Instrument to measure adherence in hypertensive patients: contribution of Item Response Theory

ABSTRACT

OBJECTIVE: To analyze, by means of Item Response Theory, an instrument to measure adherence to treatment for hypertension.

METHODS: Analytical study with 406 hypertensive patients with associated complications seen in primary care in Fortaleza, CE, Northeastern Brazil, 2011 using “Item Response Theory”. The stages were: dimensionality test, calibrating the items, processing data and creating a scale, analyzed using the gradual response model. A study of the dimensionality of the instrument was conducted by analyzing the polyechoric correlation matrix and factor analysis of complete information. Multilog software was used to calibrate items and estimate the scores.

RESULTS: Items relating to drug therapy are the most directly related to adherence while those relating to drug-free therapy need to be reworked because they have less psychometric information and low discrimination. The independence of items, the small number of levels in the scale and low explained variance in the adjustment of the models show the main weaknesses of the instrument analyzed. The “Item Response Theory” proved to be a relevant analysis technique because it evaluated respondents for adherence to treatment for hypertension, the level of difficulty of the items and their ability to discriminate between individuals with different levels of adherence, which generates a greater amount of information.

CONCLUSIONS: The instrument analyzed is limited in measuring adherence to hypertension treatment, by analyzing the “Item Response Theory” of the item, and needs adjustment. The proper formulation of the items is important in order to accurately measure the desired latent trait.

INTRODUCTION

Measuring adherence to high blood pressure (HBP) treatment is a complex task. Studies have been conducted to understand factors affecting this public health problem, as well as creating instruments which enable the degree of adherence to be measured in a reliable way.\textsuperscript{2,3,5,6,14} The most common method is to use interviews and questionnaires due to their accessibility and low cost.\textsuperscript{12}

A recent and suitable methodology for evaluating instruments which measure adherence to high blood pressure treatment is provided in the form of Item Response Theory (IRT). This theory includes a set of models for latent variables which propose to represent the relationship between the probability of a respondent giving a particular response to an item, their latent trait and the characteristics (parameters) of the item. Latent traits refer to characteristics of the individual which cannot be directly observed and are measured through secondary variables related to them (items of the instrument).\textsuperscript{5}

Few studies have evaluated measuring instruments in the sphere of public health using IRT\textsuperscript{6,7,17} and none of them involve applying IRT to adherence to HBP treatment, leaving the field open to research. Thus, the aim of this study was to analyze an instrument for measuring adherence to HBP treatment using IRT.

METHODS

Analytical study which investigated the applicability of an instrument for measuring adherence to HBP treatment, developed by Moreira,\textsuperscript{4} using IRT. This instrument has been used in studies carried out in the Brazilian capital, using a Likert style scale and measuring the medical rather than pharmacological aspect of adherence. It contains ten items: salt intake, fat intake, not smoking, not drinking alcohol, doing physical activities, dealing with stress, appropriate use of medications, attending appointments and clinical data (Body Mass Index – BMI and blood pressure – BP). Each item has five possible responses which range from 0.0 to 1.0 points. The total possible score for the questionnaire is 10 and the author standardized the values considering the score obtained for: ideal patient (9 to 10 points); slight lack of adherence (from 7 to < 9); moderate lack of adherence (from 5 to < 7); serious lack of adherence (from 3 to < 5); and very serious lack of adherence (from 1 to < 3).

The study universe was composed of the medical records of hypertensive patients registered in the HIPERDIA (Registering and Monitoring System for Hypertensive and Diabetic Patients), Fortaleza, CE, Northeastern Brazil, 2011, totaling 14,200 records. Of these, 1,315 had the associated complications and constituted the population studied. In order to calculate the sample, an estimated prevalence of adherence of 40%, according to the mean of prevalence in studies on adherence,\textsuperscript{4} a confidence interval of 95%, error of 4% and an added 10% to cover losses and refusals were adopted, making a total of 440 hypertensive patients. There were 34 subjects excluded (deaths, not at home, cognitive deficiency) and 406 people inserted in the final study. These hypertensive patients came from the 15 family primary health care units (FPHC) with the highest number of cases, distributed throughout the entire municipal area. The stages of analyzing the instrument were the following: dimensionality test, calibrating the items, processing the data and constructing the scale. The Samejima Graded Response model,\textsuperscript{15} created with the objective of obtaining other information from the individual’s responses than just the simple correct or incorrect response to the items, was used to analyze the data. This model is represented by the following equation, in which $P_i(a) = \frac{1}{1 + e^{-a(b_i-a_c)}}$

The parameter $a$ is equal in all of the categories of the scale of item $i$ and the parameters $b_i$ represent the difficulty of the $k$th category of item $i$.

Multilog\textsuperscript{6} software was used to calibrate the items and, later, to estimate the scores. Calibrating the items involved estimating the parameters $a_i$ and $b_i$, and verifying which of them was estimated satisfactorily to be used for calculating the scores of the latent traits and constructing the scale.

The anchor items were defined during the stage of constructing the scale. Andrade et al\textsuperscript{11} state that an item is an anchor for level A if it simultaneously fulfills the three conditions, for $Y < Z$:

a) $P(U = 1 | a = Z) \geq 0.65$;

b) $P(U = 1 | a = Y) < 0.50$;

c) $P(U = 1 | a = Z) \cdot P(U = 1 | a = Y) \geq 0.30$.

As it is difficult to fulfill all of these conditions, items which met two of the three conditions, so-called quasi anchors, were considered.

A study of the dimensionality of the instrument was conducted using the matrix of polychoric correlation\textsuperscript{6} and the principal components analyzed, using parallel...
analysis which enables the dimensionality of the instrument to be assessed. This study was carried out using the complete instrument: ten items with all response categories, even considering the lower levels of adherence with few responses.

The dimensionality analysis was complemented by factorial analysis of all the data, based on Multidimensional Item Response Theory (MIRT) models. This analysis enabled the correlation of the items with possible latent traits (or subjacent factors) to be assessed and verified whether all of the items belonged to a single dimensional model or whether there were sub-conjuncts of items which justified more dimensions. The parallel analysis with the polychoric correlation matrix was carried out using the psychometric “psych” package and the factorial analysis of the complete data using the “mirt” package, both free software R. Hayton et al stated that, even if little used, parallel analysis is an accurate method of estimating the number of factors.

Linear transformation was carried out to provide changes of the scale in order to relate the IRT score (estimates of adherence to HBP treatment, according to the graded response model, with a mean of zero and standard deviation of one) with the total score, for which classifications of the individuals exist regarding the intensity of adherence to HBP treatment.

This study was approved by the Research Ethics Committee of the Universidade Estadual do Ceará Process nº 10725637-1, cover sheet (CS) 401985.

RESULTS

The results of the analysis of the principal components and the corresponding parallel analysis, both in terms of the polychoric correlation matrix, indicated a stronger dimension, but one which represented little more than 20% of the variance of the items, in addition to two others, each with a participation of less than 15% of explained variance (Figure 1). The other main components had explicative power below the dotted line, i.e., they were not significant and were, therefore, rejected.

The majority of coefficients in the polychoric correlation matrix had positive, albeit low, values, indicating low correlation between the items of the instrument and it was therefore not possible to summarize the responses by one or two subjacent factors.

Figure 1. The principal component and parallel analysis of the polychoric matrix of correlation.

To understand the latent traits (or subjacent factors) linked to the patterns of the instrument’s responses, factorial analysis of the complete data based on MIRT models was carried out. Table 1 shows the factorial loads of the single, bi and tridimensional models, as well as the estimates of variance explained by each factor.

The factorial loads may be interpreted as correlations between the items and the subjacent factors. These correlations were low in the instrument, especially with item 5 (being sedentary). The fit of the bi-dimensional model showed separation between items 1 to 4 (F2) with items 7 and 8 (F1). Dimension F1 was linked to adherence to treatment with medication and dimension F2 to non-medication treatment. Analysis of the tridimensional model reinforced the interpretation of two sets of items (1 to 4 and 7 to 10), with being sedentary (item 5) as an exclusive factor. This item had a different dimension to the others and should be excluded by IRT in the construction of a single dimensional analysis.

In spite of the three models being adjusted, explained variance was low. It was decided to use the single dimensional IRT Graded Response Model to study the possible dominant dimension.

During the calibration of the items, the response categories were grouped so that the parameters could be appropriately estimated. The items referring to salt intake; fat intake; not smoking; not drinking; doing physical activity and dealing with stress ended up with two response categories. The other items (appropriate use of medication, attending appointments and data on BMI and BP) had three response categories which

---

kept the same order in the final result (Table 2). The item referring to regularly doing physical activity was excluded from the analysis.

The majority of items had two parameters, as their estimates did not converge on a scale of three points and it was necessary to dichotomize them. These factors contributed positively or negatively to adherence to HBP treatment, without graded responses.

The items with the highest parameters, or which best discriminated hypertensive patients who did and did not adhere to HPB treatment, were the references for the appropriate use of medication and attending appointments/period in which they are marked in the FPHC. The items referring to effectively dealing with stress and BMI were the most difficult to respond satisfactorily. This lead to their high parameter b values, in which the most positive responses for adherence to treatment were given only by hypertensive patients who had high adherence to their HBP treatment. The b parameter indicates the position on the scale at which the item had the most data (Table 2).

The items referring to appropriate fat intake, effectively dealing with stress and BMI had b₂ parameters located above the mean on the scale (0.1) and therefore tended to be more representative of individuals with greater adherence to treatment, although the, low, a parameters meant the items could not be classified as anchors. In spite of the items referring to appropriate use of medication, attending appointments and blood pressure having reasonable parameter a estimates for discriminating patients adhering to treatment from those who were not, it was observed that the b₂ and b₃ values were close to the mean zero, i.e., the patient did not necessarily have to adhering strictly to the treatment in order to respond to these items satisfactorily.

A parameters were low for the majority of the items. Figure 2 shows the curves characteristic of the items with two or three response categories (fat intake and medication use, respectively).

The first curve (dichotomized item) shows the location of parameter b between 0 and 1. This parameter indicates that individuals with a score of 0.536 had a 50% probability of responding to this item satisfactorily, or that 50% of individuals with a 0.536 degree of adherence had low fat intake. The a parameter at the value of 0.894 determines the slope of the curve when probability is 50%. The curve (b) represented the item with three response categories with two parameters of difficulty. The first of these (b₂ = -1.78) was at the intersection of curves 1 and 2, and the second parameter (b₃ = -0.16) at the intersection of curves 2 and 3. The sum of the probabilities of the three curves for the same score was 1 (100%), according to the condition

<table>
<thead>
<tr>
<th>Item</th>
<th>Single dimensional</th>
<th>Bi-dimensional</th>
<th>Tridimensional</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Salt intake</td>
<td>0.25</td>
<td>0.01</td>
<td>0.61</td>
</tr>
<tr>
<td>2. Fat intake</td>
<td>0.41</td>
<td>0.23</td>
<td>0.52</td>
</tr>
<tr>
<td>3. Non smoking</td>
<td>0.45</td>
<td>0.28</td>
<td>0.42</td>
</tr>
<tr>
<td>4. Not drinking alcohol</td>
<td>0.42</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>5. Regularly doing exercise</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.13</td>
</tr>
<tr>
<td>6. Dealing with stress</td>
<td>0.17</td>
<td>0.11</td>
<td>0.18</td>
</tr>
<tr>
<td>7. Appropriate use of medication</td>
<td>0.69</td>
<td>0.69</td>
<td>0.17</td>
</tr>
<tr>
<td>8. Attending appointments</td>
<td>0.57</td>
<td>0.66</td>
<td>0.03</td>
</tr>
<tr>
<td>9. Body mass Index</td>
<td>0.23</td>
<td>0.28</td>
<td>-0.03</td>
</tr>
<tr>
<td>10. Blood pressure</td>
<td>0.37</td>
<td>0.38</td>
<td>0.07</td>
</tr>
<tr>
<td>Explained variance</td>
<td>0.16</td>
<td>0.13</td>
<td>0.11</td>
</tr>
</tbody>
</table>

* Includes Varimax rotation

Table 2. Estimate of the items’ parameters. Fortaleza, CE, Northeastern Brazil, 2011.

<table>
<thead>
<tr>
<th>Item</th>
<th>a</th>
<th>b₂</th>
<th>b₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Appropriate salt intake</td>
<td>0.43</td>
<td>-0.51</td>
<td>-</td>
</tr>
<tr>
<td>2. Appropriate fat intake</td>
<td>0.89</td>
<td>0.54</td>
<td>-</td>
</tr>
<tr>
<td>3. Non smoking</td>
<td>0.84</td>
<td>-2.89</td>
<td>-</td>
</tr>
<tr>
<td>4. Not drinking alcohol</td>
<td>0.79</td>
<td>-2.18</td>
<td>-</td>
</tr>
<tr>
<td>5. Regularly doing exercise</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. Dealing with stress</td>
<td>0.62</td>
<td>3.70</td>
<td>-</td>
</tr>
<tr>
<td>7. Appropriate use of medication</td>
<td>1.42</td>
<td>-1.78</td>
<td>-0.16</td>
</tr>
<tr>
<td>8. Attending appointments/in the period in which they are made in the Family Primary Health Care Unit</td>
<td>1.18</td>
<td>-1.79</td>
<td>-0.89</td>
</tr>
<tr>
<td>9. Body mass Index</td>
<td>0.39</td>
<td>2.85</td>
<td>-</td>
</tr>
<tr>
<td>10. Blood pressure</td>
<td>0.78</td>
<td>-3.24</td>
<td>0.27</td>
</tr>
</tbody>
</table>
established in the Samejima model, in which the \( a \) parameter is equal for all of the curves.

The value of \( a \) must be > 0.1 for the item to be an anchor or quasi anchor; thus, only items 07 and 08 could be considered for constructing the scale. The items used were those referring to appropriate medication use (anchor for category 1) and attending appointments (anchor for category 1 and quasi anchor for category 2). The other items, although enshrined in the literature as related to HBP treatment adherence, were not included as they had low \( a \) parameters, i.e. they did not discriminate hypertensive patients adhering to treatment from those who were not, and were therefore not included in the scale.

Linear transformation was carried out on the estimates of the index of adherence to HBP treatment with a mean and standard deviation of 0.1 and 30.5, respectively, in order to better understand the results. The items were analyzed and placed on a scale with three levels (Table 3). Those being treated for HBP with a score < 25 were at the level of very serious non-adherence. At the next level, the patients had the same attitude with regards taking medication, but had improved appointment attendance. The patients were at the same level of adherence after level 35 of the scale, although with different scores.

The scale had few levels and little information in each of them, as it was based on only two items from the instrument.

**DISCUSSION**

Although questionnaires are very frequently used to assess adhesion to HBP treatment, this method has its limitations. Márquez Contreras et al warned that questionnaires have low negative predictive value, confirmed by the fact that 43% of users with uncontrolled high blood pressure reported that they were correctly following their treatment. Another important factor is reported by Santa Helena et al, and refers to the conditions in which the questionnaire is carried out, dealing with undesirable behavior, which may intimidate the interviewees. In such cases, the method of completing the questionnaire may lead to adherence being overestimated, especially when it is completed by an interviewer, as patients tend to optimize their adherence in their responses, seeking the interviewers approval.

It is important that the items in the instrument be assessed, so that they provide the greatest amount of information possible. IRT provides additional data, as it enables the identification of which dimension has more and less weight by evaluating latent traits. Moreover, it considers these dimensions in calculating latent traits for each respondent, in which the items are evaluated with different weightings. This differs from classical analysis in which all the items have the same weight.

**Table 3.** Scale of adherence to high blood pressure treatment. Fortaleza, CE, Northeastern Brazil, 2011.

<table>
<thead>
<tr>
<th>Level on the scale</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>Does not take medication, or does not place much importance on taking it. Only takes it when blood pressure rises or has difficulties in taking the medication due to the side effects. Does not attend appointments or only attends when feeling poorly or blood pressure rises. Wants to alter behavior but has not managed to.</td>
</tr>
<tr>
<td>30</td>
<td>Does not take medication, or does not place much importance on taking it. Only takes it when blood pressure rises or has difficulties in taking the medication due to the side effects. Rarely misses appointments.</td>
</tr>
<tr>
<td>35</td>
<td>Rarely forgets to take the medication. Is scrupulous in attending appointments.</td>
</tr>
</tbody>
</table>
The parameters of each item do not depend on the other items in the test. The score of the test, or subject’s latent trait, is found depending on their responses. In addition, it enables us to know which items can be considered easier and or more difficult to respond to satisfactorily, as one of the advantages of using IRT is that, in addition to providing a score for the latent trait, it positions the items on the same scale. This characteristic means that the exact significance of the score can be understood and it is not just an indicator or the degree, or lack of it, of adherence, as shown in classical analysis. Although individuals may have the same total score, they are not considered to be in the same case as they do not have the same profile of responses.

Another property of this model is that it enables a plan of objectives to be created for each user, according to their individual score. A patient in, for example, level 25 of the scale knows exactly that they have to be more assiduous in attending appointments in order to climb to level 30 of the scale, as at this level the hypertensive patient rarely misses appointments. This facilitates patients’ understanding and improves health care professionals’ degree of control and decision making. The health care team can use different strategies to present the results to the patient, as well as positively reinforcing their improvements on the scale for improving their adherence. The patient and the health care professional both perceive improvements in adherence to treatment with greater clarity.

The items referring to appropriate use of medication and attending appointments performed well, as they had the greatest power of discriminating between individuals who are adhering to their HBP treatment and those who are not. The items referring to treatment with medication are more or less directly related to adherence in the majority of hypertensive patients.

Items related to non-medication treatment need to be reformulated, as they possess less psychometric information and poor discrimination. This reveals that the items are poorly formulated and need to be adjusted or belong to another dimension, as they are not successful in discriminating between individuals with greater and lesser latent traits.

The item referring to doing regular physical activity was excluded, as its presence meant that the algorithm estimation did not converge. This lack of convergence may be due to the following factors: the item being poorly drawn up, low number of respondents, the item not discriminating between individuals adhering to treatment from those who are not, the item being incompatible with this dimension of analysis or a combination of these factors.

An alternative for this result is the reformulation of the items so as to be more discriminatory. The item on smoking, for example, did not take into account whether the individuals responded satisfactorily to the item were non-smokers before treatment or whether, to adhere to the treatment, they cut down or stopped smoking when they started HBP treatment.

The item referring to dealing with stress effectively was considered the most difficult to respond to satisfactorily, as it had the highest b value. This parameter discriminates patients with better adherence to treatment from those with poor adherence, if their value is above 0.70. Experiencing stressful situations may be inevitable. For patients with HBP, this is an aggravation, as they are living with a disease which requires continuous treatment, and this is generally a source of stress as it calls for lifestyle changes.

The independence of the items (maintaining the latent trait constant, the individuals’ responses to any of the items are statistically independent), the small number of levels of the scale and the low explained variance in the fit of the models show the main weaknesses of the instrument analyzed. It provides little information and therefore the items need to be altered or modified so that it is possible to identify levels of adherence to treatment with greater clarity and propose interventions according to the scale.

A weakness of this study is the analysis of an instrument with low amounts of psychometric data, producing a scale which included two items evaluating adherence to HBP treatment. It is possible that there are other gradations with more associated information, as the items of the instrument analyzed are not sufficient to create and explain other levels of adherence. Such a result denotes the need to reformulate the set of items in such a way as to enable a broader scale to be constructed.

Constructing instruments requires attention when designing the items, clarity of approach and focus on the latent trait to strengthen the discriminatory power, i.e., to produce items with high a parameters. Professionals and researchers who propose using questionnaires for evaluation are guided by enshrined theories in drawing up the items and in the analysis of their results, at the risk of arriving at vague and unhelpful conclusions.

The instrument in question, when analyzed using IRT only allowed adherence to HBP treatment with medication to be measured, as it did not have accurate items referring to lifestyle changes which could have been used as anchors. The b parameters, as they are measured on the same scale as latent traits, provide an idea of the position of the item on the scale and the a parameter decides whether the item can be used as an anchor. The analysis in Figure 2 enables it to be supposed to what extent an item may be satisfactory for constructing the scale, although it is necessary to confirm this by analyzing anchor items. This is because the latter is
more accurate, it enables a contextualized view of the behavior of each response category in each item.

The instrument is limited as it measures adherence to HBP treatment using IRT analysis and needs some adjustments. The appropriate formulation of the items is important to measure the desired latent trait accurately. It is suggested that an instrument to measure adherence to HBP treatment be constructed, with items suitable for IRT analysis, preferably of a single dimensional nature, with higher levels of validity and reliability and that they be used with patients with different levels of adherence, with the aim of reducing errors in measurement.

REFERENCES


Study based on the master dissertation of Daniele Braz da Silva, entitled: *Hipertensão arterial e complicações associadas: análise do risco cardiovascular e da adesão ao tratamento em usuários do Sistema Único de Saúde* presented to the Universidade Estadual de Fortaleza, Ceará, in 2011. The authors declare that there are no conflicts of interest.