ASSESSMENT AND DETECTION

A Psychometric Validation of the Short Alcohol Withdrawal Scale (SAWS)

Bjarne Elholm¹, Klaus Larsen², Nete Hornmes², Finn Ziera¹ and Ulrik Becker³,⁴,*

¹Alcohol Unit, Copenhagen University Hospital, Hvidovre, Capital Region, Denmark, ²Clinical Research Centre, Copenhagen University Hospital, Hvidovre, Capital Region, Denmark, ³Department of Medical Gastroenterology, Copenhagen University Hospital, Hvidovre, Capital Region, Denmark and ⁴National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark

*Corresponding author: Department of Medical Gastroenterology, 439, Hvidovre Hospital, DK-2650 Hvidovre, Denmark. Tel: +45-36-32-22-89; E-mail: becker@dahlnet.dk

(Received 29 December 2009; accepted 27 May 2010)

Abstract — Aims: The study aimed to evaluate psychometrically a Danish translation of the Short Alcohol Withdrawal Scale (SAWS) in an outpatient setting with Alcohol Dependence (AD) and Alcohol Withdrawal Symptoms/Syndrome (AWS). Methods: One hundred and twenty-two patients with AD and AWS filled in a 10-item rating scale to describe their symptoms with four graduations on five physical and five psychological items. The question of dimensionality of the construct was addressed in three different ways. First, a scree plot was constructed based on the polychoric correlations between items. Second, promax factor loadings were calculated for a two-factor model. These two steps were based on exploratory factor analysis. Third, specific violations such as local dependence and differential item functioning were investigated under the one-factor model in a confirmatory factor analysis. Results: The scree plot supported one or two dimensions while the promax rotations gave little support for a two-factor model. The confirmatory analysis also supported a one-factor model. Conclusion: The decomposition of the polychoric correlation matrix into eigenvalues and vectors suggested that there was most likely one factor underlying the 10 items in the SAWS. This was confirmed by a confirmative factor analysis with only one component when specific model violations such as local dependence and differential item findings were investigated. The SAWS is easy to use.

INTRODUCTION

Previously Alcohol Withdrawal Syndrome (AWS) was one of the cardinal diagnostic criteria for Alcohol Dependence (AD), and in the present version of the international disease classification the diagnosis of AWS is one of the six diagnostic criteria. AWS is one of the most important barriers against seeking alcohol treatment (Gossop et al., 2002). Adequate treatment of AWS is therefore imperative for the patients’ motivation for seeking treatment in inpatient as well as outpatient treatment programmes. Furthermore, sufficient treatment of AWS will reduce the risk of relapse and severity of future AWS (Becker, 1998; Malcolm et al., 2000). A number of scales have been developed to monitor the AWS in inpatient settings for example the Windsor Clinic Alcohol Withdrawal Assessment Scale (Metcalfe et al., 1995), the revised 10 item Clinical Institute Withdrawal Assessment (CIWA-Ar) Scale (Nuss et al., 2004; Sullivan et al., 1989), the revised 8-item Clinical Institute Withdrawal Assessment (CIWA-AD) (Reoux et al., 2006), the Alcohol Withdrawal Symptom Checklist (Pittman et al., 2007) and the AWS scale (Wetterling et al., 1997). Only the Short Alcohol Withdrawal Scale (SAWS) has been used and validated in outpatients (Gossop et al., 2002). Outpatient treatment for AWS is widely used but no consensus exists regarding a standardized monitoring of AWS (Saiz et al., 1997). Benzodiazepines are the drugs of choice for the treatment of AWS (National Board of Health, 2006), and at the Alcohol Unit, Hvidovre Hospital chloralhydrate has been used for decades administered according to a fixed dosage schedule — tapering the dose to zero over 8–10 days. No monitoring or systematic documentation of symptoms has been used unless the patient complained of continuing clinical symptoms in spite of treatment compliance. In that case, the dosage of AWS medication would be changed. A fixed dosage scheme suffers from the lack of individualized treatment, lack of monitoring and documentation of symptoms and a paternalistic view hampering the patients’ motivation for continued adherence to treatment. Therefore, it is important to implement a monitoring instrument in the treatment of AWS in outpatient settings. The aim of the present study was to validate the SAWS scale (Gossop et al., 2002) in the context of a randomized controlled trial comparing fixed dosage medication with symptom triggered in outpatients with AD and AWS.

MATERIALS AND METHODS

Participants

Consecutive outpatients suffering from AD and AWS were included in the study and randomized to a fixed-schedule treatment or a symptom-triggered treatment with chlordiazepoxide. All patients were evaluated by means of the European Addiction Severity Index. Patients fulfilling International Classification of Diseases 10th revision (ICD-10) diagnostic criteria for AD and diagnostic criteria for AWS were included if abstinence from alcohol had lasted for >72 h prior to inclusion and they had fluent Danish. Patients were excluded from the study if they were allergic to chlordiazepoxide, used drugs with known interaction with chlordiazepoxide, if they had been treated for AWS within the last week or had a history of three or more attempts of outpatient detoxification within the last month. Furthermore, patients were excluded if they had known severe psychiatric illness including Wernicke/Korsakoff syndrome, suicidal behaviour, severe cardiac or liver disease and type 1 diabetes. Pregnant or breastfeeding women and fertile women without safe contraception were excluded.

Finally, patients were excluded if breath alcohol concentration was >10 mg%.

Design

The patients were randomized into two groups: fixed-schedule or symptom-triggered medication. Randomization was performed as block randomization by a statistician (KL) at The
Research Centre at Hvidovre Hospital using the Randomization System-Copenhagen Trial Unit 1.04. The randomization was stratified according to SAWS score at baseline (score \(<12\) or score \(\geq 12\)). Sealed envelopes were delivered to the nurse by a secretary instructed to hand out the envelope with the lowest number from the block indicated by the nurse. Patients attended the outpatient clinic daily preferably for 10 days. In the fixed-schedule group, 200 mg chlordiazepoxide was prescribed as starting dose with daily tapering of the dose with 25 mg for patients with SAWS score \(\geq 12\) at baseline. For patients scoring SAWS \(<12\) at baseline, the starting dose of chlordiazepoxide was 80 mg with daily tapering of the dose with 10 mg. Patients were instructed to take the medication as prescribed and were offered an extra dose if necessary.

In the symptom-triggered group, patients scoring \(\geq 12\) at baseline were prescribed a maximum daily dose of chlordiazepoxide of 300 mg for 10 days. For patients scoring SAWS \(<12\) at baseline, the maximum daily dose was 120 mg for 10 days. These patients also had the possibility of taking extra doses if necessary.

All patients were instructed to bring back any unused tablets. To support abstinence, patients in both arms were offered concurrent treatment with disulfiram and/or acamprosate and controlled by breath alcohol levels.

SAWS

All patients filled in a rating scale (SAWS) with 10 items (Table 1) — five physical and five psychological items. Each item was scored 0 to 3 points retrospectively for the last 24 h, where 0 denotes no symptoms and 3 denotes severe symptoms. Patients were instructed to fill in the SAWS every day, preferably when they woke up. All patients filled out the SAWS scale until they terminated treatment or 10 consecutive days after enrolment giving a total of 100 observations per patient. The SAWS scores were transferred to the medical record by the nurse and time to SAWS \(\leq 12\). Treatment should be continued for at least 5 days, and the cut-off value of 12 was chosen after monitoring of AWS symptom scores in a pilot study of 18 patients.

**Statistical methods**

The framework for the analysis is a factor analytical model (Bollen, 1989). All measurements in all patients irrespective of treatment group were included in the analysis. As the items are ordered categorically, a threshold model is used assuming an underlying continuous and normal distributed variable (Muthén, 1984) for each of the items. All analyses are carried out for all time points separately. The question of dimensionality of the construct is addressed in three different ways. First, a scree plot (Cattell, 1966) is shown based on the polychoric correlations between items. A scree plot is the plotting of the eigenvalues of the correlation matrix in decreasing order. Second, promax factor loadings are shown for the two-factor model. These two steps are based on exploratory factor analysis. Third, specific violations such as local dependence (Chen et al., 1997) and differential item functioning (item biases) (Holland et al., 1993) are investigated under the one-factor model in a confirmatory factor analysis. All estimates are maximum likelihood estimates and all tests are likelihood ratio tests. Analysis was carried out in Mplus, version 4 (Muthén et al., 2006) and R, version 2.4.0 (R core development team, 2009).

**Ethical approval**

All participating patients gave informed consent according to the Helsinki Declaration, and the project was approved by The Danish National Committee on Biomedical Research Ethics (ref. no. 01-063/03), The Danish Medicines Agency (ref. no. 2612-2264) and Danish Data Protection Agency (ref. no. 2003-41-2937).

**RESULTS**

One hundred and fifty-three consecutive outpatients were randomized. Most of the 31 patients dropped out because of relapse before day 10. Thus, 122 patients participated in the study (99 men and 23 women (symptom triggered \(n = 56\), fixed schedule \(n = 66\)) (Table 2). All patients were abstinent for 10 days (receiving disulfiram or documented by breath alcohol test).

Table 3 shows the mean severity score for each of the 10 items in the SAWS scale from day 1 to day 10. Initially, restlessness, feeling miserable and sleep disturbance were those items with highest severity scores.

As seen from Fig. 1, there was clearly one dominating factor. Using the Kaiser-Guttman rule (to count the number of eigenvalues larger than one), it seems reasonable to conclude that there were at most two. The second largest eigenvalue is 1.35, 1.13, 0.99, 0.95, 1.06, 1.12, 1.07, 0.99, 1.09 and 1.01 for the 10 days, respectively, yielding no decisive conclusion of whether there are one or two underlying dimensions. Fig. 2 shows promax rotations of factor loadings from the exploratory two-factor model for days 1, 4, 7 and 10. Days 2, 3, 5, 6,
Table 3. Mean severity score for each item of the SAWS score from day 1 to day 10. For the sum scores, standard deviations are shown in parentheses

<table>
<thead>
<tr>
<th>Item</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>Day 6</th>
<th>Day 7</th>
<th>Day 8</th>
<th>Day 9</th>
<th>Day 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Anxious</td>
<td>1.0</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>2. Feeling confused</td>
<td>0.8</td>
<td>0.9</td>
<td>0.8</td>
<td>0.7</td>
<td>0.8</td>
<td>0.7</td>
<td>0.7</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
</tr>
<tr>
<td>3. Restless</td>
<td>2.2</td>
<td>1.8</td>
<td>1.6</td>
<td>1.4</td>
<td>1.2</td>
<td>1.1</td>
<td>1.0</td>
<td>0.9</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>4. Miserable</td>
<td>2.0</td>
<td>1.4</td>
<td>1.2</td>
<td>1.1</td>
<td>0.9</td>
<td>0.8</td>
<td>0.6</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
</tr>
<tr>
<td>5. Problems with memory</td>
<td>0.9</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>6. Tremor (shakes)</td>
<td>1.8</td>
<td>1.3</td>
<td>1.0</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>7. Nausea</td>
<td>0.7</td>
<td>0.6</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>8. Heart pounding</td>
<td>1.1</td>
<td>0.7</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>9. Sleep disturbance</td>
<td>2.0</td>
<td>1.3</td>
<td>1.0</td>
<td>0.9</td>
<td>0.8</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>10. Sweating</td>
<td>1.9</td>
<td>1.4</td>
<td>1.0</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Sum scores 14.9 (5.8) 11.1 (6.0) 9.1 (5.8) 7.7 (5.8) 6.8 (5.7) 6.3 (5.8) 6.3 (5.1) 4.6 (4.9) 4.0 (4.6) 3.6 (4.6)

8 and 9 days looked similar. Although it may seem that factor loadings change over time, this is not the main point illustrated by the plots. Their most interesting feature is the lack of evidence for a two-factor model.

The one-factor model was investigated by testing against local dependence and differential item functioning with respect to gender and age. There are 45 correlations between pairs of items at each of the 10 days. Table 4 shows results from the analysis of local dependence in Bonferroni-corrected $P$-values for the largest modification index at each of the 10 days. Although significant local dependence was found at days 5, 6 and 8, it is not immediately clear whether this is due to multiple testing. However, investigating the consistency of the findings over the 10 days sheds light on this issue; Table 4 also shows which pairs have modification indices above 3.84 (95th percentile in the chi-square distribution with one degree of freedom). Even though there is no distinct clustering of the items in two groups, some pairs of items are consistently closer correlated than others. Finally, the one-factor model was fitted with additional residual correlation between those two pairs of items, and — as to be expected from Table 4 — this model provided an excellent fit to the data at all 10 days. No differential item functioning was found with respect to gender and age.

**DISCUSSION**

Rating scales for AWS are helpful in prescribing the right dose of medication, monitoring symptoms in order to identify patients in risk of developing complications and for documenting AWS treatment. Furthermore, the use of symptom-triggered treatment schedules has been shown to be of benefit (McKay *et al.*, 2004). However, few studies have validated rating scales for AWS (Williams *et al.*, 2001).

Instead of a factor analysis, we could have used a Rasch model analysis but that model does not have the same constructivity on dimensionality as the factor analysis. The scree plot suggested either one or two subscales in the SAWS questionnaire, and the promax factor loadings from the two-factor model were not indicative of there being two factors. A few violations of the one-factor model were found when investigating the local dependence between the items and no differential item functioning was detected. We did not confirm the finding of a two-component scale (Gossop *et al.*, 2002).

Were the 10 items driven by two latent variables, of which some of the items load on one and the rest of the items load on the other, then the factor loadings would cluster in those two separate groups accordingly? As there is no such separation of the items into two groups, the distribution of the factor loadings does not support a two-factor model. This being said, there seems to be a slight though systematic change in the loadings over the time so that if you were to group the items into two, the grouping suggested by Gossop *et al.* (2002) seems reasonable at the end of the observation period (days 7–10; days 8 and 9 not shown). However, this grouping does not seem reasonable at the first 6 days of the observation period (only days 1 and 4 shown).

In particular, items 2 and 5 turn out to have a significant residual correlation in 9 out of 10 days, while the correlation between items 6 and 10 is significant in 7 out of 10 days. It is interesting that item 2 ‘feeling confused’ and item 5 ‘problems with memory’ were strongly correlated because these items represent symptoms of the AWS as well as side effects to the use of benzodiazepines. Secondly, item 6 ‘tremor/shakes’ and item 10 ‘sweating’ were correlated probably because symptoms reflect reactions from the autonomic nervous systems.

We found all 10 items to be relevant and well understood by clinicians and patients as well as the 24 hourly scoring was
well understood. Furthermore, inclusion of both somatic and psychic items is clinically relevant and easy to use in daily clinical practice. There is no reason why SAWS cannot be used at a 'present state' measure or scoring with shorter intervals or shorter monitoring periods as suggested by Gossop et al. (2002). To our knowledge, a confirmatory study of SAWS has not been done before over a period of 10 days. Gossop et al. (2002) used a 1–3 days monitoring period. Measuring and monitoring AWS in an outpatient setting may assure sufficient treatment to patients with AWS and at the same time making it possible to monitor symptoms over time and document the treatment. Because of the frequent measuring and treatment according to the scoring, it may slow down the kindling effect on the brain according to other studies (Becker, 1998). Many studies have given the patients' involvement in their own description of AWS symptoms a less important role. Object rating for AWS is in many rating scales the choice, though it seems that patients are suffering due to lack of scoring frequencies and new staff. Some AWS symptoms are difficult to score and describe from object-rating scales (i.e. feeling restless). Making the patients responsible for their own AWS scoring as in self-rating and thereby responsible for their own treatment will be a major challenge for the staff and their view on the patients suffering from AWS.

In conclusion, the decomposition of the polychoric correlation matrix into eigenvalues and vectors suggested that there was most likely one factor underlying the 10 items. This was confirmed by a confirmative factor analysis with only one component when specific model violations such as local dependence and differential item findings were investigated. The

<table>
<thead>
<tr>
<th>Day</th>
<th>Pairs of items with modification index (MI) &gt;3.84</th>
<th>Max. MI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1:2 2:5 5:7</td>
<td>10.098</td>
<td>0.067</td>
</tr>
<tr>
<td>2</td>
<td>2:5 3:5b 6:7b 6:9b</td>
<td>8.846</td>
<td>0.132</td>
</tr>
<tr>
<td>3</td>
<td>2:5 3:5b</td>
<td>8.425</td>
<td>0.167</td>
</tr>
<tr>
<td>4</td>
<td>2:5</td>
<td>8.948</td>
<td>0.125</td>
</tr>
<tr>
<td>5</td>
<td>2:5 6:7b</td>
<td>12.464</td>
<td>0.019</td>
</tr>
<tr>
<td>6</td>
<td>2:5 3:4</td>
<td>14.600</td>
<td>0.006</td>
</tr>
<tr>
<td>7</td>
<td>2:5 2:10b</td>
<td>14.000</td>
<td>0.006</td>
</tr>
<tr>
<td>8</td>
<td>2:5</td>
<td>10.856</td>
<td>0.044</td>
</tr>
<tr>
<td>9</td>
<td>2:5</td>
<td>9.906</td>
<td>0.074</td>
</tr>
<tr>
<td>10</td>
<td>2:5</td>
<td>4.837</td>
<td>1.000</td>
</tr>
</tbody>
</table>

*The 95th percentile is 3.84 in the chi-square distribution with one degree of freedom.
*The residual correlation is negative.
*Bonferroni-corrected P-values for each day.
Pairs that are significant on 5% level are shown, as well as test statistics (Max MI) and P-values for test of local dependence corrected for multiple testing.
SAWS is easy to use. Further studies should focus on the use of SAWS for individual prescription of drugs for the treatment of AWS.

Acknowledgements — The authors wish to thank the patients, the nurses and the staff at the Alcohol Unit and Clinical Research Center at Hvidovre Hospital for their participation and help with this project. We also wish to thank Janne Petersen and Yoshu Suzuki and the secretary, Bettina Rasmussen for all their help.

Conflict of interest statement: None declared.

REFERENCES