Proxy screening tools improve the recognition of dementia in old-age homes: results of a validation study

LEONORE KÖHLER, SIGFRIED WEYERER, MARTINA SCHAUFELLE

Centre: Psychogeriatric Research Unit, Central Institute of Mental Health, Mannheim, Germany
Address correspondence to: Dr. Martina Schäufele. Tel: (+49) 621 1703 6352; Fax: (+49) 621 1205. Email: martina.schaeufele@zi-mannheim.de

Abstract

Background Despite the high true prevalence of dementia, demential disorders of residents of old-age homes may often be not recognised. There is a need for a standardised tool which includes observations of nursing staff.

Objective To describe and validate the Dementia Screening Scale (DSS) for use by nursing staff in old-age homes.

Methods All residents of 20 randomly selected old-age homes in the city of Mannheim, Germany (n = 1,922) were rated by nurses using the seven-item proxy dementia rating scale. Based on a subset of residents (n = 598) the DSS was validated against independent diagnostic assessments made by trained psychologists including the Mini-Mental-State-Examination (MMSE), the Dementia Scale of the Brief Assessment Schedule (BAS DEM), and the Washington University Clinical Dementia Rating (CDR).

Results Using the CDR as a gold standard, the DSS correctly classified at a cut-off of 2/3, 85.8% of the mildly, moderately, or severely demented residents. The accuracy of the DSS was only a little worse than that of the MMSE and the BAS DEM.

Conclusion The DSS is well suited for the recognition of dementia in old-age homes. It achieved a better validity than global diagnosis-related staff assessments and compared to performance-based instruments. It is easier to apply, more economic, and associated with a lower rate of non-response.

Keywords: screening, dementia, validation study, nursing homes, proxy, elderly

Introduction

There is no doubt that in the industrialised countries the prevalence of dementia among residents of old-age homes is high. Studies carried out in Scandinavia, North America and the UK since the beginning of the 1980s revealed that 17–36% of the residents of residential homes, and 56–72% of the residents of nursing homes suffered from demential disorders [1–4].

Despite the high true prevalence of dementia, and despite the fact that patients of dementia need intense medical attention and treatment as well as specialised forms of care, demential disorders in nursing homes may not often be recognised [5].

A large survey which we have conducted recently in 11 nursing homes in the South of Germany indicated that only 27% of the demented residents were formally acknowledged to have dementia [6]. A lack of perception and adequate diagnostic work-up of demential disorders has also been reported from other European countries, as well as Canada [7–10]. Since recognition of dementia is the first stage in the provision of adequate treatment and care planning, there is an urgent need to remedy this situation [7].

One approach might be to introduce the systematic use of screening instruments in the nursing homes [11]. However, most screening instruments designed to assess dementia, such as the widespread Mini Mental State Examination (MMSE), are based on cognitive testing. The administration of a cognitive test poses requirements...
which are often not met in old-age homes: These are, on the one hand, skilled personnel and time, and on the other hand, maintained sensory and motor functions, as well as verbal comprehension skills on the part of the resident [12, 13]. To overcome these and other disadvantages of traditional assessment methods, proxy instruments have been developed [14–17]. A meta-analysis by Neri et al. [18] suggested that on average, an informant questionnaire is as effective as a screening tool for the detection of dementia as is a brief cognitive [19, 20] or a neuropsychological test [21].

However, comparable brief instruments for use by professional caregivers in old-age homes, such as the instruments included in the Multidimensional Observation Scale for Elderly Subjects [22] or in the Psychogeriatric Dependency Rating Scales [23], are rare. Generalised and diagnosis-related staff ratings (i.e. determining the type of psychiatric disorder, the presence or absence of dementia) have led to inconsistent results, and frequently, to low rates of correct classifications: A Danish nursing home study revealed a sensitivity of global staff ratings of 68%; the specificity was 84% [9]. In southeast England, the recognition was even worse. The senior nurses identified only 34.4% of the residents with an MMSE score ≤ 23 as having dementia [7]. In a Canadian study, the nurses correctly assigned the level of dementia in only 37% of the cases [10]. Therefore, the development of a proxy screening scale tailored for use by professional caregivers seems to be a promising way to improve the recognition rate for dementia in old-age homes.

Aims of the study
We constructed such a scale, the Dementia Screening Scale (DSS), by taking into account criteria described by Magaziner [24] which have turned out to improve the quality of proxy ratings. One reason for this work was the growing need for simple economic screening instruments for dementia in long-term care expressed by the care providers. In the following, we will describe the DSS for use by professional caregivers, and determine its validity against an external criterion and against approved cognitive screening tests.

Methods
Description of the DSS
The DSS was constructed as a dementia screening scale for use by professional caregivers. It comprises a series of seven items and includes two domains of cognitive functioning: memory and orientation (see Discussion). The DSS total score varies between 0 and 14 (the higher the score the more severe the cognitive impairment). In general, the caregivers need 1–3 min to complete the DSS for one person.

Validation study
The present study was part of a larger epidemiological survey on the use of psychotropic drugs in old-age homes, which was carried out in the German city of Mannheim (population: 308,000). From the total of 26 homes in Mannheim, a random sample of 20 institutions was drawn (n = 1,922; residential homes n = 299, nursing homes n = 1,623; mean age = 81.1 years; women = 77.1%). The residents were comprehensively assessed by staff members. Our requirements for the assessors were: frequent contact with the resident in the last 4 weeks, and being licenced (please see Appendix 1 in the supplementary data on the Journal website http://www.ageing.oxfordjournals.org).

The validation study was based on a subset of the total sample (please see Appendix 2 in the supplementary data on the Journal website). For this subset, a diagnostic assessment was carried out by trained clinical psychologists independently of the nurses’ assessment. The diagnostic assessment was based on instruments with established psychometric properties: the MMSE [25], the Brief Assessment Schedule (BAS) [26], and the Washington University Clinical Dementia Rating (CDR) [27]. The global diagnostic judgement according to CDR was used as the gold standard because it is a widely accepted standard, and considered to be reliable and valid [28].

Statistics
The validity of the DSS was established by determining sensitivity, specificity, positive and negative predictive values with variable cut-off points, and the correlation with two performance-based screening instruments: The Dementia Scale of the BAS (BAS DEM) and the MMSE.

For statistical analyses, subgroups of residents were formed according to the CDR: (i) a mild+− group (‘mild’, ‘moderate’ or ‘severe’ dementia versus ‘no’ or ‘questionable’ dementia) and (ii) moderate+− group (‘moderate’ or ‘severe’ dementia versus ‘no’, ‘questionable’, or ‘mild’ dementia). The coefficients for further subgroups (i.e. according to the type of old-age home or age groups) were also calculated.

Moreover, the areas under the ROC curve (AUC) for the DSS were calculated non-parametrically and compared with those of the MMSE and the BAS DEM; for the comparisons z-scores were calculated according to Hanley and McNeil [29].

All analyses were conducted using SPSS version 12.0 (SPSS Inc, Chicago, IL, USA).

Results
Excluding the cases with missing items in the DSS (n = 65) the validation sample finally consisted of 598 residents, 78.1% of the residents were female; the mean age was 81.5 years. The distribution of the residents according to age, gender, and prevalence of cognitive disorders were similar to those of a Germany-wide representative study of old-age homes [30] (please see the Table Appendix 3 in the supplementary data on the Journal website).

According to the diagnostic judgements made by clinical psychologists, only 46.2% of the residents in old-age homes (n = 598) were cognitively unimpaired or questionably demented (CDR 0–0.5). Specifically, 13.7% suffered from
Recognition of dementia in old-age homes

Table 1. Relationships between the DSS and the diagnostic judgement according to CDR

<table>
<thead>
<tr>
<th>Residents of old-age homes (n = 598)</th>
<th>Mild to severe dementia CDR 1–3 (mild+)</th>
<th>Moderate to severe dementia CDR 2–3 (moderate+)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DSS Cut-off</strong></td>
<td><strong>SE</strong></td>
<td><strong>SP</strong></td>
</tr>
<tr>
<td>1/2</td>
<td>86.6</td>
<td>83.0</td>
</tr>
<tr>
<td>2/3</td>
<td>81.4</td>
<td>90.9</td>
</tr>
<tr>
<td>3/4</td>
<td>76.1</td>
<td>94.2</td>
</tr>
<tr>
<td>4/5</td>
<td>70.5</td>
<td>94.6</td>
</tr>
<tr>
<td>5/6</td>
<td>63.7</td>
<td>97.1</td>
</tr>
</tbody>
</table>

SE, sensitivity; SP, specificity; OCC, overall correct classification; PPV, positive predictive value; NPV, negative predictive value.

Sensitivity, specificity, and ROC-analyses

The validity coefficients sensitivity, specificity, overall correct classification, positive and negative predictive value with variable cut-off points are presented in Table 1. In terms of optimal balance between sensitivity (81.4%; 95% CI 77.0–85.6) and specificity (90.9%; 95% CI 87.5–94.3), the best cut-off point for the DSS total score lay at 2/3 for the whole validation sample. Overall correct classification (OCC) was 85.8%, and the positive predictive value (PPV) was 91.3%. For more severe dementia (moderate+), the best cut-off was at 3/4 associated with an OCC of 87.4%, a sensitivity of 88.8%, a specificity of 86.0%, and a PPV of 81.6%. As expected, the OCC was somewhat better for the more severely demented group (moderate+−group) than for the group that included cases of mild dementia (mild+−group). The moderate+−group also revealed the highest sensitivity.

A single measure of the overall accuracy is the AUC. With an AUC = 0.912 (95% CI 0.89–0.94), the DSS showed a high accuracy.

Analyses of different age groups revealed that the sensitivity of the DSS improved continuously with increasing age, except for the group of the 80–84 year old people, from 68.4 upto 89.2%, while the specificity remained constantly high, fluctuating at around 90% (please see the Table Appendix 4 in the supplementary data on the Journal website). In the highest age groups (90+ years), specificity and PPV also improved. There was also a slight difference in sensitivity and specificity between male and female residents, which was due to the differences in their age structure (mean age of men = 76.2 years, SD = 11.6; mean age of women = 83.0 years, SD = 8.6). More obvious was the variation in sensitivity and specificity due to the ability to walk: For those residents who were able to walk independently, we found a considerably lower sensitivity (70.2%) than for the residents who were not able to walk at all or who needed help walking. Conversely, the specificity of the DSS for the residents who were able to walk independently was very high (96%), but rather low (60.9%) for those who needed help.

Comparison with other screening instruments for dementia

Receiver operating characteristic (ROC) comparisons were then carried out for those cases for whom all screening results (DSS, MMSE, BAS DEM) were available. The characteristics of this sample (n = 502) did not differ from the distribution of the whole validation sample (n = 598).

The ROC curves of DSS, BAS DEM, and MMSE, presented in Figure 1, were calculated for the mild+−group. In comparison with the MMSE (AUC = 0.968; 95% CI 0.953–0.982) and BAS DEM (AUC = 0.942; 95% CI 0.922–0.961), the curve of the DSS performed a little worse than the MMSE, and this difference was statistically significant (zMMSE/DSS = 4.5 and zBASDEM/DSS = 2.3 respectively). Comparing at points of evenly distributed values of misclassification errors for each instrument, a sensitivity of 84.3% and a specificity of 81.8% was calculated.

Figure 1. Receiver operating characteristic (ROC) curves for DSS total score, MMSE* and BAS DEM with CDR diagnostic judgement (mild+). *The MMSE scale is converted for illustration.
for the BAS DEM at a cut-off of 1/2. The respective values for the MMSE were 88.6 and 85.8%, respectively, and for the MMSE at a cut-off of 21/22, 92.3% (sensitivity) and 89.8% (specificity), respectively. The MMSE, thus, was the instrument that performed best in our study.

Table 2 shows the ability of the three screening instruments to discriminate different grades of severity of dementia. With regard to the differentiation between ‘healthy’/‘questionable’ and ‘mild dementia’, the DSS was significantly less successful than either the MMSE or the BAS DEM. The three screening tools performed equally well with respect to the discrimination of mild and moderate stages. However, with respect to the differentiation of moderate and severe dementia, the MMSE clearly performed better than the DSS or the BAS DEM. The value of the DSS as a screening for dementia seems best for the detection of moderate to severe cases (moderate—group) (AUC = 0.939). For these cases, its usefulness is similar to that of the BAS DEM.

Partial correlations with other screening instruments confirmed the convergent validity of the DSS. Filtering out the influence of age, gender, and type of old-age home, the correlation between the DSS total score and MMSE was Pearson’s $r = -0.78$ ($P<0.001$), between the DSS and BAS DEM $r = 0.75$ ($P<0.001$) respectively.

Discussion

The characteristics of our study sample conform to previous data from a large-scale, German-wide survey [29]. Thus, the sample appears to be representative of the target population: the residents of old-age homes in Germany, at least in regard to its main characteristics (age, gender and prevalence of cognitive disorders). The size and structure of the sample can be considered as one of the strengths of our study.

Like brief screening instruments in general, the DSS proved to be unsuitable for the reliable detection of early stages of demential illness. However, the DSS correlated highly with other, well-established screening instruments like the MMSE and the BAS DEM, which are based on cognitive testing.

In comparison to global and diagnosis-related staff judgements (i.e. assessing dementia or the level of dementia without guidelines or instruments), the DSS appears to improve accuracy. Its validity was clearly better than that of the global judgements determined in previous studies [7–10]. Although comparisons between these studies are problematic because of their methodological heterogeneity, they indicate, that global proxy assessments of dementia by nursing home staff result in a considerable underestimation of the disorder in the residents [13]. Under-recognition of cognitively impaired residents was associated with non-aggressive, proper and well adjusted behaviour, as well as with better functional status (visual, auditory, motor, expressive language) of the residents [7–10]. On the other hand, truly non-demented residents who showed problematic or unusual behaviour were likely to be falsely categorised by the staff as being demented [9]. Hence, global assessments of dementia by the nursing home staff seem to be considerably confounded by the everyday functioning and other characteristics of the residents or of the nurses.

Further analyses of our data revealed that a simple instrument like the DSS cannot entirely prevent such confounds. The validity coefficients of the DSS varied with respect to the residents’ age and functional level. Although it is not possible to give clear-cut explanations for these variations — an item analysis revealed no advice in this respect — this pattern resembles the patterns found in the earlier studies [9, 10, 13].

However, one factor that may be influencing the findings of this study is the frequency of contact between the nurses and residents: usually the staff have less frequent contact — and thus, fewer chances to observe cognitive impairments — with functionally more independent and younger residents. These findings corroborate the conclusion of Magaziner [24], that the accuracy of proxy reports is substantially influenced by the chance to observe subjects, and by subject characteristics.

Another way to identify dementia in residents of old-age homes is through the use of performance-based screening instruments. The most widespread instrument of this kind is the MMSE. In our study, the sensitivity and specificity coefficients of the MMSE for dementia (milder cases included) at a cut-off of 21/22 were surprisingly high, higher than in most of the previous studies [5, 31]: 92.3 and 89.8%, respectively. One reason for the superior performance of the MMSE in our study might be the fact that the gold standard, the experts’ diagnostic ratings based on the

Table 2. Ability of DSS, MMSE and BAS DEM to discriminate between different severity grades of dementia: Results of ROC analyses ($n = 502$) and comparisons between instruments

<table>
<thead>
<tr>
<th>Severity of dementia according to CDR</th>
<th>DSS AUC (95% CI)</th>
<th>MMSE AUC (95% CI)</th>
<th>BAS DEM AUC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No/Questionable versus Mild</td>
<td>0.772 (0.712–0.832)</td>
<td>0.915* (0.884–0.947)</td>
<td>0.875* (0.833–0.918)</td>
</tr>
<tr>
<td>Mild versus Moderate</td>
<td>0.760 (0.684–0.836)</td>
<td>0.800* (0.730–0.870)</td>
<td>0.729* (0.649–0.809)</td>
</tr>
<tr>
<td>Moderate versus Severe</td>
<td>0.791 (0.725–0.857)</td>
<td>0.928* (0.889–0.967)</td>
<td>0.847* (0.784–0.911)</td>
</tr>
<tr>
<td>No/Questionable versus Mild+</td>
<td>0.898 (0.870–0.927)</td>
<td>0.968* (0.953–0.982)</td>
<td>0.942* (0.922–0.961)</td>
</tr>
<tr>
<td>No/Questionable, Mild versus Moderate+</td>
<td>0.939 (0.918–0.959)</td>
<td>0.971* (0.957–0.986)</td>
<td>0.953* (0.934–0.972)</td>
</tr>
<tr>
<td>No/Questionable, Mild, Moderate versus Severe</td>
<td>0.935 (0.913–0.956)</td>
<td>0.985* (0.976–0.994)</td>
<td>0.968* (0.954–0.982)</td>
</tr>
</tbody>
</table>

* z-Test: $P<0.05$; ns $P>0.05$ (respective difference to AUC of the DSS) calculated with $z$-scores according to Hanley & McNeil, 1983 [29].
CDR, was strongly influenced by the MMSE results. In contrast to this, the CDR results were truly independent of the DSS results. Thus, compared to the validity of the DSS, we probably have overestimated the validity of the MMSE.

Apart from validity, further aspects of a screening instrument have to be considered: Compared to the MMSE, the DSS has proved to be faster and simpler, and was associated with less missing data. While the DSS was completed for 90% of the home residents in our validation sample, 33.5% of the residents did not complete the MMSE due to various impairments (i.e. sensory, motor, communication) or refusals. Since missing data can lead to a serious underestimation of a disorder among some cohorts, the rate of non-response should be a critical issue in the selection of a screening instrument [13].

Since the under-recognition of dementia in long-term care settings is a serious problem in many European countries [7, 9], a simple and economical instrument like the DSS can help to remedy the situation. The planning and monitoring of dementia care, as well as more comprehensive epidemiological studies in institutions could be based on the DSS. To monitor dementia care it could be applied before and after introducing changes, or routinely, once a year. Further studies are recommended to examine the DSS in other care settings (i.e. community care settings) as well as to examine the outcome and efficacy of such screening procedures with respect to the quality of care of the demented.

The Dementia Screening Scale (DSS)
The possible ratings range from 0–2, referring to the frequency of occurrence (‘never’, ‘occasionally’, ‘always’) of the following cognitive impairments/functions in the preceding 4 weeks:

(i) Did he/she recognise friends/relatives/staff?
(ii) Did he/she know friends/relatives/staff by name?
(iii) Could he/she remember what happened in the past few days?
(iv) Does he/she confuse people, and does he/she not know where he/she is?
(v) Could he/she orient him-/herself in his/her room?
(vi) Could he/she orient him-/herself in the home/living area?
(vii) Could he/she orient him-/herself in the neighbourhood of the home?

The internal consistency of the DSS proved to be good with Cronbach’s alpha = 0.94.

Key points
- The DSS proved to be a valid screening tool for use by staff in old-age homes, particularly in nursing homes.
- The DSS is easier to apply, less time-consuming, and associated with a lower number of missing cases than the cognitive tests traditionally used.

References


Received 30 November 2006; accepted in revised form 24 April 2007