correlations with other self-report measures of anxiety. But, because of the extent to which these measures have been found to correlate with measures of depression, further research focusing on discriminant validity is certainly required.

**Coexistence of somatic symptoms**

In this review, six of the most commonly used measures (BAI, BSI, GHQ, GADS, HARS and SCL-90-R) were weighted heavily with somatic symptoms, which makes it difficult to distinguish between anxiety symptoms and symptoms of other health problems (or even normal aging) among the geriatric population. This can be problematic as the experience of anxiety varies greatly in younger and older adults. Not only do older adults experience more somatic symptoms when anxious but they are also likely to have coexisting physical conditions that may produce anxiety-like symptoms. It is therefore critical for the measures to be able to distinguish between anxiety-like symptoms caused by a medical condition from symptoms caused by an anxiety disorder. By using a measure that includes many somatic symptoms, a high proportion of non-anxious older adults experiencing symptoms of a
medical condition may fall within the range used to identify clinical anxiety in a younger population. This must be taken into consideration when using these measures with older adults, particularly with older adults in medical settings or with a medical condition.

Coexistence of depression

Another issue that must be taken into consideration when evaluating the presence of anxiety in older adults is the frequent coexistence of anxiety and depression in later life. Studies suggest that as much as 38–46% of older adults meeting criteria for a mood disorder also meet criteria for an anxiety disorder (Beekman et al., 1998; Flint, 1994; Lenze et al., 2000). As these comorbid conditions can increase the complexity of anxiety assessment and diagnosis, it is important that the measures differentiate anxious and depressive symptoms in older adults. In this review, none of the most commonly used measures showed adequate evidence of discriminant validity with respect to mood disorders, which is likely to lead to a high misclassification rate among older adults. Two measures (GADS and HADS) include both a depression and an anxiety subscale but research indicates that, for both measures, the anxiety and depression subscales are highly correlated and therefore might not distinguish between the disorders (Davies et al., 1993; Flint & Rifat, 2002; Koloski et al., 2008). Consequently, researchers and clinicians should be careful when using these measures until more solid evidence of discriminant validity is obtained. Obviously, information beyond what is available from these instruments must be considered when making any diagnostic formulations.

Limitations

Findings from this review must be interpreted within the limitations of systematic reviews in general. A key issue in selecting articles for review was deciding which studies to include and which to exclude. In particular, we required that studies have samples in which all participants were at least 65 years of age. We excluded from consideration, therefore, a number of articles in which some research sample included participants in their late 50s or early 60s. Accordingly, this limited the number of studies we examined in detail. Setting the inclusion criteria to allow samples in which all participants were at least 55 years of age would certainly have increased the number of studies available for our review, but at the cost of including too broad a range of ages for our intended focus on older adults. The search was also limited to articles published in either French or English as those were the languages understood by the authors. Furthermore, we searched for studies in the most important databases for psychology (PsycINFO) and medicine (PubMed), but other databases were not considered. It is conceivable that this might have resulted in an overly narrow review of the published literature on anxiety measures used with older adults. However, our search strategy did ensure that we considered the vast majority of journals that publish research on the health and/or mental health of older adults.

Conclusions

The present systematic review shows that the anxiety measures most commonly used with older adults are mostly measures developed for a younger population. Although there is empirical support for the use of some of these measures, the majority of measures lack sufficient evidence of their psychometric soundness when used with older adults. The STAI was found to be the most commonly used measure in the reviewed articles. However, an examination of its psychometric properties yielded mixed results, suggesting that it does not yet show sufficient supporting psychometric evidence and should, therefore, be used with caution when assessing older adults. The HADS was also frequently used, but the lack of psychometric evidence for this instrument suggests that it is not a good option when evaluating older adults until more research examines its validity and reliability. The GMSE was created for older adults and shows preliminary evidence of psychometric soundness. However, this measure has not been used in recent studies and, therefore, might not be the most appropriate choice when assessing anxiety in older adults. Both the PSWQ and the BAI have shown good psychometric properties, suggesting that they may be useful tools to measure the presence of anxiety in older adults. However, considering the limited research on the psychometric properties of the measures, clinicians and researchers must be cautious and carefully consider the strengths and weaknesses of each measure before deciding which one to use for a specific purpose. Although the evidence to date is somewhat limited, the measures specifically developed for older adults, such as the WS and the GAI, should be seriously considered by clinicians and researchers when assessing anxiety in a geriatric population.

A major shortcoming evident in the reviewed measures is the inclusion of somatic symptoms of anxiety that often overlap with the symptoms of normal aging, comorbid conditions, and medication side effects. The use of measures heavily weighted for somatic symptoms should be avoided in medical settings or in samples presenting with a medical condition. Caution should also be used when assessing older adults who present possible depressive symptoms, as many of the anxiety measures are highly correlated with measures of depression. Additionally, most measures do not present age appropriate norms or clinically relevant cut-off scores, which greatly limit their use with older adults. For the evaluation of anxiety in older adults to be more evidence-based,
there is a pressing need for the validation of measures that were created for young adults and that are used with older adults. Research is needed on the development and validation of anxiety measures created specifically for older adults that evaluate a wide range of anxiety symptoms and disorders.

References


