Evaluating the Accuracy of Minimum Data Set Bed-Mobility Ratings Against Independent Performance Assessments: Systematic Error and Directions for Improvement

Barbara M. Bates-Jensen, PhD, RN, CWOCN, Sandra F. Simmons, PhD, John F. Schnelle, PhD, and Cathy Alessi, MD

Purpose: The Minimum Data Set (MDS) Activities of Daily Living (ADL) bed-mobility item, which rates the staff-assistance level necessary for bed movement, is used to target scheduled repositioning interventions and to identify physical function changes in nursing home residents; however, accuracy of the item is uncertain. The purpose of this study was to evaluate the accuracy of the MDS ADL bed-mobility item as completed by nursing home nurses with independent performance assessments conducted by research staff. Design and Methods: A convenience sample of 197 long-stay residents from 26 California nursing homes participating in a larger project was used in this cross-sectional study to compare independent research-staff performance assessments (using graduated assistance protocols of residents’ ability to move in bed) and nursing home nurse MDS bed-mobility ratings. Participants also wore movement monitors to verify performance assessments. Results: Poor agreement existed between the nursing home nurse MDS bed-mobility ratings and the research-staff performance assessments across all assistance levels (kappa range, $\kappa = 0.007$, $p = .918$ to $\kappa = 0.484$, $p < .001$), with better agreement seen in totally dependent participants and with fewer elapsed days between MDS ratings and performance assessments. The odds of nursing home nurse errors (underestimating or overestimating dependency) on the MDS bed-mobility item were 2.1 times higher for participants judged independent by research staff compared with participants judged as requiring physical assistance by research staff (95% confidence interval, 1.14–4.03) when adjusted for number of days between nurse MDS ratings and research-staff performance assessments. Implications: Nursing home nurses overestimated resident dependency in bed mobility. The systematic inaccuracies in MDS bed-mobility ratings have implications for their use as a basis for targeting residents for repositioning programs and determining changes in residents’ physical function. Performance assessments utilizing graduated assistance protocols are recommended as a method of improving the accuracy of MDS bed-mobility ratings.

Key Words: Activities of daily living assessment, Bed mobility, Minimum data set accuracy, Physical performance assessments

The Minimum Data Set (MDS) is a periodic comprehensive assessment that has been required for nursing home residents since implementation of the Omnibus Reconciliation Act of 1987. The MDS assessment is completed by nursing home nurses, and it is done on admission, once each quarter, and when there is a change in a patient’s condition. Selected MDS items trigger resident assessment protocols for standardized care planning. Implementation of the MDS assessment and resident assessment protocols were originally intended to improve the quality of nursing home care through uniform resident assessment and care planning. Use of MDS data has expanded considerably from the original purpose. Most notably, MDS data are used by the Centers for Medicare and Medicaid Services (CMS) to calculate...
scores on quality measures that are published on a Web site (www.nursinghomecompare.org) for all nursing homes in the United States. The public reporting system is available to consumers, contains quality-related information on every nursing home, and implies judgments about the quality of care delivery in nursing homes (e.g., facilities with a higher prevalence of pressure ulcers, according to MDS data, provide poorer care quality). Further, MDS data are used to inform regulators (e.g., state nursing home survey agencies), as a basis for prospective reimbursement for care in some states and are used to judge the efforts of state quality-improvement organizations. However, there is controversy about the accuracy of MDS data.

Studies using different methodologies to assess MDS accuracy have reported different conclusions. Most studies have evaluated MDS accuracy by comparing the ratings of research nurses to ratings of nursing home nurses. These studies have generally concluded that the MDS ratings are accurate because there is a high agreement between the two sources (Mor et al., 2003; Phillips et al., 1997; Snowden et al., 1999). The problem with this approach is that the research nurses often rely on data that are collected from nursing home staff sources (e.g., asking nurses about what the residents can do or reviewing staff medical-record documentation). Thus, the MDS ratings completed by research nurses and nursing home nurses, at least partially, rely on the same data sources.

In contrast, studies that have used research staff to conduct independent assessments (e.g., direct observation or resident interviews) of depression, pain, oral food and fluid intake, and incontinence frequency without either asking nursing home nurses for information or relying on medical-record data have reported discrepancies between nursing home nurse MDS ratings and independent assessments conducted by research staff (Cadogan, Schnelle, Yamamoto-Mitani, Cabrera, & Simmons, 2004; Pokrywka et al., 1997; Schnelle et al., 2003; Simmons & Reuben, 2000; Simmons et al., 2004). Independent performance-based assessments of residents’ cognitive status and communication ability, however, have demonstrated closer agreement with nursing home nurse MDS ratings (Frederiksen, Tariot, & De Jonghe, 1996; Hartmeier et al., 1995).

One explanation for the discrepancies between independent performance-based assessments and MDS data is that nursing home nurses often do not use standardized assessments to determine resident status (Cadogan et al., 2004; Pokrywka et al., 1997; Schnelle et al., 2003; Simmons & Reuben, 2000; Simmons et al., 2004). For example, if nursing home nurses do not conduct assessments that allow residents to self-perform activities of daily living (ADLs) with the least amount of assistance necessary and instead rely on the assistance provided by nurse aides (whether needed or not) during daily ADL-care delivery, the MDS data might underestimate ADL self-performance when compared with independent performance assessments using standardized protocols. Nurse aides often provide excessive physical assistance for nursing home residents because it requires less time than allowing the resident to complete the task independently. This practice reinforces physical dependency in nursing home residents (Baltes, Honn, Barton, Orzach, & Lago, 1983; Baltes, Kindermann, Reisenzin, & Schmid, 1987). Licensed nursing home nurses who complete the MDS based on observations of nurse aide care delivery or who ask nurse aides how much assistance they provide to residents would likely yield measures of resident dependency and helplessness instead of a resident’s ability to self-perform the activity. If this is the case, then there are implications for how data from the MDS ADL items are used in care planning and as indicators of quality. Further, relying on data obtained by nurse aides, either by verbal communication or by observation of daily care, is in contrast to how ADLs are typically assessed in the community, where nurses conduct physical performance tests to evaluate an individual’s abilities. Thus, comparisons of ADL function between nursing home residents and community-dwelling elders may be invalid.

One of the ADL items evaluated in the MDS is bed mobility. Dependency in bed mobility or decline in bed-mobility function is used by nursing home staff to target residents who should be placed on a scheduled repositioning program. If residents are inappropriately rated as dependent, these residents may be incorrectly targeted for labor-intensive scheduled repositioning interventions, and clinically significant declines in ADL bed-mobility functioning (e.g., increases in dependency) are likely to go unrecognized. The failure by nursing home staff to recognize and intervene when a resident declines in bed mobility could contribute to poor resident outcomes, such as pressure ulcer development. Current efforts to monitor ADL bed-mobility decline as part of overall ADL function as an indicator of nursing home care quality (e.g., indicator; prevalence of residents with decline in physical functioning) also might be misleading.

Our purpose in this study was to evaluate the accuracy of the MDS bed-mobility item. If this physical function item is inaccurate because it is based on a comparison between nursing home nurse ratings and a standardized performance-based assessment, then it is likely that other MDS ADL items also are inaccurate because of a lack of performance-based assessments being used by nursing home nurses. In this study we report data that address the accuracy of the bed-mobility item as one part of the MDS ADL assessment. We examined the following research questions:

1. How accurate are nursing home nurse MDS self-performance ratings of bed mobility when they
Table 1. Definitions of Minimum Data Set Bed-Mobility Self-Performance Ratings and the Independent Physical Performance Assessment Assistance Levels

<table>
<thead>
<tr>
<th>Minimum Data Set Bed-Mobility Self-Performance Ratings</th>
<th>Independent Physical Performance Assessment Assistance Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 = Independent: no help or staff oversight</td>
<td>0 = Request only, no physical assistance (e.g., “please turn on your side”)</td>
</tr>
<tr>
<td>1 = Supervision: oversight, encouragement, cueing provided</td>
<td>1 = No physical assistance but verbal cues or instructions on how to perform the activity (e.g., “grab the siderail and pull yourself over”)</td>
</tr>
<tr>
<td>2 = Limited assistance: physical help in guided maneuvering of limbs or non-weight-bearing assistance</td>
<td>2 = Verbal cues plus manual physical guidance to start the movement (e.g., take hand and move hand to the siderail)</td>
</tr>
<tr>
<td>3 = Extensive assistance: full staff performance during part of last 7 days and weight-bearing assistance</td>
<td>3 = Partial physical assistance (e.g., flex arm and shoulder or hip to assist with turning movement)</td>
</tr>
<tr>
<td>4 = Total dependence: full staff performance during entire 7 day period</td>
<td>4 = Unable to perform without complete physical assistance</td>
</tr>
</tbody>
</table>

are compared with an independent performance assessment?

2. Are discrepancies between independent performance assessments and MDS ratings of bed mobility random or systematic (i.e., is there a consistent trend for overestimating or underestimating dependency level)?

Methods

Participants and Setting

We analyzed data from a larger project including 26 California nursing homes that scored in either the lower or upper quartile on seven MDS quality indicators. The larger project is more fully described elsewhere in a series of articles (Bates-Jensen et al., 2003; Cadogan et al., 2004; Schnelle et al., 2003; Schnelle, Bates-Jensen, et al., 2004; Schnelle, Simmons, et al., 2004; Simmons et al., 2004). The University of California Los Angeles Office for the Protection of Research Subjects approved the consent procedures for the project.

In this article, we report on a convenience sample of 197 long-stay (i.e., excluding Medicare short-stay or subacute rehabilitation residents) nursing home residents from the 26 facilities in the larger project who had independent bed-mobility performance assessments by research staff and had corresponding movement monitor data (subsequently described). Persons who are admitted to the nursing home as Medicare short-stay residents are in the facility to receive rehabilitation services, and, if rehabilitation is successful, they are discharged within 90 days. The study sample included only long-stay residents, that is, those admitted with chronic care needs or those who failed rehabilitation efforts.

MDS Bed-Mobility Item

The MDS bed-mobility item (MDS item G1Aa) is part of the ADL physical function section, which requires nursing home nurses to rate the resident on the basis of the most frequent level of assistance required over the past 7 days for self-performance. Nursing home nurses also rate the type and amount of support received by the resident from nursing home staff (item G1Ba). Instructions advise that the two ratings be made independently and that self-performance be determined first (Health Care Financing Administration [HCFA], 1999). The intent of the ADL self-performance ratings is to record what the resident actually did independently along with the level and type of staff assistance needed to perform the action. This implies that, in order for ADL function to be assessed, the resident must be given the opportunity to perform activities with the minimum level of assistance. Nursing home nurses rate residents’ self-performance assistance levels on a 5-point scale from 0 (independent) to 4 (totally dependent; see Table 1). The definitions of each level of assistance on the 5-point scale suggest the use of a graduated assistance protocol to determine the abilities of a resident to perform each activity, although the instructions do not explicitly state this and a specific assessment protocol is not provided (HCFA).

Measures

Research staff extracted descriptive data for all participants from the medical record and the most recent annual or quarterly MDS assessment. As the “gold standard,” independent research staff (blinded to medical record and MDS data) assessed participants’ bed mobility by utilizing a standardized graduated assistance protocol, developed prior to study onset. Participants were first asked to perform the desired action (i.e., “please sit up in bed”), using the minimum level of human assistance possible, and then provided with graduated levels of assistance, as needed. We defined the amount of assistance provided according to five levels ranging from 0 (able to perform on command) to 4 (unable to perform without full physical assistance; see Table 1). Assistance
was provided in graduated levels (e.g., request only, verbal cues, manual physical guidance, and partial physical assistance) until the resident was able to complete the action or it was determined the resident was unable to complete the action without complete physical assistance. For example, if a resident failed to complete the action requested, the request was made again and verbal cueing was provided (e.g., “place your hand on the side rail and pull yourself over to turn”). If the resident remained unable to perform the action, the action was requested again, verbal cues were again provided, and minimal physical guidance was given (e.g., placing the resident’s hand on the side rail). We evaluated each participant’s ability to sit up in bed and turn in bed because both abilities are included in the MDS bed-mobility item (HCFA, 1999).

Upon completion of the physical performance assessment, participants were categorized as independent (Table 1, Levels 0 and 1) or dependent (Table 1, Levels 2, 3, and 4) on the basis of their need for physical assistance to perform the actions. There was perfect agreement between two observers classifying participants’ assistance level during the performance assessments (κ = 1.0, n = 44 observations in 22 participants). Performance assessments were repeated for a different subsample of 45 participants on 2 separate days within 1 week, with results demonstrating good stability for assistance needs (r = .90, p = .007). The mean time for research staff to conduct the performance assessments was 3.4 min (SD = 2.4). All physical performance assessments conducted by research staff occurred after nursing home nurse MDS ratings.

Each participant also wore a wireless movement monitor on the thigh above the knee for 8 to 24 hr with a mean wearing time of 10 hr (SD = 2.7) in the daytime and 11.5 hr (SD = 1.0) in the nighttime. The movement monitors (AugmenTech, Inc., Pittsburgh, PA) measured horizontal and vertical orientation of the thigh with respect to gravity every 4 s. Prior research has demonstrated the validity of the movement monitors compared with direct observations of movement, including repositioning and activity movements (Bates-Jensen et al., 2003). Interrater agreement for movement-monitor data interpretation was high (n = 63 participants, κ = 0.691 and 0.609, both ps < .001 for identifying movement and repositioning, respectively). The correlation between raters for total frequency of movements was also high (r = .92, p < .001). Movement data were used to verify the participants’ ability to move independently in bed compared with research-staff ratings based on the performance assessments.

Data Analysis

We conducted kappa statistics to evaluate strict agreement between bed-mobility ratings on the MDS at each of the five ratings of assistance and the independent research-staff performance assessments. We recoded data to allow a comparison between each category versus all other categories collapsed (SPSS Advanced Statistics, Version 11.5, 2002). We also conducted kappa statistics by using subsamples of participants with 30, 45, and 60 days or fewer between MDS ratings and research-staff performance assessments.

As a more lenient measure, we also defined agreement as the MDS rating that matched the performance-assessment level plus MDS ratings one category lower or higher. For example, if performance was assessed as Level 2, then we considered MDS ratings of 1, 2, or 3 to be in agreement. We compared nursing home nurse MDS bed-mobility (item G1Aa) ratings with research-staff performance assessments for participants’ ability to sit up in bed and turn in bed separately, using both the strict (exact match) and lenient (exact match ± one category) criteria.

Utilizing t tests for continuous variables and chi-square for categorial variables, we compared nursing homes that scored in the extreme quartiles for MDS bed-mobility rating accuracy on size, profit status, staff turnover rate, total staffing hours per resident per day (HPRD), nurse aide HPRD, registered and licensed vocational nurse HPRD, expenditures for direct resident care per resident day, and proportion of Medicaid and non-Caucasian residents. These data were all available from a public reporting system in California (www.calhhs.org).

We also used t tests (e.g., average number of movements per hour) to determine if participants judged by research staff as independent on the basis of performance assessments (e.g., Levels 0 and 1) actually did move according to the movement monitors. We calculated the average number of movements per hour by summing all movements according to movement monitor data and dividing by the total number of hours the monitor was worn.

We performed a post hoc logistic regression analysis, with nursing home nurse errors on the MDS bed-mobility item as the outcome variable (the strict measure of agreement), and performance assessment as the predictor variable. We defined nursing home nurse errors as exact match or no match between the MDS rating and the performance assessment. We rated participants as 0 or 1 on performance assessment versus those rated as 2, 3, or 4. We adjusted for the number of days between nursing home nurse MDS rating and performance assessments.

Results

Participants were predominantly Caucasian females with some level of cognitive impairment (see Table 2). Median length of residency was 15.2 months, with 22% of the participants residing in the
facility for less than 6 months and 22% residing there 36 months or longer. The number of participants per nursing home ranged from 2 to 24 ($M = 7.6$, $SD = 5.2$), and 11 facilities had fewer than 6 participants.

Within participants, research-staff assessments of sit-up-in-bed assistance levels were highly correlated with research-staff assessments of turn-in-bed assistance levels (Spearman’s rho, $r = .85$, $p < .001$). Results were comparable for all analyses utilizing sit-up-in-bed and turn-in-bed assessments; thus, we present only turn-in-bed performance assessment results here.

**MDS Rating and Performance Assessment: Strict Criterion (Exact Match)**

Poor agreement existed between research staff turn-in-bed performance assessments and nursing home nurse MDS bed-mobility ratings, with kappa agreement according to the strict criterion (exact match) ranging from $\kappa = 0.007$ ($p = .918$) for Level 2 to $\kappa = 0.484$ ($p < .001$) for Level 4 (Table 3). Better agreement existed for all participants at the extreme of dependence. Agreement improved among participants with 30, 45, and 60 days or fewer between MDS ratings and performance assessments, although this improvement was not consistent between levels. Participants with 30 days or fewer between MDS ratings and performance assessments and who were judged as independent (e.g., MDS rating = 0) by nursing home nurses showed a high level of agreement with performance assessments ($\kappa = 0.639$, $p = .021$), but this better agreement at rating 0 declined in participants with 45 days or fewer ($\kappa = 0.505$, $p = .009$) and was not seen in participants with 60 days or fewer between MDS ratings and performance assessments ($\kappa = 0.363$, $p = .013$). Within Level 4 (totally dependent), those residents with 45 days or fewer between MDS ratings and performance assessments showed the best agreement ($\kappa = 0.653$, $p = .001$). Overall, there was a trend for agreement to improve as elapsed time between performance assessments and the MDS ratings decreased. The small sample of residents with 30 and 45 days or fewer between performance assessments and MDS ratings was distributed fairly evenly across levels based on performance assessments (Table 3).

**MDS Rating and Performance Assessment: Lenient Criterion**

The agreement results comparing nursing home nurse MDS bed-mobility ratings with turn-in-bed research-staff performance assessments using the more lenient definition of agreement (exact match ± one category) are shown in Table 4. MDS ratings overestimated participants’ dependence compared with research-staff performance assessments. Specifically, 61% (Table 4, column 3, Levels 0,1; $n = 38$) of the 62 participants rated by research staff as independent (Table 4, column 1, Levels 0,1) were rated by nursing home nurses as requiring physical assistance for bed mobility (Table 4, column 2) on the MDS.

Increased agreement existed between nursing home nurse MDS ratings and research-staff performance assessments for residents judged more dependent by research staff. Nursing home nurse ratings agreed for 84% (Table 4, column 1, Levels 2, 3, 4; $n = 113$) of the 135 participants rated by research staff as dependent according to performance assessments. Only 13% ($n = 17$) of the 135 participants rated by research staff as physically dependent (Levels 2, 3, 4) were rated by nursing home nurses as less dependent (Levels 0, 1) on the MDS bed-mobility item. Thus, the error pattern in the MDS ratings by nursing home nurses was one of overestimation of dependence.

**Logistic Regression Analysis**

The odds of nursing home nurse MDS bed-mobility rating errors (according to the strict criterion) for participants judged independent (Levels 0, 1) by research staff were 2.1 (confidence interval, or CI = 1.14–4.03) times higher than for participants judged by research staff as physically dependent (Levels 2, 3, 4), when adjusted for days between MDS rating and performance assessment.

**Facility-Level Analyses**

The proportion of residents with MDS bed-mobility rating errors (according to the strict
criterion) per facility ranged from 17% to 100% with a mean of 51% (SD = 19.49). There were no study homes with 0% errors in MDS bed-mobility ratings. No differences existed between homes scoring in the extreme quartiles (lower quartile, 32% or fewer MDS errors, n = 6; upper quartile, 67% or more MDS errors, n = 8) for MDS bed-mobility rating accuracy on any of the following facility characteristics: size, profit status, staff turnover rate, total staffing HPRD, nurse aide HPRD, registered and licensed vocational nurse HPRD, expenditures for direct resident care per resident day, and proportion of Medicaid and non-Caucasian residents.

Movement-Monitor Data

Participants judged by research staff as independent on performance assessments (Levels 0, 1) did indeed move more frequently than participants judged by research staff as physically dependent (Levels 2, 3, 4) according to movement-monitor data (M = 0.80, SD = .40 and M = 0.28, SD = .43 moves per hour, respectively, t = 7.213, p < .001).

Table 3. Agreement Between Nursing Home Nurse MDS Bed-Mobility Ratings and Research Staff Performance Assessments Defining Agreement as Assistance Levels Match Exactly

<table>
<thead>
<tr>
<th>Research Staff Performance Assessment Assistance Levels</th>
<th>Corresponding MDS Bed-Mobility Self-Performance Ratings</th>
<th>Participants With ≤ 30 Days Between MDS Rating &amp; Performance Assessment (n = 13)</th>
<th>Participants With ≤ 45 Days Between MDS Rating &amp; Performance Assessment (n = 25)</th>
<th>Participants With ≤ 60 Days Between MDS Rating &amp; Performance Assessment (n = 46)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 0: Request only</td>
<td>Rating 0: Independent</td>
<td>.369 (&lt;.001)</td>
<td>.639 (.021)</td>
<td>.505 (.009)</td>
</tr>
<tr>
<td>Level 1: Verbal cues</td>
<td>Rating 1: Supervision</td>
<td>.007 (.918)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Level 2: Manual physical guidance</td>
<td>Rating 2: Limited assistance</td>
<td>.051 (.438)</td>
<td>.435 (.057)</td>
<td>.432 (.009)</td>
</tr>
<tr>
<td>Level 3: Partial physical assistance</td>
<td>Rating 3: Extensive assistance</td>
<td>.135 (.016)</td>
<td>.030 (.913)</td>
<td>.213 (.278)</td>
</tr>
<tr>
<td>Level 4: Unable to perform</td>
<td>Rating 4: Totally dependent</td>
<td>.484 (&lt;.001)</td>
<td>.418 (.125)</td>
<td>.653 (.001)</td>
</tr>
</tbody>
</table>

Note: MDS = Minimum Data Set.

*Participants were evenly spread across assistance levels 0, 2, 3, 4 with n = 4, 3, 3, and 4, respectively.

bUnable to evaluate, too few participants.

cParticipants were spread across assistance levels 0, 2, 3, 4 with n = 4, 6, 6, and 9, respectively.

Table 4. Agreement Between Nursing Home Nurse Generated MDS Bed-Mobility Ratings Compared to Research Staff Turn in Bed Performance Assessments Defining Agreement as Assistance Levels Match Plus or Minus One Category

<table>
<thead>
<tr>
<th>Turn-in-Bed Performance Level Based on Research Staff Assessment (N)</th>
<th>MDS Bed-Mobility Ratinga Agrees, b N (%)</th>
<th>MDS Bed-Mobility Rated More Dependent by &gt;1 Category, N (%)</th>
<th>MDS Bed-Mobility Rated Less Dependent by &gt;1 Category, N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 0: Request only (60)</td>
<td>23 (38)</td>
<td>37 (62)</td>
<td>—</td>
</tr>
<tr>
<td>Level 1: Verbal cues (2)</td>
<td>1 (50)</td>
<td>1 (50)</td>
<td>—</td>
</tr>
<tr>
<td>Level 2: Manual physical guidance (17)</td>
<td>11 (65)</td>
<td>5 (30)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Level 3: Partial physical assistance (20)</td>
<td>18 (90)</td>
<td>—</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Level 4: Unable to perform (98)</td>
<td>84 (86)</td>
<td>—</td>
<td>14 (14)</td>
</tr>
</tbody>
</table>

Note: MDS = Minimum Data Set.

aMDS bed mobility (item G1aA) rated on a 0–4 scale, with 0 = independent, 1 = supervision required, 2 = limited assistance, 3 = extensive assistance, and 4 = totally dependent.

bAgreement was defined as MDS ratings matching the performance assessment levels plus MDS ratings one category lower or higher than the performance assessment level.

Discussion

In this study we evaluated the accuracy of the MDS bed-mobility item. Results showed poor agreement between nursing home nurse MDS bed-mobility ratings and performance-based assessments conducted by research staff. Results also showed that the direction of error was that nursing home nurses systematically overestimated residents’ assistance needs for bed mobility. These results suggest that nursing home nurses do not conduct valid performance assessments when rating bed mobility. It is likely that nursing home nurses also do not conduct valid performance assessments when rating other MDS ADL items (e.g., transfer, walking, and eating). The systematic overestimation of resident

dependency found in this study is consistent with reports from other health care settings wherein caregivers overestimate the dependency level of the care recipient, and with reports from the nursing home setting that staff reinforce resident dependent behavior (Dresselhaus, Luck, Peabody, 2002; Harari, Gurwitz, Avorn, Choodnovskiy, & Minaker, 1994; Lawrence, Tennstedt, & Almy, 1997; Magaziner, Zimmerman, Gruber-Baldini, Hebel, & Fox, 1997; Milisen et al., 2002; Stange et al., 1998). These findings contradict the results of other studies that evaluated MDS accuracy by comparing nursing home nurse ratings with research nurse ratings based on nursing home data sources (e.g., research nurses used staff self-report of assistance provided and medical-record documentation to determine ratings). If nursing home nurses do not perform valid performance-based assessments, then their MDS ratings, medical-record documentation, and verbal reports to research nurses will likely agree yet be inaccurate.

The tendency for nursing home staff to engage in dependency-enhancing care patterns (Baltes et al., 1987; Bowers, Lauring, & Jacobson, 2000; Rogers et al., 2000) and the extra time required of staff to allow resident independence in daily care activities (Schnelle, Sowell, Hu, & Traughber, 1988) may contribute to the bias that nursing home nurses have toward overestimating resident’s physical dependency on the MDS. Thus, the MDS ADL bed-mobility item may reflect what nursing home staff allow the resident to do during daily care provision or the resident’s learned-dependency behaviors rather than what the resident is able to do. This failure to perform valid assessments may explain previously identified discrepancies between standardized assessments of pain, food and fluid intake, and incontinence frequency and MDS data (Cadogan et al., 2004; Pokrywka et al., 1997; Schnelle et al., 2003; Simmons & Reuben, 2000; Simmons et al., 2004). We believe the approach utilized in this study of a standardized performance assessment of residents’ abilities compared with nursing home nurse MDS ratings should be utilized for evaluating accuracy of other MDS ADL items (e.g., walking, transfer ability).

Study limitations that may limit generalizability include small sample size, geographic homogeneity, and nonrandom selection of participants. The small samples in some nursing homes hindered evaluation of facility effects on nursing home nurse MDS errors, and this should be evaluated in the future. The small samples also hindered analyses at different levels for number of elapsed days between MDS ratings and performance assessments. Further, participants were examined on 1 or 2 days and bed mobility was not evaluated over the previous week, as recommended in the MDS user’s manual. However, others have shown that most ADL assessments are stable over 1 and 2 weeks; results of this study showed ADL bed-mobility stability over 2 days (Engle & Graney, 1993; Graney & Engle, 2000). No attempts were made to control for completion date of the MDS assessment, although the most recent MDS assessment was used to compare with performance assessments. We did find some evidence for closer agreement when the MDS and research-staff assessments were performed fewer than 30 days apart and an overall trend for improved agreement with fewer elapsed days between the MDS and research-staff assessments. However, error rates remain high and appear to reflect a systematic overestimation of dependency, regardless of number of elapsed days. This is likely due to the absence of a standardized performance-based assessment that allows nursing home residents to demonstrate ability with minimum staff assistance. The pattern of increased agreement with less time between MDS ratings and performance assessments at the independent level may reflect genuine increasing independence of the residents. However, this is not likely, given that these residents were long-stay residents and not in the facility for rehabilitation. Further, the general course for long-stay nursing home residents is one of gradually increasing dependency. Hence, because all MDS ratings were completed before the performance assessment, the expectation would be for MDS errors to underestimate the dependency level of the residents. These findings should be interpreted cautiously because of the small number of residents with 30 and 45 days or fewer between MDS ratings and performance assessments.

The systematic error rates reported in this study may reflect ambiguous instructions about how to complete the MDS ADL bed-mobility ratings and the failure to provide nursing homes with a specific example of how to implement a performance-based assessment protocol that allows residents to perform tasks at their highest level of independence. That use of a graduated assistance protocol is implied in the instructions for completing the MDS ADL ratings but, based on findings from this study, not likely used by nursing home nurses is an essential factor contributing to the differences between MDS ratings and performance assessments. Further, it is possible that a cumulative measurement effect occurs with the nursing home nurse who completes the MDS, wherein reports or observations of greater levels of dependency achieve more prominence in the mind of the nursing home nurse and thus contribute to the overestimation of bed-mobility dependency. The performance assessment described in this study should be considered for the training of nurses in how to complete the MDS ADL bed-mobility item, because it requires approximately 3 min per resident for ADL bed-mobility assessment, demonstrated stability over 2 days, and had high interrater reliability. This would represent an important improvement in nursing home nurse MDS bed-mobility assessment data. Further, licensed nurse
assessments of a resident’s self-performance of bed mobility using graduated assistance protocols would more accurately identify residents who should be placed on a scheduled repositioning program, which might reduce workload for nursing home staff. Better targeting of residents for scheduled repositioning is important, because repositioning is often part of the care plan for many residents but does not occur even when medical-record documentation indicates that a repositioning program is taking place (Bates-Jensen et al., 2003).

Licensed nurse assessments of a resident’s self-performance of bed-mobility ADLs using graduated assistance protocols could be compared with nurse aide reports of actual assistance provided to the resident, thus providing important data for staff management. For example, a resident who needed no physical assistance for bed-mobility ADLs based on performance assessments using graduated assistance protocols yet who was provided physical assistance by nurse aides indicates problematic care that increases the risk for resident decline in bed-mobility ADL function. The agreement between accurate licensed nurse bed-mobility-ADL resident-performance ratings and actual assistance provided for bed mobility by nurse aides provides a method of monitoring and improving bed-mobility-ADL care processes to enhance residents’ independence and prevent further decline.

References

Received October 22, 2004
Accepted June 10, 2005
Decision Editor: Linda S. Noeker, PhD