Construct Validity of Symptom Checklist-90-Revised (SCL-90-R) and General Health Questionnaire-28 (GHQ-28) in Patients with Drug Addiction and Diabetes, and Normal Population

Abolfazl ARDAKANI †, *Tahereh SEGHATOLESLAM †,2, Hussain HABIL †,2, Fahimeh JAMEEI †, Rusdi RASHID †, Alireza ZAHIRODIN ‡, Farid MOTLAQ †, Abbas MASJIDI ARANI ‡

1. Dept. of Psychological Medicine, University of Malaya Center of Addiction Sciences (UMCAS), Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
2. Dept. of Clinical Psychology, Shahid Beheshti University of Medical Sciences, Tehran, Iran
3. Dept. of Psychiatry, Mahsa University, Kuala Lumpur, Malaysia
4. Dept. of Psychiatry, Behavior Research Centre of Shahid Beheshti University of Medical Sciences, Tehran, Iran

*Corresponding Author: Email: drtseghatoleslam@yahoo.com

(Received 26 May 2015; accepted 12 Dec 2015)

Abstract

Background: Given that validity is the baseline of psychological assessments, there is a need to provide evidence-based data for construct validity of such scales to advance the clinicians for evaluating psychiatric morbidity in psychiatric and psychosomatic setting.

Methods: This comparative cross-sectional study aimed to investigate the construct validity of the Malaysian version of the GHQ-28 and the SCL-90-R. The sample comprised 660 individuals including diabetics, drug dependents, and normal population. The research scales were administered to the participants. Convergent and discriminant validity of both scales were investigated by Confirmatory Factor Analysis (CFA) using AMOS. The Pearson correlation coefficient was utilized to obtain the relationship between the two scales.

Results: The internal consistency of the GHQ-28 and SCL-90-R were highly acceptable, and confirmatory factor analysis confirmed the convergent validity of both scales. The results of this study revealed that the construct validity of GHQ-28 was acceptable, whereas discriminant validity of SCL-90-R was not adequate. According to Pearson correlation coefficient the relationships between three common subscales of the GHQ-28 and SCL-90-R were significantly positive; somatization (r=0.671, P<0.01), Anxiety (r=0.728, P<0.01), and Depression (r=0.660, P<0.01).

Conclusions: This study replicated the construct of the Malaysian version of GHQ-28, yet failed to support the nine-factor structure of the SCL-90-R. Therefore, multidimensionality of the SCL-90-R as clinical purposes is questionable, and it may be a better unitary measure for assessing and screening mental disorders. Further research need to be carried out to prove this finding.

Keywords: SCL-90-R, GHQ-28, Construct validity, Mental health

Introduction

Drug addiction and diabetes are two prevalent chronic diseases, and comorbidity of psychiatric disorders with these stress-related diseases is well documented (1-7). Moreover, such comorbidity interferes with the management of these patients (8). Oftentimes, evaluating clinicians are faced with considerable confusions for diagnosing comorbid psychiatric disorder in these patients (5, 8-
10); therefore, this points a great need to provide proper psychological scales to evaluate and understand such comorbidities in patients with chronic disease (11).

Nowadays, self-rating assessments have become necessary to advance the clinician in evaluating and understanding psychiatric comorbidities in chronic patients. Nevertheless, use of these scales can yield benefits if they are validated (7, 12-15). Indeed, validity is the baseline of measurement as a crucial indicator of psychometric quality, and construct validity is essential in these disciplines (16).

Among valid and reliable self-rating scales, the Symptom CheckList-90-Revised (SCL-90-R) and General Health Questionnaire-28 (GHQ-28) are both appropriate psychological instruments for assessing psychiatric disorder in chronic patients (17) i.e. drug addicts and diabetics. The concurrent validity of the GHQ-28 against SCL-90-R indicated a meaningful correlation between scores of samples within the scales of these instruments (18). Several studies have investigated the construct validity of the SCL-90-R and GHQ-28. The factor structures of GHQ-28 have been consistent with its original dimensions (18-22); whereas, the documents for construct validity of the SCL-90-R are controversial.

The SCL-90-R nine subscales are originally proposed as representing a multidimensional structure (29). Several studies (19, 23-29) replicated the construct validity of SCL-90-R, and supported the multidimensional structure of the SCL-90-R; meanwhile, other studies failed to replicate this fact, and suggested a unitary measure for global distress instead of multi-dimensional measure (19, 23, 30-34). Although, the recent study done in Malaysia revealed a good reliability for the SCL-90-R (35); yet, there is no study has examined the construct validity of this self-report inventory in Malaysia.

With regard to the contributions of psychological assessments in case finding, case conceptualization, and treatment planning of patients with comorbid psychiatric disorders (18, 36), validation of the SCL-90-R and GHQ-28 would be important to ensure accurate psychiatric information especially in patients with stress-related disease like drug addiction and diabetes. Furthermore, continual improvement and innovation in validating such scales is necessary to all research settings; in fact, they must be replicated in different setting, language, and culture, it is because, they rely on the judgment of the respondents, and are sensitive to the linguistic or cultural factors (34).

Despite the significance of studying the psychometric properties of psychological instruments, there do not appear to be any study investigating the construct of SCL-90-R and GHQ-28 scales in Malaysia.

Therefore, the aim of this study was to investigate the constructs (convergent and discriminant validity) of SCL-90-R and GHQ-28 scales in chronic patients (i.e. drug addiction and diabetes) and normal population in Malaysia. This study would be a step toward establishing the construct validity of the Malaysian version the SCL-90-R and GHQ-28, and add substantial information on the construct validity of SCL-90-R and GHQ-28. Furthermore, the present study advances the literature through the comparative investigation of the mentioned scales.

Materials and Methods

Sample
This comparative cross-sectional study was carried out in Kuala Lumpur, Malaysia in 2012-13. The sample frame consisted of Malaysian adults aged 18 and above with chronic disease and normal people as a control group. The clinical samples were patients with drug addiction and diabetes (two prevalent stress-related diseases), and were recruited from six clinics, including three diabetic clinics and three rehabilitation clinics. The non-clinical sample was consisted of the normal people free of any chronic disease, and was selected nearby the selected clinics in the study. With regard to the results of sample size calculations, a total of 660 subjects were included in the study through simple random sampling method. The method of calculation for the sample size was the method described by Kelsey et al. (37).

Exclusion criteria

Available at:  http://ijph.tums.ac.ir
The following subjects have been excluded in this study (the registration criteria):

(i) Females: to prevent the gender bias, the females were excluded from the study populations, as only 12 females MMT patients referred to the clinics.

(ii) Non-Malaysian citizens (for the purpose of homogeneity).

(iii) Those unwilling to answer the questionnaire.

(iv) Severe ill patients or who needed to be referred to emergency room.

(v) Those with communication difficulties; e.g. deafness, blindness, muteness, or dysarthria.

Adding to exclusions, to obtain accurate data, if any patient revealed any withdrawal symptoms, acute intoxication, or acute psychosis would be excluded from the study, and also if he had an organic brain problem, or an organic brain problem made it difficult for giving reliable information, he would be excluded, as well.

**Ethical notes**

Prior to collect the data, the ethical approval was obtained from the Medical Ethics Committee of University of Malaya Medical Center (UMMC). The ethical conditions of participation including voluntary participation, privacy, anonymity and confidentiality were explained to the respondents. Based on the written informed consent, they would be free to withdraw from the study at any point. All subjects in this study declared their agreement, and to ensure anonymity, they were mentioned that putting their names or other identifying notation on the questionnaires is optional. Afterward, they were invited to complete the screening package translated to the Bahasa Malaysian.

**Instruments**

The instruments applied in this study were the Bahasa Malaysian version of Symptom Checklist-90-Revised (SCL-90-R) and General Health Questionnaire-28 items (GHQ-28). These scales were designed to determine the mental health status and psychiatric symptoms at a specific time. They are brief instruments being sensitive, time efficient and well-validated for assessing common psychiatric disorders (38), and are successfully used in both clinical and epidemiological studies of mental disorders (18, 39, 40).

**The Symptom CheckList-90- Revised:** The SCL-90-R developed by Derogatis is a self-report instrument containing 90 items and designed to measure nine current psychiatric symptoms, as well as psychological distress. The SCL-90-R subscales assess followings psychiatric symptoms: Somatization, Obsessive Compulsive Disorder, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism. Each item has five following response categories: 0 = Not at all, 1= little, 2 = some, 3 = very, 4 = severe. Albeit this instrument has been developed in 1970’s, it is still useful to understand psychiatric disorder as classified in DSM-IV (28).

**The General Health Questionnaire-28:** The GHQ-28 developed by Goldberg has only 28 items completed in only 10-12 min. The scale helps measuring 4 groups of psychiatric morbidity containing somatization, anxiety, social dysfunction, and depression, and provides overall mental health at a specific point in time, as well. Each item has four response categories such as "better than usual"; "same as usual"; "worse than usual"; or "much worse than usual". Respondents rate each item according to how they have recently felt their experience. The higher GHQ-28 scores, the severe psychiatric disorder in which the subject may suffer from.

**Statistical analysis**

The data was analyzed using Statistical Package for the Social Sciences software (SPSS) version 20, and those with more than 20% missing data were excluded. The confirmatory factor analysis (CFA) and measurement model were conducted by utilizing AMOS 21. Convergent and discriminant validity of the constructs were examined by the following indexes: Composite Reliability (CR), Average Variance Extracted (AVE), Maximum Shared Squared Variance (MSV), and Average Shared Squared Variance (ASV). These indexes provide useful evidences to establish construct reliability and validity. Pearson correlation was utilized to
prove the relationship between the research instruments, as well as their dimensions.

Results

The entire sample fully completed the Malaysian versions of GHQ-28 (BM) and SCL-90-R (BM) in the study. The study samples were partly similar in the socio-demographics characteristics. The mean age of the study sample was 40.45 (SD = 10.735), and the majority of them were Malay (78.9%), Muslim (79.2%). Result of the present study is based on Goodness of Fit of the measurement models obtained for both scales by using AMOS-21. Indeed, the measurement model was adjusted to provide a few useful indexes putting to test the construct validity and reliability of the GHQ-28 and SCL-90-R.

Convergent and discriminant validity of the GHQ-28

The measurement model showed satisfactory fit statistics; (Chi-squared=810.8, df=305, RMR=.023, CFI = .962, AGFI = .898, GFI = .918, RMSEA=0.05). To measure the consistency of related questions in the GHQ-28, the subscales of the scale were subjected to reliability analysis. According to the results of Composite Reliability (CR), all factor loadings were higher than acceptable threshold level 0.50. The composite reliability of GHQ-28 subscales ranged from 0.915 to 0.859, accordingly, the highest consistency was related to Somatization items (0.915), and the lowest consistency was related to Depression items (0.85); thus, the reliability of the scale was established.

Following Confirmatory Factor Analysis (CFA) and determining the measurement model the construct validity was evaluated to calculate the discriminant and convergent validity of the subscales. All variables have met a sufficient convergent validity in the study; indeed, Composite Reliability (CR) is between 0.875 and 0.915 in the study, and the standardized factor loadings are significant; likewise, as illustrated in Table 1, Composite reliabilities (CR) of all dimensions have exceeded the minimum limit (0.70), as well as the average variance extracted (AVE). Furthermore, AVE is all above the acceptable threshold level (0.50) implying that greater than fifty percent of the variances have been observed.

To obtain discriminant validity, the maximum shared squared variances (MSV) has been compared between factors and Average shared squared variance (ASV) as well as average variance extracted (AVE). Indeed, the discriminant validity was obtained by putting to test the average variance extracted for all construct against squared correlations (shared variance) between the construct and all other constructs in the model. As all of the constructs have exceeded the test, the discriminant validity has been confirmed in the current study. Furthermore, AVE squared root of each dimension surpassed the squared correlation between two dimensions suggesting that an adequate discriminant validity for all of the constructs (Table 2).

Convergent and discriminant validity of the SCL-90-R

The measurement model showed satisfactory fit statistics (Chi-squared = 12587, df = 3265, RMR = 0.038, CFI = 0.785, AGFI = 0.636, GFI = 0.656, RMSEA =0.06). To measure the consistency of related questions in the GHQ-28, the subscales of the scale were subjected to reliability analysis.

<table>
<thead>
<tr>
<th>Construct and indicators (items/parcels)</th>
<th>CR</th>
<th>AVE</th>
<th>MSV</th>
<th>ASV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somatization</td>
<td>0.915</td>
<td>0.607</td>
<td>0.336</td>
<td>0.311</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.891</td>
<td>0.541</td>
<td>0.473</td>
<td>0.351</td>
</tr>
<tr>
<td>Social Dysfunction</td>
<td>0.875</td>
<td>0.507</td>
<td>0.473</td>
<td>0.328</td>
</tr>
<tr>
<td>Depression</td>
<td>0.859</td>
<td>0.509</td>
<td>0.336</td>
<td>0.277</td>
</tr>
</tbody>
</table>

Composite Reliability (CR), Average Variance Extracted (AVE), Maximum Shared Squared Variance (MSV), and Average Shared Squared Variance (ASV)
Composite Reliability index revealed that all factor loadings of the scale were significant and greater than 0.50 level, and the consistency of the SCL-90-R was in a high acceptable level (Table 3) in which the consistency of its subscales ranged from 0.955 to 0.904, the highest consistency was related to the Anxiety items (0.955), and the lowest consistency was related to Aggression items (0.904). Therefore, the reliability of the questionnaire was established, as well.

On the basis of the CFA and measurement model the SCL90-R was evaluated for construct validity (Table 3); a sufficient convergent validity have been provided for all variables. The results show that Composite Reliability (CR) for all dimensions are above 0.9 and also are larger than the AVE. Estimated AVE were all above the acceptable threshold 0.5, suggesting that greater than one-half of the variances observed. MSV and ASV were larger than AVE for SCL-90-R dimensions; thus, the discriminant validity was not met for this questionnaire. Based on Table 4, the squared root of AVE of each dimension is less than the correlations between two dimensions. Thus, for all of the constructs the discriminant validity is not adequate.

Pearson correlation coefficient was utilized for evaluation of the relationship between total mean scores of the GHQ-28 and SCL-90-R as well as their common subscales including; Somatization, Anxiety, and Depression. The relationships were significantly positive; total mean scores (r= 0.765 P<0.01), Somatization (r=0.671, P<0.01), Anxiety (r=0.728, P<0.01), and Depression (r=0.660, P<0.01).

**Discussion**

The current study aimed to examine the construct validity of the Malaysian version of GHQ-28 and SCL-90-R as two psychometric scales in two clinical sample, as well as normal sample. Following the measurement model and Confirmatory Factor Analysis (CFA), convergent and discriminant validity of both scales were evaluated.

<table>
<thead>
<tr>
<th>Construct and indicators (items/parcels)</th>
<th>CR</th>
<th>AVE</th>
<th>MSV</th>
<th>ASV</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOM</td>
<td>0.940</td>
<td>0.613</td>
<td>0.949</td>
<td>0.898</td>
</tr>
<tr>
<td>OCD</td>
<td>0.934</td>
<td>0.585</td>
<td>0.966</td>
<td>0.915</td>
</tr>
<tr>
<td>IS</td>
<td>0.931</td>
<td>0.601</td>
<td>0.970</td>
<td>0.936</td>
</tr>
<tr>
<td>DEP</td>
<td>0.941</td>
<td>0.618</td>
<td>0.966</td>
<td>0.927</td>
</tr>
<tr>
<td>ANX</td>
<td>0.955</td>
<td>0.679</td>
<td>0.994</td>
<td>0.952</td>
</tr>
<tr>
<td>AGG</td>
<td>0.904</td>
<td>0.611</td>
<td>0.899</td>
<td>0.845</td>
</tr>
<tr>
<td>PHOB</td>
<td>0.915</td>
<td>0.606</td>
<td>0.994</td>
<td>0.936</td>
</tr>
<tr>
<td>PAR</td>
<td>0.912</td>
<td>0.634</td>
<td>0.955</td>
<td>0.897</td>
</tr>
<tr>
<td>PSY</td>
<td>0.938</td>
<td>0.606</td>
<td>0.980</td>
<td>0.929</td>
</tr>
</tbody>
</table>

Composite Reliability (CR), Average Variance Extracted (AVE), Maximum Shared Squared Variance (MSV), and Average Shared Squared Variance (ASV). Sel_71 and Sel_60 had a loading factor >0.5 and were excluded from the analysis.
Table 4: Construct correlation matrix

<table>
<thead>
<tr>
<th>Dimension</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOM</td>
<td>0.783</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCD</td>
<td>0.943</td>
<td>0.765</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IS</td>
<td>0.970</td>
<td>0.955</td>
<td>0.775</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEP</td>
<td>0.956</td>
<td>0.983</td>
<td>0.977</td>
<td>0.786</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANX</td>
<td>0.974</td>
<td>0.983</td>
<td>0.978</td>
<td>0.980</td>
<td>0.824</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGG</td>
<td>0.909</td>
<td>0.945</td>
<td>0.931</td>
<td>0.929</td>
<td>0.948</td>
<td>0.782</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHOB</td>
<td>0.964</td>
<td>0.954</td>
<td>0.982</td>
<td>0.967</td>
<td>0.997</td>
<td>0.920</td>
<td>0.778</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAR</td>
<td>0.928</td>
<td>0.929</td>
<td>0.975</td>
<td>0.955</td>
<td>0.970</td>
<td>0.892</td>
<td>0.977</td>
<td>0.796</td>
<td></td>
</tr>
<tr>
<td>PSY</td>
<td>0.955</td>
<td>0.963</td>
<td>0.985</td>
<td>0.959</td>
<td>0.990</td>
<td>0.886</td>
<td>0.987</td>
<td>0.969</td>
<td>0.778</td>
</tr>
</tbody>
</table>

Note: Correlations are below the diagonal, and AVE is presented on the diagonal, in bold.

On the basis of confirmatory factor analysis, all dimensions of the Malaysian version of GHQ-28 and SCL-90-R scales had sufficient internal consistency (Table 1 and 3) which implies a sufficient reliability and convergent validity for psychometric testing. In the other words, the results generally pointed that almost all items of the scales were significantly correlated with the factors (symptoms) of the original versions. Therefore, the finding of the study replicated those confirmed the convergent validity of the GHQ-28 (19-22, 41), as well as the SCL-90-R (19, 23-27).

On the basis of confirmatory factor analysis, the index value for all constructs has indicated no concerns about the construct validity of the GHQ-28 and support the four-factor structure of the scale. Putting simple terms, the sample of study could recognize inter-correlation between items or recognize the similarities among items; likewise, they could distinguish the differences between items, suggesting an adequate discriminant validity. Accordingly, the finding of the study is consistent with previous studies supporting construct validity and reliability of this scale (19-22, 41); however, it is inconsistent with another study (17).

In contrast, this study failed to support the construct validity of the SCL-90-R, in fact discriminant validity was not sufficient. Technically, AVE for all dimensions was less than MSV and ASV, in addition, squared root of AVE is less than the correlation between two constructs (Table 3). The SCL-90-R subscales are originally proposed to represent a nine-factor structure; however, the confirmatory factor analysis failed to support this fact; therefore, the study failed to support the multidimensionality of the scale with a stable multi-factor structure. It may be assumed that this assessment tool is more useful as a unitary measure for global distress instead of multidimensional measure (32). Nonetheless, the study supported the previous studies suggesting the SCL-90-R as a unitary measure rather than multidimensional structure (19, 23, 30-34), and was inconsistent with those studies replicated the multidimensionality of the original version of this scale (19, 23-28).

However, the current finding lends further support of the limits of the SCL-90-R as reported elsewhere (19, 25, 32). One way may account for is, the scores of the SCL-90-R are subjected to response biases, and severity and acuteness of mental crisis is offered as an explanation. Putting simple terms, some respondents believe that reporting their severe psychopathology such as paranoid status, hostility, and psychotic symptoms are undesirable and may lead to view of the involuntary hospitalization or other probable adverse consequences such as social and working problems. Therefore, a positive response bias, that is, denying and underreporting with the aim of preventing the probable adverse consequences would lessen the variability of the scores leading to incline toward favor of a general psychopathology factor. Another reason could be due to
the relationships between the sub-scales of the SCL-90-R. Some of sub-scales represented in the SCL-90-R are correlated with each other. For example, interpersonal sensitivity is associated with anxiety (42), and also depression is associated with somatization (43). Finally, the factor structure of the SCL-90-R may vary throughout the time of observation in different samples (32). Yet, these factors need to be more interpreted in light of the future studies.

Conclusion

On the bias of confirmatory factor analysis, the construct validity of the Malaysian version of GHQ-28 has been confirmed in the current study; however, this study failed to do so with the SCL-90-R. Since the SCL-90-R nine subscales are originally proposed as representing a multi-dimensional structure, this study failed to replicate this fact, and suggested a unitary measure for global distress instead of multi-dimensional measure; thus, it can be argued that multidimensionality of the SCL-90-R as clinical purposes is questionable. Therefore, it is may be a better unitary measure for assessing psychological status, general distress, and screening for mental disorders, as well as measuring change in outcome studies. Further research need to be carried out to prove this finding. This study might be considered unique, for, no study has focused on construct validity of the Malaysian version of the GHQ-28 and SCL-90-R across three different samples simultaneously; therefore, it may be contributed to add knowledge to the utility of two widely used psychological assessment tools.

Ethical considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Acknowledgements

The authors wish to thank the staffs of the UM Hospital and Hospital KL, and would like to appreciate Dr. Khafitz as well as the University Malaysia Center of Addiction Sciences (UMCAS) team for their collaboration with the study. The authors declare that there was no conflict of interest, and the study has not received any financial support.

References


Available at: http://ijph.tums.ac.ir


