Psychometric properties of the symptom check-list-90-R in prison inmates

Yuriy Ignatyev, Rosemarie Fritsch, Stefan Priebe, Adrian P. Mundt

Abstract

The aim of this study was to investigate the reliability, construct and criterion validity of the Symptom Check-List-90-R (SCL-90-R) for prison inmates. A sample of 427 adult prisoners was assessed at admission to the penal justice system in the metropolitan region of Santiago de Chile using the SCL-90-R and the mini international neuropsychiatric interview. We tested internal consistency using Cronbach’s alpha. We examined construct validity using Principal Components Analysis and Confirmatory Factor Analysis (PCA and CFA) as well as Mokken Scale Analysis. Receiver Operating Characteristic (ROC) analysis was conducted to examine external criterion validity against diagnoses established using structured clinical interviews. The SCL-90-R showed good internal consistency for all subscales (α = 0.76–0.89) and excellent consistency for the global scale (α = 0.97). PCA yielded a 1-factor structure, which accounted for 70.7% of the total variance. CFA and MSA confirmed the unidimensional structure. ROC analysis indicated useful accuracy of the SCL-90-R to screen for severe mental disorders. Optimal cut-off on the Global Severity Index between severe mental disorders and not having any severe mental disorder was 1.42. In conclusion, the SCL-90-R is a reliable and valid instrument, which may be useful to screen for severe mental disorders at admission to the prison system.

1. Introduction

Several studies have pointed to high prevalence rates of mental health problems in prison populations worldwide (Lamb and Weinberger, 2001; Fazel and Danesh, 2002; Mir et al., 2015) and in South America (Ponde et al., 2011; Mundt et al., 2013, 2015b; Andreoli et al., 2014). Decreasing psychiatric bed numbers were linked with increasing prison population rates in South America (Mundt et al., 2015a). Prisoners with mental illness are at risk to become victims of other inmates and at risk of suicide (Fazel et al., 2011). The detection of mental disorders in prisoners at admission is essential to initiate adequate treatment and protection. Furthermore, it can contribute to the wellbeing of other prisoners, correctional staff, and the community (Martin et al., 2013).

Most of the screening tools to detect mental health problems in prison inmates were developed especially for correctional facilities without being used or validated in the general populations (Martin et al., 2013). However, instruments that could be used for the same people while imprisoned and while living in the community may have advantages for longitudinal studies and clinically. People with mental disorders living in the community have high rates of incarcerations. Severe mental disorders have shown to predict incarceration in the general population (Greenberg and Rosenheck, 2008). The integration of treatments in correctional institutions and in the community is often still poor. The use of the same screening instrument, which is valid both in community and in prison service, could improve this.

The Symptom Checklist-90-Revised (SCL-90-R) has been translated in 24 languages and is validated in different communities. It may be useful for screening purposes (Derogatis, 1994). The tool was designed to evaluate a broad range of psychological problems and symptoms. It has also been used to measure the outcome of severe mental disorders in clinical or research contexts (Burlingame et al., 2005). Descriptive statistics such as mean values of psychopathological subscales and indices of the SCL-90 as well as the SCL-90-R...
among prisoners have been already established in the UK (Wilson et al., 1985), the USA (Gibbs, 1987; Steadman et al., 1999; Harris et al., 2003; Moser et al., 2004), Germany (von Schönfeld et al., 2006; Dudeck et al., 2009; Obschonka et al., 2010), Spain (Echeburua et al., 2003; Villagra Lanza et al., 2011), and Iran (Sepehrmanesh et al., 2014). Several prison studies examined the criterion validity of the tool with respect to clinical (Wilson et al., 1985; Bulten et al., 2009) and legal (Harris et al., 2003; Taylor et al., 2010) attributes. The criterion validity of the SCL-90-R among severely violent psychiatric inpatients was established in Norway (Bjerkly, 2002). The convergent validity of the SCL-90-R was examined in a British prison sample (Wilson et al., 1985).

Studies assessing the factor structure of the SCL-90-R in samples of prison inmates are lacking. In a review on psychometric properties of the SCL-90-R it was concluded that the dimensionality may vary across different diagnostic and social groups (Cyr et al., 1985). A study of people in crisis with high levels of suicide risk and aggressive behaviors showed that a one-factor model (global psychological distress) may best represent the data (Bonyne, 1993).

The aims of the study were to investigate the reliability of the SCL-90–R, to examine whether the unidimensional structure of the instrument applies to the prison context, to validate the instrument against a structured diagnostic interview in prisoners, and to suggest a threshold score for detecting severe mental disorders in prison inmates.

2. Methods

2.1. Setting and design

We conducted a cross-sectional observational study of consecutively committed prison populations. The sample of 229 male and 198 female prisoners at admission were randomly selected from lists in the three remand prison facilities serving the metropolitan region of Santiago de Chile. The field team consisted of three clinical psychologists trained and supervised by a senior consultant psychiatrist in using the instruments. The assessments including socio-demographic variables, the mini international neuropsychiatric interview (MINI) and application of the SCL-90-R lasted for 45–60 min and were held in separate rooms. In the case of difficulties answering any of the items on the SCL-90-R, participants had the opportunity to immediately consult and resolve this with an assessor. The data were collected between February and September 2013. All the females admitted in the study period were approached for inclusion; every third male on the daily printed admission lists were approached for inclusion. Exclusion criteria for the study were the inability to communicate in the Spanish language and a lack of capacity to provide informed consent. The study was approved by the Ethics Review Board of the University of Chile (Acta de Aprobación 01 from 25.01.2012) and by the Ministry of Justice of the Republic of Chile (reference: Subsecretaria de Justicia 15.03.2012). For details on the sampling and population also see Mundt et al. (2015b).

2.2. Instruments

2.2.1. Symptom checklist 90-Revised (SCL-90-R)

The questionnaire consists of 90 symptom statements that respondents rate on a five-point scale of severity based on their experience in the previous week. The nine subscales of the SCL-90-R are as follows: Somatization, Obsession compulsion, Interpersonal sensitivity, Depression, Anxiety, Hostility, Phobic anxiety, Paranoid ideation, and Psychoticism. There are seven additional items that explore disturbances in appetite and sleep. The Global Severity Index (GSI), the mean score on the instrument, is a widely used global index of distress.

2.2.2. Mini international neuropsychiatric interview (MINI)

Participants were assessed for the presence of psychiatric disorders using the Spanish version of the MINI 5.0 as ‘gold standard’. The MINI was developed by Sheehan et al. (1998) to classify mental disorders according to the fourth version of the DSM-IV. The tool covers a wide range of current and lifetime psychiatric diagnoses including current severe mental disorders, such as major depression, recurrent major depression, major depression with melancholic features, current manic episode, current psychotic disorder, and current psychotic mood disorder.

2.3. Data analysis

Socio-demographic and psychiatric characteristics of the sample were assessed using descriptive statistics. Internal consistency was explored calculating the inter-correlations between the original nine subscales and between the subscales and the GSI. The Cronbach’s α coefficient was established for the items and the subscales as well as for the subscales and the GSI. A Cronbach’s α between 0.6 and 0.7 is considered an acceptable value. A value between 0.7 and 0.9 is a good value, and a value of 0.9 or higher indicates excellent reliability (Fayers and Machin, 2007).

To examine the theoretical one-dimensional structure of the SCL-90-R regardless of multivariate normal assumption, first Principal Components Analysis (PCA) and then Confirmatory Factor Analysis (CFA) were conducted (Gerbing and Hamilton, 1996; Wang and Du, 2000). As suggested by Kline, the ratio between the sample size and the number of items in a questionnaire should approach 10:1 to indicate an optimal condition for factor analysis. Since our sample included only 427 cases, we examined the factor structure at the subscale-level (Kline, 2011). The sample was split in two subsamples stratified for gender and age. Ten cases were excluded due to missing data on age. To test for the adequacy of factor analysis for both subsamples, we used the Kaiser-Meyer-Olkin measure (should be ≥ 0.5) and the Bartlett test of sphericity. The Mardia’s tests of multivariate skewness and kurtosis were used to examine a data deviation from multinormality. A PCA of mean scores from the nine subscales was conducted in the first subsample. To determine the number of factors we used the Kaiser-Guttman eigenvalue ≥ 1 criterion and the Cattell’s scree plot. Further, we conducted a CFA with the maximum-likelihood solution in the second subsample. A global fit of our model was examined by different fit indices. The chi-square goodness-of-fit test should ideally be non-significant, or at least have evidence of a chi-square/df ratio between two and five (Tabachnik and Fidell, 2006). The Comparative Fit Index (CFI) should be above 0.95, and the Root Mean Square Error of Approximation (RMSEA) Index should ideally be around 0.05, and not higher than 0.10 for a good model-data fit (Blunch, 2008). The Tucker-Lewis-Index (TLI) should have values as low as 0.90. For the Standardized Root Mean-square Residual (SRMR), a cut-off as low as 0.08 has been suggested (Blunch, 2008). To test the model for hidden factors, the use of modification indices was applied. Mokken Scala Analysis (MSA) with the model of monotone homogeneity was used to subsequently examine the construct validity. The model of monotone homogeneity assumes unidimensionality, monotonicity, and local independence of items within a scale. The fit of the model was evaluated using the Loewinger’s scalability H coefficients. H between 0.3 and 0.39 indicates a weak scale, H between 0.4 and 0.49 indicates a moderate scale, and H of 0.5 or higher indicates a strong scale (Meijer and Baneke, 2004). We conducted the MSA using the command line of the mokken package in the R free software for the polytomous SCL item scoring from 0 to 4 (Van
A Receiver Operating Characteristic (ROC) curve analysis was conducted in order to determine SCL-90-R criterion validity and cut-off scores for severe mental disorders, which are the most relevant psychiatric disorders in correctional setting to screen for. To confirm criterion validity coefficients, we used the two previously generated random subsamples. Using the MINI as the ‘gold standard’ for severe mental disorders, areas under ROC curves (AUCs) of each subsample were calculated. AUCs provide an indication of a particular scale’s diagnostic ability to discriminate between those with and without diagnoses. Values between 0.50 and 0.70 represent a scale with low accuracy, values between 0.70 and 0.90 are indicative of a useful screening scale and a value of 0.90 and above indicates high accuracy for screening (Swets, 1988). To confirm the screening properties of the SCL-90-R, a comparison of the ROC-curves was conducted with the Venkatraman’s permutation test for unpaired ROC-curves (Venkatraman, 2000). The test is included in the R-package pROC (Robin et al., 2011). The operation was done with the method “venkatraman”, the permutation of sample ranks. The p-value of the test was computed using 2000 permutations, which was set by default. The command line in R for conducting the Venkatraman’s permutation test for unpaired ROC-curves is shown in the legend of Fig. 1. To confirm the cut-off score, we established in the first subsample sensitivity, specificity, positive predictive values (PPV), negative predictive values (NPV), diagnostic positive (DLR+) and negative (DLR−) likelihood ratios as well as the optimal cut-point based on sensitivity–specificity criterion (Hosmer Jr et al., 2013). The testing was carried out with the command line of the R-package OptimalCutpoints (López-Ratón et al., 2014). The cut-point was then used in the second subsample and appropriate validation coefficients were established. Their conformity with confidence intervals of the corresponding validation coefficients from the first subsample was tested. The visualization of receiver operating characteristic (ROC) graphs was conducted using the command line of the ROCR package in R (Sing et al., 2005). In addition, the proportions of positive (true vs. false) and negative (true vs. false) cases of both samples using Yates’ chi-square test were compared. Descriptive analysis, reliability tests and factor analysis were performed using statistical packages SPSS 21, AMOS 21. The MSA and the ROC-analysis were carried out with the R packages running in R version 3.2.0 (R Core Team, 2015).

3. Results

3.1. Socio-demographic and psychiatric characteristics of the sample

N=473 prisoners were identified on the admission lists as potential participants. Three did not follow the call to the assessment area and could not be screened for eligibility. Seven potential participants were excluded due to mental or psychological incapacities to participate. Out of the remaining 463, n=30 rejected participation; n=433 agreed to participate in the study; n=6 prematurely ended the assessment and were excluded from further analysis; the sample used for the final analyses included a total of n=427 participants (229 male and 198 female prisoners). Since questions could be resolved immediately and people were reminded if they accidentally omitted a question, we did not observe missing items of the SCL-90-R unless participants discontinued or interrupted the entire session and did not return. The mean age of all participants was 31.6 ± 11.5 years. A majority of 76% had low levels of education and 75% of participants worked for income before imprisonment. A majority of 59% was living without partner; 44% of all participants had previous imprisonment(s).

According to the MINI, suicidal risk was present in 208 (48.7%) of participants. At least one psychiatric disorder was present in 327 (76.6%) participants. The mean number of diagnoses was 2.95. Among those with disorders, n=76 (23.2%) had one disorder, n=86 (26.3%) had two disorders, and n=165 (50.5%) had three or more disorders. Any current severe mental disorder was present in n=222 (52%) prisoners.

3.2. Internal consistency

Table 1 shows the mean scores of the nine original subscales of the SCL-90-R. Cronbach’s α on the item-level showed good consistency for all subscales except for the subscale Paranoid ideation, which had a value of 0.70 at the limit between acceptable and good consistency if item 76 (Others do not give you proper credit for your achievements) was deleted. Cronbach’s α on the subscale-level presented in Table 2 indicated good internal consistency for all

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Total mean score (SD) N=427</th>
<th>Females mean score (SD) N=198</th>
<th>Males mean score (SD) N=229</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somatization</td>
<td>1.50 (1.54) 1.56 (1.58)</td>
<td>1.63 (1.66) 1.70 (1.68)</td>
<td>1.39 (1.42) 1.70 (1.68)</td>
</tr>
<tr>
<td>Obsession</td>
<td>1.58 (1.51) 1.44 (1.50)</td>
<td>1.44 (1.58) 1.70 (1.44)</td>
<td>1.70 (1.44) 1.70 (1.44)</td>
</tr>
<tr>
<td>Compulsion</td>
<td>1.23 (1.48) 1.17 (1.54)</td>
<td>1.17 (1.54) 1.29 (1.42)</td>
<td>1.39 (1.42) 1.29 (1.42)</td>
</tr>
<tr>
<td>Interpersonal sensitivity</td>
<td>1.79 (1.64) 1.58 (1.68)</td>
<td>1.84 (1.76) 1.75 (1.54)</td>
<td>1.75 (1.54) 1.75 (1.54)</td>
</tr>
<tr>
<td>Depression</td>
<td>1.64 (1.58) 1.70 (1.68)</td>
<td>1.70 (1.68) 1.60 (1.49)</td>
<td>1.60 (1.49) 1.60 (1.49)</td>
</tr>
<tr>
<td>Hostility</td>
<td>0.72 (1.23) 0.55 (1.13)</td>
<td>0.55 (1.13) 0.87 (1.32)</td>
<td>0.87 (1.32) 0.87 (1.32)</td>
</tr>
<tr>
<td>Phobic anxiety</td>
<td>0.68 (1.38) 0.90 (1.46)</td>
<td>0.90 (1.46) 0.87 (1.32)</td>
<td>0.87 (1.32) 0.87 (1.32)</td>
</tr>
<tr>
<td>Paranoid ideation</td>
<td>1.41 (1.55) 1.41 (1.66)</td>
<td>1.41 (1.66) 1.40 (1.45)</td>
<td>1.40 (1.45) 1.40 (1.45)</td>
</tr>
<tr>
<td>Psychoticism</td>
<td>1.11 (1.49) 0.96 (1.51)</td>
<td>0.96 (1.51) 1.22 (1.47)</td>
<td>1.22 (1.47) 1.22 (1.47)</td>
</tr>
<tr>
<td>GSI</td>
<td>1.41 (0.81) 1.39 (0.79)</td>
<td>1.39 (0.79) 1.42 (0.82)</td>
<td>1.42 (0.82) 1.42 (0.82)</td>
</tr>
</tbody>
</table>
Table 2: Correlations on the SCL-90-R subscales and GSI with Cronbach’s alpha (α) and Loewinger’s scalability coefficient (H).

<table>
<thead>
<tr>
<th>Subscale</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>Somatization</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obsession</td>
<td>0.68</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpersonal</td>
<td>0.63</td>
<td>0.76</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compulsion</td>
<td>0.73</td>
<td>0.80</td>
<td>0.76</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>0.78</td>
<td>0.75</td>
<td>0.76</td>
<td>0.82</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hostility</td>
<td>0.42</td>
<td>0.58</td>
<td>0.60</td>
<td>0.31</td>
<td>0.54</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychoticism</td>
<td>0.62</td>
<td>0.67</td>
<td>0.70</td>
<td>0.69</td>
<td>0.76</td>
<td>0.43</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paranoid ideation</td>
<td>0.62</td>
<td>0.71</td>
<td>0.76</td>
<td>0.71</td>
<td>0.75</td>
<td>0.59</td>
<td>0.65</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>GSI</td>
<td>0.63</td>
<td>0.74</td>
<td>0.72</td>
<td>0.72</td>
<td>0.78</td>
<td>0.61</td>
<td>0.63</td>
<td>0.75</td>
<td>1.00</td>
</tr>
<tr>
<td>Hostility</td>
<td>0.84</td>
<td>0.89</td>
<td>0.86</td>
<td>0.91</td>
<td>0.93</td>
<td>0.65</td>
<td>0.80</td>
<td>0.84</td>
<td>0.86</td>
</tr>
<tr>
<td>Somatization</td>
<td>0.87</td>
<td>0.85</td>
<td>0.81</td>
<td>0.88</td>
<td>0.89</td>
<td>0.82</td>
<td>0.76</td>
<td>0.76</td>
<td>0.82</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>0.39</td>
<td>0.38</td>
<td>0.38</td>
<td>0.42</td>
<td>0.49</td>
<td>0.55</td>
<td>0.37</td>
<td>0.37</td>
<td>0.41</td>
</tr>
</tbody>
</table>

first sample are summarized in Table 3. The optimal mean GSI cut-point score to screen for severe mental disorders was 1.42. At this threshold, the validation coefficients in the second subsample were as follows: sensitivity = 0.72, specificity = 0.78, PPV = 0.76, NPV = 0.78, DLR+ = 3.27 and DLR− = 0.36. All of these values were within confidence intervals of the validation coefficients of the second subsample. The comparison of proportions of true-false positive (chi-square = 0.410, df = 1, p = 0.52) and true-false negative (chi-square = 0.021, df = 1, p = 0.89) cases between two samples confirmed the appropriateness of the cut-point of 1.42 for the screening of severe mental disorders in prison inmates.

4. Discussion

4.1. Main findings

The SCL-90-R showed good reliability for each subscale and excellent reliability for the summary scale. Classical test theory and the non-parametric IRT approach suggest that the unidimensional structure of the SCL-90-R could not be rejected. ROC-analysis showed the validity of the SCL-90-R to screen for severe mental disorders. The optimal cut-point to screen for severe mental disorders covered by the MINI was established.

4.2. Strength and limitations

This is the first study testing unidimensionality of the SCL-90-R in prison inmates and evaluating criterion validity of the tool to screen for severe mental disorders in prisoners establishing a cut-point. Convergent validity of the instrument (overlap with a similar measure) was not tested. The one-factor structure of the SCL-90-R was confirmed on the subscale-level, not on the item-level because the sample was too small. Since CFA was conducted on the subscale-level, the unidimensional model may reflect a second-order factor. However, the MSA based on the item-level also showed unidimensionality. In our study, we examined only a simple one-factor model of global psychological distress. The presence of assessors while going through the questionnaire was different to the genuine self-report situation. Finally, this study had a cross-sectional design. Additional longitudinal studies are needed to establish psychometric properties of the SCL-90-R over time.

4.3. Comparison against the literature

More than half of prison inmates in our sample had severe mental disorders. This rate is higher than those reported for most other countries (Fazel and Seewald, 2012) and may be related to assessing the inmates unlike most other studies at the moment of admission to the penal justice system. The mean GSI score detected in our sample was higher than the ones reported from North American (Harris et al., 2003) and European (von Schönfeld et al., 2006; Dudeck et al., 2009; Echeburua et al., 2003) prison...
populations. This result could be explained by assessing at admission, which may be a moment of particular distress, and by the specific cultural context. It has been reported in clinical studies that self-administered questionnaires tend to overestimate psychopathology in samples from Latin America (Lewis and Araya, 1995). Sample characteristics may have accounted for a high Lövénier’s H coefficient on the subscale Hostility compared to other subscales. The level of Hostility in prisoners detected in our study was lower than in the general population in Chile (Fuentesalba et al., 2004). High internal consistency of this subscale may correspond to a common item response pattern reflecting social desirability of lower hostility in prisons (Vigil-Colet et al., 2012).

High inter-correlations between subscales, one-factor structure and sufficient Lövénier’s scalability H coefficients of the SCL-90-R in our study are consistent with previous validation studies, which found unidimensionality of the tool (Rief and Fichter, 1992; Bonnye, 1993; Schmitz et al., 2000; Elliott et al., 2006). The findings contrast recent Norwegian studies (Paap et al., 2011, 2012) in a large clinical sample, which demonstrate multidimensionality of the SCL-90-R in severely ill patients. The differences may be explained by characteristics of the populations. Our sample was non-clinical, had 52% of cases of severe mental disorders and high levels of psychiatric comorbidity. The unidimensionality of the SCL-90-R detected in our study may reflect a lower level of severe psychopathology than in clinical (i.e. inpatient) populations but higher levels of comorbidity. A recent study of the SCL-90-R in a large non-clinical community sample showed that a bifactor structural model had the best fit (Urbán et al., 2014). To clarify whether the bifactor or other hierarchical models apply to the prison context, further research with larger samples is necessary.

Since psychopathological symptoms in prison populations can rarely be ascribed to one single clearly delineated diagnostic category, a dimensional approach may be more appropriate to conceptualize and study psychopathology. Resources in prison mental health care are particularly limited. Self-report screening could be an efficient way to detect severe mental disorders in newly admitted prisoners. Due to the unidimensionality of the SCL-90-R, shorter versions of the instrument (i.e., Brief Symptom Inventory) may be explored in prison settings.

5. Conclusion

The SCL-90-R has good internal consistency, plausible construct validity, and useful diagnostic accuracy. The SCL-90-R could be a useful screening instrument for severe mental disorders at admission to prison.

Contributors

Concept and design: YI, APM. Critical revision of the manuscript and final approval: YI, RF, SP, APM.

Conflict of interests

None.

Acknowledgments

Adrian P. Mundt was a European Union, Marie Curie International Outgoing Fellow, Grant number PIOF-2011-302346. The authors would like to acknowledge Sinja Kastner, Sarah Larrain, Danilo Jimenez, Catalina Lira and Daphne Yevenes for contributions to the data collection.

References


Mundt, A.P., Castner, S., Larrain, S., Fritsch, R., Priebe, S., 2015b. Prevalence of mental disorders at admission to the penal justice system in emerging countries: a