Mokken scaling analysis of the Hospital Anxiety and Depression Scale in individuals with cardiovascular disease.

Theodore D. Cosco  
Royal College of Surgeons in Ireland

Frank Doyle  
Royal College of Surgeons in Ireland, fdoyle4@rcsi.ie

Roger Watson  
University of Sheffield

Mark Ward  
Trinity College Dublin

Hannah McGee  
Royal College of Surgeons in Ireland

Citation
Manuscript Number: GHP-11-263R1

Title: Mokken scaling analysis of the Hospital Anxiety and Depression Scale in individuals with cardiovascular disease

Article Type: Regular Article

Corresponding Author: Mr. Theodore David Cosco, M.Sc.

Corresponding Author's Institution: Royal College of Surgeons in Ireland

First Author: Theodore David Cosco, M.Sc.

Order of Authors: Theodore David Cosco, M.Sc.; Frank Doyle, PhD; Roger Watson, PhD; Mark Ward, MSc; Hannah McGee, PhD
Mokken scaling analysis of the Hospital Anxiety and Depression Scale in individuals with cardiovascular disease

Theodore D. Cosco¹*, Frank Doyle¹, Roger Watson², Mark Ward³, and Hannah McGee¹

¹ Division of Population Health Sciences (Psychology), Royal College of Surgeons in Ireland
² School of Nursing and Midwifery, University of Sheffield
³ School of Social Work and Social Policy, Trinity College Dublin

*Corresponding Author: Theodore D Cosco, 123 St. Stephens Green, Dublin 2, Ireland, e-mail: coscot@tcd.ie; phone: +353 01 402 2177.
1. Introduction

Anxiety and depression are predisposing psychosocial risk factors for cardiovascular disease (CVD) (1). The relationship between these risk factors and CVD is not unilateral, but rather complex and reciprocal in nature. It is a peculiar phenomenon that the negative effects of this comorbidity expand to both psychological and physical dimensions, reciprocally. According to several reviews (2, 3) and a recent meta-analysis (4), the presence of depression and/or anxiety disorders is linked with increased incidence of recurrent cardiac events, sudden death and poorer prognoses in individuals with CVD. Within the context of CVD, physical ailments inhibit functioning which can lead to anxiety and depression, which can in turn lead to the worsening of CVD (2). Furthermore, such symptoms are generally uncorrelated with disease indices, highlighting their importance as an issue to be considered in addition to the physical ailment (5). Anxiety and depression are more predictive of functional impairment in individuals with medical conditions, such as CVD, than the severity of the physical ailment itself: the severity of the anxiety/depression is indicative of the severity of the risk to the individual (2).

Based on the assertion that anxiety and depression are predisposing risk factors for CVD and that inextricably links between these disorders exist, the ability of the Hospital Anxiety and Depression Scale (HADS) (6) to capture these phenomena should be subject to scrutiny. While a phenomenological assessment of anxiety and depression provides distinct disparities, an empirical divide between these closely related disorders has been elusive (7). The degree
of overlap between symptoms of anxiety and depression across HADS items is marked, calling into question the ability of the scale to differentiate between these disorders. While the disparate subscales of the HADS are intended to measure mutually exclusive levels of anxiety (HADS-A) and depression (HADS-D) a recent systematic review of the latent structure of the HADS does not support the traditional anxiety-depression bidimensional structure (8). Furthermore, a systematic review of the case-finding ability of the HADS indicates that the HADS is a useful tool in the identification of non-psychiatric patients with emotional distress; however, concern for the ability to distinguish between emotional disorders is also voiced (9). For clinicians, cognizance of the structure of the HADS can enhance the ability to swiftly assess symptom severity, given the inherent hierarchical structure of psychiatric morbidity (10). The marked overlap between symptoms of anxiety and depression in the context of the HADS items highlights the need for further investigation into the ability of the HADS to accurately differentiate between anxiety and depression.

Previous studies with individuals with CVD have provided mixed evidence as to the latent structure of the HADS, revealing one- (11), two- (12, 13), and three-factor models (14-21). If a two-factor latent variable structure of the HADS cannot be confirmed, it cannot be conclusively deduced that the HADS is accurately measuring, and differentiating between, anxiety and depression (22). While the majority of studies have supported a three-factor structure of the HADS, the single-scale model of the HADS has received nearly unanimous support amongst extant studies using item response theory (IRT) methods (23-
Despite the popularity of CTT methods, such as factor analysis, IRT has significant advantages over CTT, potentially providing much more compelling evidence of the latent structure of the HADS (26). Results derived from CTT methods are confined to the population in which the original study was conducted; however, through the incorporation of item information into the ability-estimation process and conversely the incorporation of examinee ability into item-parameter estimation, IRT results can be applied beyond the scope of the original study (27). Given the limitations of CTT methods and the robust evidence provided by IRT, further investigation using IRT methods is warranted. A potential drawback of many IRT methods are difficulties in fitting an appropriate model to the data, these shortcomings are most pronounced in parametric IRT methods, e.g. Rasch analysis. Therefore, a non-parametric IRT method, Mokken scaling, is used in the current study.

2. Methods

A secondary analysis of cross-sectional HADS data collected from four independent studies of CVD in Ireland was conducted. These HADS data were aggregated into a single dataset and analysis of the latent variable structure of the HADS was conducted via Mokken scaling.
2.1. Participants

Four studies including individuals with heart failure, acute coronary syndrome, and post-coronary artery bypass graft surgery were included in the data conglomerate: the Coronary Care Unit (CCU) study, the Heart Failure (HF) study, the Psychological Well-Being (PWB) study, and the Coronary Artery Bypass Graft (CABG) study. HADS data from 893 participants was collected from the contributing studies (CCU, n=444; HF, n=90; PWB, n=85; CABG, n=274); these data were collected via different methods, within different diagnostic populations and at different times (further details of the study methods are available from the corresponding author).

2.2. Materials

Zigmond & Snaith’s (1983) HADS is a 14-item self-assessment scale used to detect possible (>8) and probable (>11) depression and anxiety disorder cases in non-psychiatric hospital outpatients. The HADS is composed of two seven-item subscales, the HADS-A and HADS-D measuring anxiety and depression, respectively.

2.3. Data Analysis

Missing data (comprising less than 1% of the total data points) were imputed via a regression based impute function in STATA, calculating missing items based on completed scale items. After imputation the data were rounded to the nearest number, as the HADS deals exclusively with whole numbers.
Data were analyzed using PASW 18.0 (30), with Mokken analysis conducted in STATA 9.2 (31) and R 2.11.1 (32). MSP was conducted with the “msp” function in STATA, written by Jean-Benoit Hardouin (33). The scale extracted from MSP was examined for monotone homogeneity and invariant item ordering using R.

Mokken Scaling Procedure (MSP), a nonparametric IRT model, is a probabilistic version of Guttman scale analysis used to assess the dimensionality and scalability of psychometric measures (34, 35). Guttman scaling involves a set of increasingly “extreme” binary items to which an examinee responds, under the assumption that an examinee responding positively to a more extreme item will respond positively to a less extreme item. MSP posits a stochastic relationship between the item and underlying variables, improving upon the deterministic nature, and the uninformative step shaped item characteristic curve, associated with Guttman scaling (36).

Unidimensional scales are formed in MSP using a “bottom up” hierarchical clustering procedure (34). Using the homogeneity coefficient (Hij) as the clustering criterion, the pair of items with the highest Hij is selected (34). After the first pair of items has been extracted, the item with the next best fit in the scale is selected (34). Inclusion of items onto a respective scale proceeds based on the fulfillment of the criterion of positive covariation with the previous item, a Loevinger $H$-value ($H$) above the coefficient value ($c$) and of the item increasing the overall $H$ value of the previously selected items. This procedure is repeated until no remaining items meet inclusion criteria for that scale (37). Once a scale
HADS Mokken Scaling

has been saturated with items meeting its inclusion criteria, a new scale is formed and the process continues until all of the items have been used or specifically excluded due to failure to fulfill inclusion criteria. For each user-defined coefficient level a potentially unique set of scales can be created; the most interpretable scales are then extracted from a given c level (36, 38). Due to minimal model assumptions in relation to parametric IRT models, a larger number of items can be included in the creation of a scale, which has lead some researchers to describe MSP as “exemplifying the purest form of unidimensionality” (39). According to Meijer & Baneke (2004) the minimum sample size for MSP is 400.

The ability of a scale to discriminate via its component items and subsequently differentiate and order examinees is determined by the scale’s $H$ value (36). The $H$ value for a scale is calculated using the formula: $H = 1 - \left( \frac{\text{Observed Guttman Error}}{\text{Predicted Guttman Errors}} \right)$. Predicted Guttman errors are based on the probability that an item will be chosen by chance, given the item difficulty, whereas observed Guttman errors are responses that are “out of order” in the Guttman sequence (36). Higher $H$ values are obtained when fewer observed Guttman errors occur, in other words, when a greater number of items are responded to in a manner that corresponds with the proper sequence. For the purposes of interpretation $0.3 \leq H \leq 0.4$ indicates a weak scale, $0.4 \leq H \leq 0.5$ indicates a moderate scale, and $H \geq 0.5$ indicates a strong scale (36).

Increasing user defined coefficient values (c) were entered, according to Meijer & Baneke’s (36) MSP procedure. After an initial screen at .01, c values
were incrementally increased by .05 from .05 to .6, inclusive, using c as the lower-bound cutoff score for scale inclusion. Output from the various c increments were evaluated in terms of interpretability, extracting the most interpretable solution (36). Scales extracted at higher c values, with correspondingly high $H$-values, and those including a greater number of items were deemed more interpretable. Of these potential scales, the scales that included the most items that could be intuitively grouped, i.e. according to similar symptoms or disorders, were extracted. These extracted scales were then examined for violations of the assumption of monotone homogeneity as well as for invariant item ordering. Invariant item ordering, i.e. the same ordering of item difficulty across varying levels of the latent variable, is often considered desirable, but not integral, to the interpretability of scale results (40).

3. Results

Demographics and details of disease indices are shown in Table 1. All of the respondents had at least one form of cardiovascular disease, 90 had experienced heart failure, 274 underwent coronary artery bypass surgery, and 529 had acute coronary syndrome. Further details of the specific study demographics are available from the corresponding author.

| Insert Table 1 here |

Mokken Scaling Procedure

The results of the Mokken analyses are shown in Table 2.
At c=.30 a moderate unidimensional scale ($H = .41$) including 12 of 14 HADS items was created. At c=.40 one moderate ($H = .49$) and one strong ($H = .60$) scale were formed with items 1, 3, 5, 6, 7, 9, 10, 11, 13 and 2, 4, 12, respectively. At c=.50 three strong scales were formed: items 1, 3, 5, 6, 13 ($H = .60$); items 2, 4, 12 ($H = .60$); items 7, 9 ($H = .51$).

Visual inspection of the 12 item scale extracted at c=.30 showed no violations of monotone homogeneity. Analysis of invariant item ordering revealed six items with no violations (items 4, 7, 8, 10, 11, 12; $H^T = .32$).

4. Discussion

The HADS has been extensively used under the assumption that it captures two dimensions (anxiety and depression); however, the current study reveals that the HADS captures a single overarching dimension: general psychological distress. The inclusion of 12 of the 14 HADS items compounded by the strength of the scale made the general psychological distress scale the most interpretable of the extracted models.

Several alternative models were revealed at the c=.40 and .50 levels; however, the low levels of interpretability did not provide robust evidence of the existence of the proposed two- and three-dimensional models. At c=.4, the scales provide neither evidence for the original anxiety-depression model nor a clear depiction of uniformly captured symptoms. Scale 2 ("Anhedonia") is
composed solely of anhedonic depression items, a dimension that has received support in previous models (41); however, anhedonic depression items 6 and 10 are also present on scale 1 (“General Distress”). Both the General Distress scale and the Anhedonia scale contain anhedonia items, in contrast to all of the prolific models of depression and anxiety, i.e. Caci (42), Dunbar (41), Friedman (43), Moorey (44), Razavi (45) and Zigmond (6), that do not split the depression items.

The c=.50 model reveals a three-dimensional structure combining four anxiety items and one depression onto a scale (“General Distress”), eliminating three depression items and forming a scale with the remaining depression items (“Anhedonia”) whilst creating a third, two anxiety item, scale (“Anxiety”). This three-dimensional model bears no resemblance to any of the established three-dimensional structures (41-43) or tripartite theories of depression and anxiety (46). Furthermore, this model also excludes four items, three of four items deemed by Zigmond & Snaith’s (1983) to be “mandatory”; therefore, providing insufficient evidence of the interpretability of this model.

The single-scale model provides a statistically viable, clinically useful measure of general psychological distress. Similar to the two- and three-scale models, the one-scale model could not differentiate between anxiety and depression; however, the inclusion of all but items 8 (“I feel as if I am slowed down”) arguably the only somatic/fatigue item, and 14 (“I can enjoy a good book or radio or TV program”), indicates that the scale is capturing something broader than the specific disorders of anxiety and depression. While invariant item ordering is a desirable attribute of a scale, it is not necessary (40); therefore,
failure to create an interpretable item invariant scale, due to the exclusion of more than half (8) of the 14 HADS items, does not impede the viability of the general psychological distress scale.

Although two items are not included in the scale, neither of these items were included amongst the eight items deemed “mandatory” by Zigmond & Snaith during the creation of the HADS (6). Given the average age of the respondents, item 8 (“I feel as if I am slowed down”) may have captured, and subsequently been confounded with, the benign slowing of activities associated with ageing. It is plausible that “slowed down” may have been misinterpreted as a generalized slowing rather than the intended depressive symptom of fatigue; capturing not only mood, but manifestations of somatic illness and/or cognitive slowing (11, 47). Item 14 could be criticized as having lost its cultural relevance and ability to assess anhedonia as it is intended to and has been shown to have poor discriminative abilities (28, 48). The validity of this item has been questioned in acute medical wards (11), suggesting that these activities may never be enjoyed in the hospital environment in which the HADS is completed. Despite eliminating two HADS items, the one scale model maintains its psychometric viability based on the questionable validity of the eliminated items.

While the symptomatic overlap demonstrated by the HADS is seen as a flaw in the bidimensional model, it provides strong evidence for the unidimensional general psychological distress model. In the general psychological distress model the intent is to capture and quantify symptoms of both anxiety and depression; therefore, whether or not a particular item captures
just one of the two disorders is irrelevant. In individuals with CVD, the clinical relevance of the general psychological distress model rests in the inextricable links between anxiety and depression in the course of the disease.

As outlined previously, anxiety and depression are both prognostic risk factors for CVD, highlighting the need to be able to assess these disorders via the HADS (49-55). Having the ability to quantify mutually exclusive levels of anxiety and depression is an integral component of assessment when the prognoses are specifically predicted by one or the other disorder; however, in the case of CVD, both anxiety and depression are prognostic risk factors, negating the need to differentiate between these disorders. Consequently, using the HADS as a unidimensional measure of general psychological distress will not impede its ability to provide timely and invaluable information to clinicians regarding the prognoses of individuals with CVD.

The unidimensional general psychological distress model of the HADS should be adopted by both clinicians and researchers alike. Given the degree of controversy that has arisen with regards to the structure of the HADS, in this case, simplicity may be the best policy. Rather than concerning themselves with categorically dividing symptoms with increasing levels of precision, adopting a more holistic viewpoint will enable clinicians to be able to be more cognizant of the overall psychological status of the individuals, avoiding a “missing the forest for the trees”-type situation. While a need may exist for other somatic disorders to be able to accurately differentiate between anxiety and depression, the current
study demonstrates that in the case of prognostic risk factors in CVD, the general psychological distress model of the HADS is an appropriate psychometric fit.

A limitation of the current study is the potential for a small number of participants to have been included across the samples. While this is highly unlikely, given that the samples were recruited over a 7-year timeframe, and across disparate locations, eliminating this possibility is not possible, due to the subsequent anonymization of the data and the passage of time since these data were collected.

The results from the current study provide strong evidence of the existence of a unidimensional model of general psychological distress in the HADS. The two-scale and three-scale models exhibited cross-loading of anxiety and depression items and confounding of symptoms across scales, resulting in the rejection of these models and provision of additional support for the ability of the unidimensional general psychological distress model to accurately capture symptoms of both anxiety and depression. The single latent variable model of general psychological distress found in the current study was both statistically viable and included 12 of the 14 HADS items, providing further evidence of its ability to capture symptoms of anxiety and depression.

The HADS has been used prolifically under the assumption of a bidimensional structure of anxiety-depression; however, the current study reveals that the underlying variable structure of the HADS is unidimensional. Being able to provide more compelling evidence of the structure of any psychometric tool should be a top priority for researchers; therefore, employing statistical methods
that reflect a higher degree of accuracy is imperative. The current study uses IRT methods in a relatively novel way to examine the structure of the HADS, a trend that will hopefully continue. Studies employing Rasch analysis in the examination of the latent factor structure of the HADS have emerged only in recent years (23, 25, 56), hopefully prompting a trend that will see the proliferation of IRT methods in psychometric validation studies. Employing a statistical method superior to the methods used in the vast majority of prior studies, a comprehensive and in-depth look at the underlying structure of the HADS was undertaken. The results from the current study provide compelling evidence of the unidimensional structure of the HADS and strongly advocate the adoption of the general psychological distress model.

Acknowledgements

The authors would like to thank Dr. Stefan Hoefer, Dr. Karen Morgan, and Dr. Neil Oldridge for generously contributing their data to the current study.
References


21. Denollet J, Pedersen SS, Daemen J, De Jaegere P, Serruys PW, Van Domburg RT. Reduced positive affect (anhedonia) predicts major clinical events


43. Friedman S, Samuelian J-C, Lancrenon S, Even C, Chiarelli P. Three-dimensional structure of the Hospital Anxiety and Depression Scale in a large


50. van Melle JP, de Jonge P, Spijkerman TA, Tijssen JGP, Ormel J, van Veldhuisen DJ, et al. Prognostic Association of Depression Following Myocardial ...
Infarction With Mortality and Cardiovascular Events: A Meta-analysis.


### Table 2. Loevinger H-values at c=.30, .40, .50

<table>
<thead>
<tr>
<th>Item</th>
<th>Prompt</th>
<th>Mean Score</th>
<th>c=.30</th>
<th>c=.40</th>
<th>c=.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I feel tense or wound up</td>
<td>.22</td>
<td>.45</td>
<td>.51</td>
<td>.59</td>
</tr>
<tr>
<td>2</td>
<td>I still enjoy things I used to enjoy</td>
<td>.44</td>
<td>.33</td>
<td>.61</td>
<td>.61</td>
</tr>
<tr>
<td>3</td>
<td>I get a sort of frightened feeling as if something awful is about to happen</td>
<td>.28</td>
<td>.46</td>
<td>.53</td>
<td>.60</td>
</tr>
<tr>
<td>4</td>
<td>I can laugh and see the funny side of things</td>
<td>.66</td>
<td>.40</td>
<td>.58</td>
<td>.58</td>
</tr>
<tr>
<td>5</td>
<td>Worrying thoughts go through my mind</td>
<td>.31</td>
<td>.46</td>
<td>.54</td>
<td>.60</td>
</tr>
<tr>
<td>6</td>
<td>I feel cheerful</td>
<td>.54</td>
<td>.46</td>
<td>.51</td>
<td>.56</td>
</tr>
<tr>
<td>7</td>
<td>I can sit at ease and feel relaxed</td>
<td>.36</td>
<td>.39</td>
<td>.40</td>
<td>.51</td>
</tr>
<tr>
<td>8</td>
<td>I feel as if I am slowed down</td>
<td>.12</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>I get a sort of frightened feeling like &quot;butterflies&quot; in the stomach</td>
<td>.42</td>
<td>.38</td>
<td>.42</td>
<td>.51</td>
</tr>
<tr>
<td>10</td>
<td>I have lost interest in my appearance</td>
<td>.59</td>
<td>.39</td>
<td>.44</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>I feel restless as if I have to be on the move</td>
<td>.23</td>
<td>.32</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>I look forward with enjoyment to things</td>
<td>.55</td>
<td>.33</td>
<td>.59</td>
<td>.60</td>
</tr>
<tr>
<td>13</td>
<td>I get sudden feelings of panic</td>
<td>.37</td>
<td>.49</td>
<td>.56</td>
<td>.64</td>
</tr>
<tr>
<td>14</td>
<td>I can enjoy a good book or radio or TV program</td>
<td>.68</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Overall scale*: .41 .49 .60 .60 .60 .51