The MDS Challenging Behavior Profile for long-term care

D. L. GERRITSEN1,2,6, W. P. ACHTERBERG2, N. STEVERINK3, A. M. POT2,4, D. H. M. FRIJTERS2,5, & M. W. RIBBE2

1Heyendaal Institute, Radboud University Nijmegen, 2Department of Nursing Home Medicine Amsterdam, VU University Medical Center/EMGO Institute for Research in Extramural Medicine, 3Department of Health Sciences/Health Psychology, University Medical Center Groningen, University of Groningen, Groningen, 4Trimbo-institute, Netherlands Institute of Mental Health and Addiction in Utrecht, 5PRISMA NT, Dutch Center for Health Care Information, Utrecht, The Netherlands, and 6Department of Nursing Home Medicine, Radboud University Nijmegen Medical Centre, Nijmegen

(Received 3 June 2005; accepted 20 February 2007)

Abstract
The objective was to construct a reliable and valid challenging behavior scale with items from the Minimum Data Set (MDS). Exploratory factor analyses of a sample of 656 nursing home residents yielded a 16-item Behavior Profile containing four internally consistent and valid subscales measuring conflict behavior, withdrawn behavior, agitation and attention seeking behavior (alpha range: 0.69–0.80). On a second dataset of 227 nursing home residents, internal consistency, inter-rater reliability and validity against the Behavior Rating Scale for Psychogeriatric Inpatients (GIP) were established. Internal consistency of the subscales ranged between 0.54 and 0.78. The overall inter-rater reliability of the items was 0.53 (kappa); of the scale it was 0.75 (ICC). The MDS Challenging Behavior Profile could potentially be an important contribution to existing clinical MDS-scales but additional studies on reliability, validity and usefulness are needed.

Introduction
In long-term care, many residents express behavior that is considered burdensome for the residents themselves as well as a burden and a challenge for their caregivers (Aalten, de Vugt, Jaspers, Jolles & Verhey, 2005; Cohen-Mansfield & Mintzer, 2005; Zuidema et al., 2007). Care strategies for reducing this behavior can be most successful if they are based on adequate behavioral assessment (Kovach, Kelber, Simpson & Wells, 2006). Numerous instruments for assessing and documenting this ‘challenging behavior’ are available (Lam, Chan, Mok, Li & Lam, 2006) but there is little agreement on how these instruments can be used routinely by nursing home staff (Snowden, Sato, Roy-Byrne, 2003). The only widely used standard assessment procedure in daily long-term care is the Minimum Data Set (MDS).

The MDS is a structured and comprehensive questionnaire that produces a large amount of clinical information about a resident (Morris, Murphy & Nonemaker, 1995). The questions comprise information on several aspects of the patients’ functioning, health and well-being. Over the years, several measurement scales have been developed on the basis of these MDS items [e.g. the Cognitive Performance Scale (Morris et al., 1994); the Index for Social Engagement (Mor et al., 1995); and the CHESS (Hirdes, Frijters & Teare, 2003)]. A useful addition to the MDS would be to provide a measurement scale for challenging behavior. This may assist nurses in everyday care planning and help them to determine when it is necessary to intervene and what kind of care strategy is required.

Different types of challenging behaviors, e.g. apathetic or aggressive behavior, require different care strategies. Thus, specific information on residents’ type of challenging behavior would provide clues to the kind of approach that should be selected (Opie, Doyle & O’Connor, 2002). Subsequently, these care strategies could be evaluated with a scale for challenging behavior. That is why, in this study, we have made an effort to construct an internally consistent and valid scale for challenging behavior, based on items currently in the MDS.

Methods
Participants
The dataset used to construct the MDS-behavior profile consisted of MDS 2.0 assessments for...
656 residents of four nursing homes in the Netherlands, assessed between September 2002 and April 2003. These homes were the four nursing homes in the Netherlands that had, at that time, implemented the MDS as an integrated part of daily care (Gerritsen, 2004). The mean age was 81 years, 74% were women and 71% suffered from moderate to severe cognitive problems (i.e. scored >2 on the Cognitive Performance Scale) (Morris et al., 1994).

A second dataset was used to study reliability and validity of the MDS-behavior profile and the subscales. It consisted of MDS 2.0 assessments of 227 nursing home residents of ten nursing homes in the Netherlands, assessed between January 2000 and June 2002 (Gerritsen et al., 2004). These were 10 of the 11 Dutch homes that were in the process of implementing the MDS at that time. Dual assessments by a second rater were available for 151 residents. The mean age of these 227 residents was 79.9 years, 78% were women and 54% suffered from moderate to severe cognitive problems (i.e. scored >2 on the Cognitive Performance Scale).

Although the four nursing homes of the first sample were also part of the second sample, no individual residents were part of both samples.

Measures

The MDS is collected from observations by formal caregivers and interviews with residents and family members. In the Netherlands, it is completed by Licensed Practical Nurses (LPN) involved in daily care, who know the resident well. This assessment is performed within two weeks after admission, once every three months and when there is an important change in health status.

The MDS contains several behavioral items. These have shown good reliability and validity (Frederiksen, Tariot & De Jonghe, 1996; Hawes et al., 1995; Morris et al., 1990; Snowden et al., 1999; 2003) and have been found useful as indicators of behavioral problems, but they lack a severity score (Snowdon et al., 2003).

For a study of the concurrent validity, the Behavior Rating Scale for Psychogeriatric Inpatients (GIP) was used in the second sample. The 82-item GIP is a behavior rating scale consisting of 14 subscales that is administered by LPNs. It covers aspects of behavioral, affective and cognitive functioning. Each subscale has five to eight items to be scored on a four-point scale (never/sometimes/often/very frequent). The GIP is widely used and of known reliability and validity in Dutch long-term care facilities (de Jonghe, Kat & De Reus, 1994; Verstraten, 1988).

Complete GIP-subscases were available for 211 to 218 residents. For this paper GIP-subscases were used that measure concepts that corresponded best with the concepts that are measured with the new subscales. These were the GIP non-compliant behavior subscale, the GIP restless-behavior subscale, the GIP repetitive-behavior subscale, the GIP dependent-behavior subscale and the GIP socially withdrawn-behavior subscale. The latter scale consists of eight items, the others five.

For each resident the same LPN assessed the MDS and the GIP within a 4-week period.

Selection of the items

Five clinical experts (two nursing home physicians and three nursing home psychologists) used to working with the MDS in clinical practice were asked to select items on challenging resident behavior from the current MDS that they considered to be important for routine assessment and subsequent care planning. Items that were selected by none or only one expert were excluded.

The MDS-items that were selected by two or more of the experts were: periods of restlessness, negative statements, repetitive questions, repetitive verbalizations, repetitive persistent anger with self or others, repetitive health complaints, repetitive anxious complaints/concerns, repetitive physical movements, withdrawal from activities of interest, reduced social interaction, wandering, verbally abusive behaviors, physically abusive behaviors, socially inappropriate or disruptive behavior, resisting care, covert/open conflict with or repeated criticism of staff, recurrent statements that something terrible is about to happen, unhappy with roommate, unhappy with residents other than roommate and openly expressing conflict/anger with family/friends.

Subsequently, frequency distributions of the selected items were studied in our first sample, which consisted of a subgroup of nursing home residents with Alzheimer’s disease; a subgroup with other dementias; a subgroup without dementia; and in the group as a whole. If an item had a very low endorsement rate in one or more of these groups (frequency below 10%) it was considered to be non-discriminative (Streiner & Norman, 1995) and was therefore discarded. This pertained to the following items: recurrent statements that something terrible is about to happen; unhappy with roommate; unhappy with residents other than roommate; and openly expresses conflict/anger with family/friends.

Data analysis: Scale construction

Depending on their MDS-section, the selected MDS items have different response categories (the items and their description can be found in Table I). The response categories of the items with three categories were: 0 = behavior not exhibited; 1 = behavior exhibited up to five days a week; 2 = behavior exhibited daily or almost daily (6–7 days a week). The response categories of the
Table 1. Internal consistency, factor structure and descriptives of the MDS Challenging Behavior Profile and subscales in the first sample of nursing home residents (N= 656).

<table>
<thead>
<tr>
<th>Item-description</th>
<th>Item-name in MDS</th>
<th>Response categories(^1)</th>
<th>PCA(^2) with oblique rotation</th>
<th>Alpha</th>
<th>Range (max of scale)</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conflict</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetitive persistent anger with self or others</td>
<td>E1d</td>
<td>3</td>
<td>0.017</td>
<td>−0.376</td>
<td>0.061</td>
<td>−0.573</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbally abusive behaviors</td>
<td>E4ba</td>
<td>3 (4)</td>
<td>0.275</td>
<td>−0.156</td>
<td>−0.064</td>
<td>−0.621</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physically abusive behaviors</td>
<td>E4ca</td>
<td>3 (4)</td>
<td>0.308</td>
<td>0.204</td>
<td>−0.166</td>
<td>−0.677</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resists care</td>
<td>E4ca</td>
<td>3 (4)</td>
<td>0.023</td>
<td>0.114</td>
<td>0.140</td>
<td>−0.656</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflict with or repeated criticism of staff</td>
<td>F2a</td>
<td>2</td>
<td>−0.289</td>
<td>−0.154</td>
<td>0.130</td>
<td>−0.578</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Withdrawal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Withdrawal from activities of interest</td>
<td>E1o</td>
<td>3</td>
<td>0.070</td>
<td>−0.034</td>
<td>0.889</td>
<td>0.013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced social interaction</td>
<td>E1p</td>
<td>3</td>
<td>0.154</td>
<td>0.052</td>
<td>0.868</td>
<td>−0.025</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Agitation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Periods of restlessness</td>
<td>B5d</td>
<td>2 (3)</td>
<td>0.630</td>
<td>−0.207</td>
<td>0.049</td>
<td>−0.019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetitive physical movements</td>
<td>E1n</td>
<td>3</td>
<td>0.760</td>
<td>−0.038</td>
<td>0.209</td>
<td>0.038</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wandering</td>
<td>E4aa</td>
<td>3 (4)</td>
<td>0.603</td>
<td>0.174</td>
<td>0.119</td>
<td>−0.126</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socially inappropriate/disruptive behavior</td>
<td>E4da</td>
<td>3 (4)</td>
<td>0.527</td>
<td>−0.156</td>
<td>−0.095</td>
<td>−0.378</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attention seeking</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative statements</td>
<td>E1a</td>
<td>3</td>
<td>−0.127</td>
<td>−0.566</td>
<td>0.111</td>
<td>−0.391</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetitive questions</td>
<td>E1b</td>
<td>3</td>
<td>0.335</td>
<td>−0.704</td>
<td>−0.019</td>
<td>0.104</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetitive verbalizations</td>
<td>E1c</td>
<td>3</td>
<td>0.417</td>
<td>−0.566</td>
<td>−0.083</td>
<td>−0.018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetitive health complaints</td>
<td>E1h</td>
<td>3</td>
<td>−0.170</td>
<td>−0.725</td>
<td>0.008</td>
<td>−0.014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetitive anxious complaints and/or concerns</td>
<td>E1i</td>
<td>3</td>
<td>0.014</td>
<td>−0.779</td>
<td>0.029</td>
<td>−0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)Number of response categories after (before between brackets) recoding.
\(^2\)Principal component analyses.
items with four categories were: 0 = behavior not exhibited; 1 = behavior of this type occurred on 1–3 days in last seven days; 2 = behavior of this type occurred on 4–6 days, but less than daily; 3 = behavior of this type occurred daily. For an optimally balanced contribution of each item to the scale score, categories 1 and 2 of the items with four response categories were taken together and recoded into 1, and category 3 was recoded into 2.

In addition, the item ‘periods of restlessness’ (B5d) was recoded. Its original response categories were: 0 = behavior not present; 1 = behavior present, not of recent onset; 2 = behavior present, over last seven days appears different from resident’s usual functioning. For this item, categories 1 and 2 were taken together and coded into 1.

The selected items were analyzed for internal consistency and principal components. Internal consistency was calculated by means of Cronbach’s alpha (Cronbach, 1951). Cronbach’s alpha is considered to be good if higher than 0.70 but should not be higher than 0.90 (Steiner & Norman, 1995). In principal component analysis, the component solution (eigenvalues above 1) was rotated with oblique rotation (OBLIMIN) in order to find subscales with meaningful content. Oblique rotation was used as it was expected that the scales would be correlated.

Data analysis: Testing the scale’s properties on the second dataset

Dual MDS-assessments were completed by trained LPNs in the participating facilities, with the second assessment completed after at least one week, but no later than one month, after the first assessment. This enabled the calculation of inter-rater reliability of the items and the scales. The inter-rater kappa values of the items (Cohen, 1968; Landis & Koch, 1975) and intra-class correlation coefficients of the scale-scores (Shrout & Fleiss, 1979) were calculated to determine inter-rater reliability. The Landis and Koch classification was used to interpret both the kappa results and the ICC coefficients (Montgomery, Graham, Evans & Fahey, 2002) (0.00–0.20 = slight; 0.21–0.40 = fair; 0.41–0.60 = moderate; 0.61–0.80 = substantial; and 0.81–1.0 = almost perfect). Principal component analysis with four factors and oblique rotation was performed to provide insight into whether the four identified behavioral dimensions could also be found in this second sample.

Information on concurrent validity was obtained by calculating Spearman’s correlation coefficients of the subscale-scores with the GIP. For this analysis, each new subscale was paired with a GIP-subscale, based on content. It was hypothesized that: (1) the two scales in a pair would correlate significantly to each other and that (2) each pair would have the highest coefficient if all new subscales were correlated to the GIP-subscgales.

Results

Scale construction

Principal component analysis with oblique rotation revealed four subscales with meaningful content. These were specified as conflict, withdrawal, agitation and attention seeking. The item ‘repetitive verbalizations’ (E1c), loaded above 0.40 on two factors (see Table I). Subsequent internal consistency analyses of the subscales revealed alpha’s ranging from 0.69–0.80. All subscales were positively skewed (see Table I for descriptives). The Spearman correlations among the four scales were significant (p < 0.01) and ranged from 0.18–0.50 (not in Table I), confirming the need to use oblique rotation.

Although the subscales will have the greatest clinical relevance when they are calculated separately and considered as a behavioral profile, the scores could be summated because the items form an internally consistent overall scale. Cronbach’s alpha of the 16 MDS items used in this study was 0.83 and all items contributed to the scale.

The total score can be used as a basic indicator for the presence of challenging behavior. For calculating a total score, it was decided to simply summate the residents’ scores on all 16 items. The subscales have different ranges and, as a consequence, they contribute with varying strength to the overall scale. On this MDS-behavior profile, which had a range of 0–24 in our sample (with a theoretical maximum of 30), 82% of all residents had a score above 0; 50% had a score of 4 or higher; and 25% had a score of 8 or higher.

Properties of the challenging behavior profile in the second sample

Table II shows some basic descriptives and the results of the reliability analysis in the second sample. With the exception of conflict (alpha 0.54), the subscales appeared to be sufficiently internally consistent in the second sample (alpha agitation: 0.70, attention seeking: 0.75 and withdrawal: 0.78). The squared weighted kappas were satisfactory, with only one item (‘resisting care’ [E4ea]) having a kappa value below 0.40. The mean and 95% confidence intervals (CI) of Intra-Class Correlation coefficients were moderate to almost perfect, although the 95% CI of withdrawal included 0.35, which is fair.

Principal component-analysis with four factors and oblique rotation revealed a factor structure in the second sample (see Table III) that was not as strong as in the first sample. The withdrawal and agitation subscales emerged with loadings ranging...
Table II. Descriptives and reliability of the MDS Challenging Behavior Profile and subscales in the second (validation) sample of nursing home residents (N = 227.).

<table>
<thead>
<tr>
<th>(sub-) scale</th>
<th>N items</th>
<th>Theoretical max. score</th>
<th>Range in sample</th>
<th>Median</th>
<th>Mean</th>
<th>SD</th>
<th>Alpha</th>
<th>Mean K</th>
<th>Range</th>
<th>ICC¹</th>
<th>95% CI²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conflict</td>
<td>5</td>
<td>9</td>
<td>0–6</td>
<td>0</td>
<td>0.8</td>
<td>1.3</td>
<td>0.54</td>
<td>0.49</td>
<td>0.34–0.70</td>
<td>0.59</td>
<td>0.47–0.68</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>2</td>
<td>4</td>
<td>0–4</td>
<td>0</td>
<td>0.7</td>
<td>1.3</td>
<td>0.78</td>
<td>0.44</td>
<td>0.44–0.44</td>
<td>0.48</td>
<td>0.35–0.59</td>
</tr>
<tr>
<td>Agitation</td>
<td>4</td>
<td>7</td>
<td>0–7</td>
<td>0</td>
<td>1.3</td>
<td>1.3</td>
<td>0.70</td>
<td>0.61</td>
<td>0.47–0.71</td>
<td>0.80</td>
<td>0.73–0.85</td>
</tr>
<tr>
<td>Attention seeking</td>
<td>5</td>
<td>10</td>
<td>0–9</td>
<td>0</td>
<td>1.4</td>
<td>2.2</td>
<td>0.75</td>
<td>0.53</td>
<td>0.43–0.62</td>
<td>0.68</td>
<td>0.58–0.76</td>
</tr>
<tr>
<td>Overall CBP</td>
<td>16</td>
<td>30</td>
<td>0–20</td>
<td>3</td>
<td>4.2</td>
<td>4.6</td>
<td>0.81</td>
<td>0.53</td>
<td>0.34–0.71</td>
<td>0.75</td>
<td>0.67–0.81</td>
</tr>
</tbody>
</table>

¹ICC: Intraclass Correlation Coefficient.
²95%CI = 95%CI of Intraclass Correlation Coefficient.
from 0.765–0.807, and 0.551–0.855 respectively. However, it was difficult to distinguish the conflict subscale from the attention seeking subscale; three of the five conflict items loaded higher on the attention-seeking factor than on the conflict factor.

The new subscales were all significantly correlated with their respective GIP subscales: Spearman’s rho of Conflict behavior with GIP non-compliant behavior was 0.53; the rho of withdrawal with GIP socially withdrawn behavior was 0.35; rho of Agitation with GIP restless behavior and GIP repetitive behavior were 0.50 and 0.51 respectively; and the rho of attention seeking with GIP dependent behavior was 0.23. The results can be found in Table IV.

Each scale was correlated highest with its corresponding GIP-subscale, with the exception of one: attention seeking correlated about the same to GIP-dependent behavior and GIP-restless behavior (0.23 and 0.24 respectively).

### Discussion

In this study, we have constructed the MDS Challenging Behavior Profile, which is available for all long-term care facilities using the MDS. This scale is designed for a better understanding and management of challenging behavior in long-term care. It consists of four subscales, measuring the behavioral dimensions: conflict behavior, withdrawal, agitation and attention seeking. The subscales were found to relate significantly to other scales measuring the same constructs.

Although the subscales were quite distinguishable in the first sample, internal consistency and principal component analyses in the second sample showed

![Table III. Factor analyses of the MDS Challenging Behavior Profile on the second (validation) sample (N = 227).](image)

![Table IV. Spearman correlation coefficients of the MDS Challenging Behavior Profile (CBP) with GIP subscales (Behavior Rating Scale for psychogeriatric inpatients). N_range: 208–226.](image)

*p<0.05; **p<0.01.
less distinct relationships between the items. In the second sample, the conflict subscale was difficult to distinguish from the attention-seeking subscale and had a low alpha. Indeed, the observed behaviors in these two scales are more similar than those in the other two subscales. Conflict and attention seeking behavior may be expressions of one behavioral dimension in which conflict behavior is the stronger form. Further research on independent samples should confirm whether the conflict subscale can be distinguished from the attention-seeking subscale.

In contrast to the other three GIP subscales, the GIP-dependent behavior scale did not have strong relationships with the new subscales. This may have been caused by the low internal consistency of the GIP-dependent behavior (alpha of 0.61), which may explain why the new attention-seeking subscale did not correlate more strongly with the GIP-dependent behavior scale than with the other GIP-scales.

A limitation of our study is that, by using items included in the MDS, it is likely that we have missed some behavioral features. The Neuro-Psychiatric Inventory of Cummings and colleagues (1994) for instance, contains items on irritated and disinhibited behavior. Suspicious behavior is an important feature in other behavioral scales (e.g. Verstraten, 1988). Moreover, some items may not have been detailed enough. For instance, the subtext of the item ‘socially inappropriate or disruptive behavior’ contains very different behaviors (e.g. ‘made disruptive sounds’, ‘sexual behavior’ and ‘hoarding’). Although we included a broad range of behavior and ensured that all scales differ in the behavioral features they address, further research should study the need to include other items in the scale and, thus, in the MDS, as well as the need to make the MDS-items more detailed. Notwithstanding these concerns, one great advantage of using MDS-items for a behavior scale is that thousands of long-term care facilities could have instant access to a specific profile of behavior of their residents that requires no additional primary collection of data. If followed by adequate additional analysis (with more extensive instruments if needed) and interventions, this can have a positive impact on the quality of life of residents of nursing homes worldwide. The scores on the subscales may prove to be helpful in evaluating the actions taken in the care plan on these dimensions of behavior.

We have developed a measurement scale for challenging behavior based on MDS-items, which is shown to have reasonable psychometric properties. Because the MDS is the mandated standard assessment instrument for all nursing homes in the US and is used in many other countries (Hirdes et al., 1999), the Challenging Behavior Profile may be very useful in care planning and the evaluation of care. However, there is a need for additional studies on reliability, validity, usefulness and sensitivity to change in other countries, in non-research settings and against other instruments. Additional studies are needed to investigate whether the MDS should be expanded to incorporate challenging behavior features more extensively and meticulously. The ability of this scale to identify specific groups of patients that will benefit from additional behavioral assessment instruments needs to be established. The impact of behavior on both caregiver distress and resident well-being makes this behavioral profile an important contribution to the suite of clinical MDS-scales.

Acknowledgements

We want to thank Marcel Ooms for his important contribution in the design and analysis of this study and Dirk Knol of EMGO for his advise on the analysis.

References


