American Journal of Alzheimer's Disease & Other Dementias®
2014, Vol. 29(1) 84-89

DOI: 10.1177/1533317513506777
aja.sagepub.com
SAGE

The Bedford Alzheimer Nursing-Severity Scale to Assess Dementia Severity in Advanced Dementia: A Nonparametric Item Response Analysis and a Study of Its Psychometric Characteristics

Francisca Galindo-Garre, PhD[1], Simone A. Hendriks, MD[2], Ladislav Volicer, MD, PhD[3], Martin Smalbrugge, MD, PhD[2], Cees M. P. M. Hertogh, MD, PhD[2], and Jenny T. van der Steen, PhD[2]

Abstract

The Bedford Alzheimer Nursing-Severity Scale (BANS-S) assesses disease severity in patients with advanced Alzheimer's disease. Since Alzheimer is a progressive disease, studying the hierarchy of the items in the scale can be useful to evaluate the progression of the disease. Data from 164 Alzheimer's patients and 186 patients with other dementia were analyzed using the Mokken Scaling Methodology to determine whether respondents can be ordered in the trait dementia severity, and to study whether an ordering between the items exist. The scalability of the scale was evaluated by the H coefficient. Results showed that the BANS-S is a reliable and medium scale ($0.4 \leq H < 0.5$) for the Alzheimer group. All items with the exception of the item about mobility could be ordered. When later item was eliminated from the scale, the H coefficient decreased indicating that the scalability of the scale in the original form is more accurate than in the shorter version. For the other dementia group, the BANS-S did not fit any of the Mokken Scaling models because the scale was not unidimensional. In this group, a shorter version of the scale without the sleeping cycle item and the mobility item has better reliability and scalability properties than the original scale.

Keywords

Alzheimer's disease, Bedford Alzheimer Nursing-Severity scale, Mokken scaling, item response theory

Introduction

Several instruments to assess physical and mental functioning have been developed and validated in nursing research and practice. These instruments can support practitioners in making health care decisions. Two examples are the activities of daily living (ADLs) questionnaire developed by Katz et al[1] and the Mini-Mental State Examination developed by Folstein et al.[2] Two approaches are commonly used to study the reliability and validity of these instruments. Classical test theory is concerned with the estimation of measurement error and the estimation of the true score, and item response theory (IRT) evaluates the responses to individual items. Another alternative approach that is becoming popular in nursing research[3] is the Mokken scaling.[4,5] This scaling methodology follows the principles of IRT for assessing the relationship between items but it requires less rigid assumptions.

One interesting property of IRT models is that items and measured constructs or traits are measured in the same scale. Thanks to this property, items can be ordered along latent trait levels and a hierarchy of symptoms can be established. Hierarchical scales have been useful for measuring a range of constructs for instance, feeding behavior in dementia,[6] distress,[7] or happiness.[8] All these articles used Mokken scaling to determine whether some symptoms are expected to be more frequently observed than other symptoms in the scale.

Further, for an Alzheimer's disease severity scale, assessing the ordering of the items within the scale may be useful. Alzheimer's disease is a progressive disease characterized by limitations in

[1] Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, the Netherlands
[2] Department of General Practice & Elderly Care Medicine, EMGO institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands
[3] School of Aging Studies, University of South Florida, Tampa, FL, USA

Corresponding Author:
Francisca Galindo-Garre, PhD, Department of Epidemiology and Biostatistics, VU University Medical Center, Postbus 7057, 1007MB Amsterdam, the Netherlands.
Email: f.galindo@vumc.nl

cognitive and physical performance.[9] Although the progression is not uniform for patients, the first symptoms are usually cognitive deficits, followed by functional impairments, and finally pathological symptoms.[10] For dementia severity, ordering scales' items implies that the ordering of the items is the same for all patients, irrespective of dementia severity. This means that people with low-dementia severity are expected to have difficulties only with complex items or it is expected that, in general, some problems will appear earlier than others in the dementia disease process.

The Bedford Alzheimer Nursing-Severity Scale (BANS-S) was developed to assess disease severity in patients with advanced Alzheimer's dementia. The scale is based on clinical information about the development of Alzheimer-type dementia. The BANS-S combines measurements of cognitive and functional deficits with the occurrence of other symptoms. It is composed of 7 polytomous items, 2 cognitive items (speech and eye contact), 3 functional items (dressing, eating, and ambulation), and 2 items referring to pathological symptoms (sleep–wake cycle disturbance and muscle rigidity/contractions). The BANS-S total score ranges from 7 to 28, summing the 7 items each ranging from 1 to 4.

The BANS-S has been used extensively in nursing practice and has been quoted in 39 publications (eg, in a large prospective study on advanced dementia in nursing homes by Mitchell et al[10]). The first validation of the current version of the scale[11] showed that the scale is psychometrically strong. Bellelli et al[12] performed a new validation study and they demonstrated that this instrument is valid and that it discriminates between groups of patients with different dementia severity. Volicer et al[13] performed a study on the progression of Alzheimer's dementia with the BANS, which is a previous version of the BANS-S. They estimated dementia duration after which at least 50% of the patients had problems with each BANS item. The patients first had problems with dressing themselves (after 5 years), then sleep–wake cycle dysfunctions (after 6 years), then they lost the ability of feeding themselves and ambulating independently (after 8 years), and finally the ability to keep eye contact (after 12 years). Although this pattern did not apply to all patients because some patients retained some functions despite a long duration, these results indicate a possible hierarchy in the appearance of dementia symptoms.

Establishing the hierarchical properties of the BANS-S provides information additional to the total score obtained by summing patient responses. A scale with hierarchical properties has items that can be ordered according to their mean scores in the total group. Dementia severity is the latent trait assessed by the scale. Patients with a higher dementia severity score are expected to have higher scores in items that are high in the hierarchy than patients with a lower dementia severity score.

The Mokken scaling methods to study the hierarchical properties of a scale with polytomous items are more complex than for a scale with dichotomous items.[14] A set of polytomous items with ordered categories forms a hierarchical scale when the ordering of the items according to their mean score is the same across different values of the latent trait or the measured construct. This property is also named invariant item ordering (IIO). Recently, Ligtvoet et al[15] have developed a method to assess IIO for polytomous items.

The present study assesses the hierarchical properties of the items of the BANS-S using Mokken scaling. First, we assess whether the probability of presenting difficulties with the BANS-S' item scores is higher for patients with higher scores in the trait dementia severity. Then, we use Ligvoet et al[15] method to investigate whether the BANS-S items can reliably be invariantly ordered as severity indicators of dementia. Since the BANS-S was developed for patients with Alzheimer's disease, we first study the subgroup of patients with Alzheimer's disease and then we study whether the ordering of the items found for patients with Alzheimer's disease applies to the group of patients with other types of dementia, because this instrument is often used in research in nursing homes in the United States, Italy, and the Netherlands to assess patients with different types of dementia.[16-22] Finally, we study the ordering of the BANS-S items for the complete scale to investigate whether the BANS-S measures different traits for the different groups.

Methods

Description of the Sample

The data were collected as part of the Dutch End of Life in Dementia study describing quality of dying and end-of-life care and assessing associated factors. We enrolled 372 residents in 28 long-term care facilities upon admission. A comprehensive description of the participants of this study can be found in van der Steen et al.[23] The diagnoses of dementias were based on international guidelines.[24-26]

Description of the Instrument

The BANS-S is a nursing staff-administered questionnaire comprising 7 items with 4 ordered categories. Respondents are evaluated in their ability to perform 3 ADLs ("dressing," "eating" [dependence], and "mobility" [ability to walk independently]), their ability to speak ("speech"), their capacity to maintain eye contact ("eye contact"), the regularity of their sleep–wake cycle ("sleeping"), and the state of their muscles ("muscles"). The item categories have different labels, and they range from 1 to 4. The total score is the sum of the item scores, and it ranges from 7 (no impairment) to 28 (complete impairment).

In our study, the BANS-S was administrated by a nurse or a physician every 6 months. For this analysis, we used the first measurement approximately 8 weeks after admission to the long-term care facility.

Statistical Methods

The R package *Mokken*[27,28] was used to study the hierarchy of the BANS-S instrument. First, we fit the Monotone Homogeneity model (MHM). If the MHM fits, the mean of the latent trait increases as the total score increases,[29] and the sum score can be used to order patients stochastically on the trait in most practical situations.[30] To fit this model, 3 model assumptions are tested, (1) unidimensionality: all items in the instrument measure the same latent trait (the construct dementia severity); (2)

Table 1. Mean Item Scores and Scalability Coefficients (H) With the Standard Errors (SEs) in Parentheses for the BANS-S Items.[a]

	Alzheimer's Dementia (N = 164)		Other Dementias (N = 186)		Complete Group (N = 350)	
Item Label	Mean Scores	H (SE)	Mean Scores	H (SE)	Mean Scores	H (SE)
1. Dressing	2.76	.58 (.04)	2.96	.58 (.03)	2.86	.58 (.03)
2. Sleeping	1.64	.31 (.06)	1.48	.22 (.08)	1.56	.25 (.05)
3. Speech	1.83	.37 (.06)	1.77	.40 (.05)	1.79	.38 (.04)
4. Eating	1.76	.49 (.05)	1.87	.50 (.04)	1.82	.50 (.03)
5. Mobility	1.76	.55 (.04)	2.11	.49 (.04)	1.94	.51 (.03)
6. Muscles	1.82	.50 (.05)	1.93	.42 (.05)	1.88	.45 (.03)
7. Eye contact	1.43	.47 (.05)	1.48	.39 (.07)	1.45	.42 (.04)

Abbreviations: BANS-S, Bedford Alzheimer Nursing-Severity Scale; SE, standard error; MS, Molenaar Sijtsma statistic.
[a] Scale: Alzheimer dementia: H =.47 (.04); reliability MS =.82, Cronbach's α =.81. Other dementias: H =.44 (.04); reliability MS =.81, Cronbach's α =.80.

monotonicity: the probability of choosing a higher category of the item increases with increasing dementia severity; and (3) conditional independence: The responses regarding the same patient to different items are only related to his dementia severity level. Assumptions 1, 2, and 3 can be tested by checking the following restrictions on the scalability coefficients H[28] (Theorem 4.3): the total H coefficient value, the H coefficient for each item, and the H coefficient for each pair of items must be between 0 and 1. The procedures to check these restrictions are the automated-item selection procedure[31] and the item rest score regression. The scalability coefficient H^4 was computed to determine the strength of the relationship of each item with the latent trait. A set of items form a scale if the H coefficient for each pair of items is higher than or equal to .3. Furthermore, scales are classified according to the following criteria for the H value: (1) $.3 \leq H < .4$: weak scale, (2) $.4 \leq H < .5$: medium scale, and (3) $H \geq .5$: strong scale. The unidimensionality assumption was also assessed by exploratory factor analysis, but the results are not reported because they were equivalent to the results obtained with the MHM. The reliability of the scale was checked with the Cronbach's α and the Molenaar Sijtsma statistic (MS), which is a more accurate reliability coefficient. For a description of the properties of these coefficients see van der Ark.[32]

Next, we fitted Double Monotonicity Model (DMM) for polytomous items. This model fits when the previously described assumptions hold, and when the items are ordered among patients. This means that people with a higher dementia severity have a higher probability to experience more difficulties to perform complex activities without help. This—IIO—is a necessary condition for a scale to be hierarchical, and it can be tested by the method of manifest IIO (MIIO).[15] Items involved in violations of the IIO assumption are removed from the questionnaire by the backward method.[32] After IIO was established, the H^T coefficient was calculated to assess the precision of the item ordering.[15] The H^T coefficient was evaluated following the criteria described for the H coefficient.

Results

Of the 372 patients assessed with the BANS-S questionnaire, 350 had completed all the items. Almost half (47%, n = 164) of these patients had Alzheimer's dementia, 22% (n = 77) had vascular dementia, 17% (n = 60) had Alzheimer's and vascular dementia, and 14% (n = 49) had another type of dementia. Since the BANS-S was built for patients with Alzheimer's disease, the psychometric characteristics of 2 groups of patients were studied separately: patients (n = 164) with Alzheimer's disease and the other type of dementia (n = 186) group which includes combinations of Alzheimer's dementia with other dementias.

The MHM

Table 1 shows the mean scores and the scalability coefficients (H) for the BANS-S items computed for the Alzheimer's, the other dementia, and the complete groups. For the Alzheimer's group, the BANS-S scale was a medium scale (H = .47) and had a high reliability according to both reliability coefficients used (MS = .82 and Cronbach's α = .81). The scale was unidimensional and there were no violations in the assumption of monotonicity. Therefore, we can conclude that the MHM model fits for this scale.

For the other dementia group, the scalability and the reliability coefficients were very similar (H = .44, MS = .81, and Cronbach's α = .80) to the coefficients reached by the Alzheimer's group. There was no violation in the monotonicity assumption for both the groups. However, the results from the Mokken's automated-item selection algorithm to check unidimensionality showed that the "sleeping" item did not belong to the same dimension as the other items in the scale. After eliminating this item, the remaining 6 items formed a strong scale with H = .51 (standard error [SE] = .04), and the reliability coefficients for the new scale were MS = .82 and Cronbach's α = .81. Therefore, we cannot conclude that the MHM fits for the complete BANS-S scale for the other dementia group, because the assumption of unidimensionality is violated. Finally, the results for the complete group were close to the results for the other dementia group (H = .45, MS = .82, and Cronbach's α = .80). Again, the "sleeping" belonged to another dimension. The scalability coefficient for the scale without the "sleeping" item was H = .52 (SE = .03), and the reliability coefficients were MS = .83 and Cronbach's α = .82. For both the other dementia and the complete groups, the MHM fits for a 6-item subscale without the "sleeping" item.

Table 2. Mokken Scale of the BANS-S Checked for Violations of Invariant Item Ordering for the Alzheimer Dementia Group (N = 164) and for the Other Dementias Group (N = 186): Mean Item Scores and Scalability Coefficients (H) With the Standard Errors (SEs) in Parentheses.[a]

	Alzheimer Dementia (N = 164)		Other Dementias (N = 186)	
Item Label	Mean Scores (Ordering)	H (SE)	Mean Scores (Ordering)	H (SE)
1. Dressing	2.76 (1)	.54 (.04)	2.96	.57 (.04)
2. Sleeping	1.64 (5)	.29 (.06)		
3. Speech	1.83 (2)	.37 (.06)	1.77	.45 (.06)
4. Eating	1.76 (4)	.45 (.05)	1.87	.54 (.04)
5. Muscles	1.82 (3)	.44 (.05)	1.93	.41 (.06)
6. Eye contact	1.43 (6)	.46 (.05)	1.48	.43 (.07)

Abbreviations: BANS-S, Bedford Alzheimer Nursing-Severity Scale; SE, standard error; MS, Molenaar Sijtsma statistic.
[a] Scale: Alzheimer dementia: H =.42 (.04); reliability MS =.77, Cronbach's α =.76. Other dementia: H =.48 (.04); reliability MS =.79, Cronbach's α =.77.

The DMM

As with MHM, to fit the DMM, the ordering of the items was evaluated for the Alzheimer, the other dementia, and the complete groups. In the Alzheimer's group, 5 items ("mobility," "muscles," "eating," "speech," and "sleeping") were involved in several significant violations of MIIO. The items for which MIIO violations occur do not follow the same ordering by difficulty for all individuals in the population of interest. The backward selection procedure suggested that the item "mobility" should be eliminated from the scale. After removing the "mobility" item, no violations were left. The new scale has a H^T coefficient of .57 that suggests strong support for IIO ($H^T > .5$). This means that the item ordering found has a high accuracy. Table 2 shows the coefficients for the scale after excluding the "mobility" item. Lower mean scores indicate that these deficits appear with higher dementia severity. After adjusting for IIO, the scalability and reliability coefficients for the scale without the mobility item decreased (H = .42, MS = .77, and α = .76). The scalability coefficients for all the items decreased and for the "sleeping" item, it became lower than the cutoff for the H coefficient of .3.These results indicate that, although the "mobility" item cannot be ordered in the hierarchy, the scale should stay in its original form for the group of patients with Alzheimer's disease, because it achieves better values for reliability and scalability in this form.

For the other dementia group, the "sleeping" item was removed from the scale, and the DMM model was fitted for the remaining items. The "mobility," "muscles," "eating," and "speech" items were involved in several significant violations of MIIO. The backward selection procedure also indicated that the "mobility" item should be eliminated from the scale. After removing the "mobility" item, no violations were left, and the new scale had a H^T coefficient of .62. This means that the item ordering found has a high accuracy. After adjusting for IIO, the reliability coefficients for the scale without the "mobility" decreased (MS = .79 and α = .77), but the scalability coefficient increased from H = .44 to H = .48. The scalability coefficients for all the items increased or remained the same.

Finally, we fit the DMM model for the complete group to assess whether the BANS-S measured different traits for the different groups. Four items ("mobility," "muscles," "eating," and "speech") were involved in several significant violations of MIIO. The backward selection procedure indicated that the item mobility should be eliminated from the scale for this group too. After removing the mobility item, no violations were left ($H^T = .59$). The item ordering found for the complete group was very similar to the ordering obtained for the other dementia group.

Discussion

In this article, we have fitted Mokken models to the BANS-S to study its psychometric properties. We found that the BANS-S meets the criteria for an ordinal scale for the patients with Alzheimer's disease. The DMM did not fit well because the "mobility" item could not be accurately ordered in the scale. However, if we remove the "mobility" item from the scale the reliability and the scalability of the scale decrease indicating that the "mobility" item must be retained in the scale.

We found that the BANS-S also meets the criteria for an ordinal scale for other dementias, but the "sleeping" item could not be accurately ordered in the scale. The scale without the "sleeping" item did not fit well with DMM because the "mobility" item could not be accurately ordered in the scale for other dementias. Removing the "mobility" item from the scale increases the scalability of the scale and only slightly decreased the reliability. Our results pointed out that the ordering of the symptoms was different for the patients with Alzheimer's disease compared with the other dementia group but the differences vanished when patients with Alzheimer's disease and other dementia patients were combined.

The reliability of the instrument was already studied for the development population (see Voilcer et al[11]) using classical test theory. They found that a Cronbach's α ranged from .64 to .80, an excellent correlation between raters' score and Spearman correlations higher than .5 with other related test measuring physical functioning, cognitive functioning, speech ability, and dementia progression. In our population, we also studied the reliability of the instrument. We found a Cronbach's α of .81 for the Alzheimer's group and .80 for the other dementia group.

The range of the mean scores suggests that the items can discriminate between patients with different degrees of dementia. These findings confirm the results reported in Bellelli et al.[12]

We found that patients with Alzheimer's disease had the highest mean score for the dressing item and the lowest for the eye contact item. Volicer et al[13] also found that patients with a short dementia duration often have problems with dressing themselves and that a high proportion of patients could keep eye contact 12 years after diagnosis.

For both the Alzheimer's and the other dementia groups, the "mobility" item could not be ordered in the dementia intensity scale. This means that the scores for this item do not have the same ordering for all the values of the latent trait. The reason may be that not only dementia but also other diseases such as stroke, arthritis or the effects of a fall may affect a person's ability to walk independently.

The results differed between the group of patients with Alzheimer's disease and the group of other dementias. The last group comprised patients who had vascular dementia, a combination of vascular dementia and Alzheimer's dementia or other types of dementias and, therefore, this group was more heterogeneous. The scale was not unidimensional for the other dementia group, because the "sleeping" belonged to a different dimension. Problems with the "sleeping" item were already reported by van der Steen et al.[20] Furthermore, a lower mean score for the other dementia group in Table 1 suggest that people with other types of dementia had sleeping problems less often than patients with Alzheimer's disease. The proportion of patients who report an irregular sleep–wake cycle in the other dementia group was 44% versus 56% of the patients in the Alzheimer's group. Although an irregular sleep–wake rhythm is a symptom that may occur for all dementia types, differences in sleep symptoms and signs may vary according to the dementia (sub) type.[33] Sleep disturbances may occur more frequently and in an earlier stage of the Alzheimer's disease in comparison with other dementia types. In a population of patients with autopsy-confirmed Alzheimer's disease, a unique profile of disordered activity was found when compared to those with other neurodegenerative dementias. The hypothesized mechanism of circadian rhythm disturbance includes damage to the suprachiasmatic nucleus, circadian pacemaker damage, and alterations in pineal gland function and melatonin secretion.[33]

Another difference in the results for the Alzheimer's and the other dementia groups was that the place of the item "speech" in the hierarchy was different. This result is difficult to interplet clinically because the moment in the course of the dementia in which this item is affected may vary between type of dementia.[34] For example, speech is often affected early in frontotemporal dementia[35,36] while it may be a later symptom in Alzheimer's disease.[37-39] However, whether this symptom is affected in vascular dementia or not depends on the location of the lesion.[40]

The present study has some limitations that warrant comment. First, this study is based on cross-sectional analyses limited to the first measurement of a longitudinal study. Further work may replicate the analyses for measurements obtained later after admission to study, to investigate whether the relationships between the items change, and to study individual disease progression. Second, we had no external criterion against which to evaluate the responsiveness of the scale to clinical changes. Third, we could not explain associations between mobility and comorbidity, because we do not know if the mobility problems were caused by the dementia or by other diseases. Finally, the differentiation between dementia types was mostly based on clinical findings, which may not always correlate with neuropathological evaluation.[41]

Determining IIO gives a clear meaning to test scores because we learn about the ordering of the problems. The probability of having problems with an item with a higher mean score (higher in the hierarchy) was higher for patients with high-dementia severity than for people with low-dementia severity. This result is relevant because many scales do not discriminate between patients with more severe dementia. However, this scale may present a floor effect for patients with lower levels of dementia, because they did not have difficulties with most of the items. This was not the case because only 26% of the patients has a sum score of $\leq$9). Furthermore, it should be also taken into account that the data were from baseline measurements and that the patient population at this point was not always severely demented. Further research should be done to study whether the dementia patterns found for this population apply to the course of the dementia for an individual and to evaluate the responsiveness of the scale to individual changes.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by a career award for JTS of the Netherlands Organisation for Scientific Research (NWO, the Hague; Veni 916.66.073), and by the Netherlands organisation for health research and development (ZonMw, the Hague; grant number 1151.0001).

References

1. Katz S, Ford A, Moskowitz R, Jackson B, Jaffe M. Studies of illness in the aged. *JAMA*. 1963;185:914-919.
2. Folstein M, Folstein S, McHugh P. MiniMental State: a practical method for grading the cognitive state of patients for the clinitian. *J Psychiatr Res*. 1975;12(3):189-198.
3. Watson R, van der Ark A, Lin L, Fieo R, Deary I, Meijer R. Item response theory: how Mokken scaling can be used in clinical practice. *J Clin Nurs*. 2011;21(19-20):2736-2749.
4. Mokken R. *A Theory and Procedure of Scale Analysis with Applications in Political Research*. New York: De Gruyter; 1971.
5. Molenaar I. Nonparametric Models for Polytomous Responses. In: van der Linden W, Hambleton R, eds. *Handbook of Modern Item Response Theory*. New York: Springer-Verlag; 1997; 367-380.
6. Watson R. Mokken scaling procedure (MSP) applied to feeding difficulty in elderly people with dementia. *Int J Nurs Stud*. 1996;33(4):385-393.

7. Watson R, Deary I, Shipley B. A hierarchy of distress: mokken scaling of the GHQ-30. *Psychol Med*. 2008;38(4):575-579.
8. Stewart M, Watson R, Clark A, Ebmeier K, Deary I. A hierarchy of happiness? Mokken scaling analysis of the oxford happiness inventory. *Pers Indiv Differ*. 2010;48(7):845-848.
9. Bossers W, van der Woude L, Boersma F, Scherder E, van Heuvelen M. Recommended measures for the assessment of cognitive and physical performance in older patients with dementia: a systematic review. *Dement Geriatr Cogn Dis Extra*. 2012;2(1): 589-609.
10. Mitchell SL, Teno JM, Kiely DK, et al. The clinical course of advance dementia. *N Engl J Med*. 2009;361(16):1529-1538.
11. Volicer L, Hurley A, Lathi D, Kowall N. Measurement of severity in advanced Alzheimer's disease. *J Gerontol*. 1994;49(5): 223-226.
12. Bellelli G, Frisoni G, Bianchetti A, Trabucchi M. The bedford Alzheimer nursing severity scale for the severely demented: validation study. *Alzheimer Dis Assoc Disord*. 1997;11(2):71-77.
13. Volicer L, Seltzer B, Rheaume Y, et al. Progression of Alzheimer-type dementia in institutionalized patients: a cross-sectional study. *J Appl Gerontol*. 1987;6(1):83-94.
14. Sijtsma K, Meijer R, van der Ark LA. Mokken scale analysis as time goes by: an update for scaling practitioners. *Pers Indiv Differ*. 2011;50(1):31-37.
15. Ligtvoet R, van der Ark LA, te Marvelde JM, Sijtsma K. Investigating an invariant item ordering for polytomously scored items. *Educ Psychol Meas*. 2010;70(4):578-595.
16. Brandeis GH, Berlowitz DR, Coughlin N. Mortality associated with an influenza outbreak on a dementia care unit. *Alzheimer Dis Assoc Disord*. 1998;12(3):140-145.
17. Frisoni GB, Gozzetti A, Bignamini V, et al. Special care units for dementia in nursing homes: a controlled study of effectiveness. *Arch Gerontol Geriatr*. 1998;26(suppl 1):215-224.
18. Camberg L, Woods P, Ooi WL, et al. Evaluation of simulated presence: a personalized approach to enhance well-being in persons with Alzheimer's disease. *J Am Geriatr Soc*. 1999;47(4): 446-452.
19. Colombo M, Vitali S, Cairati M, et al. Wanderers: features, findings, issues. *Arch Gerontol Geriatr Suppl*. 2001;7:99-106.
20. van der Steen JT, Ooms ME, van der Wal G, Ribbe MW. Pneumonia: the demented patient's best friend? discomfort after starting or withholding antibiotic treatment. *J Am Geriatr Soc*. 2002; 50(10):1681-1688.
21. Van der Steen JT, Ooms ME, David R, van der Wal G, Ribbe MW. Severe dementia and adverse outcomes of nursing home-acquired pneumonia: evidence for mediation by functional and pathophysiological decline. *J Am Geriatr Soc*. 2002;50(3):439-448.
22. Pasman HR, Onwuteaka-Philipsen BD, Kriegsman DM, et al. Discomfort in nursing home patients with severe dementia in whom artificial nutrition and hydration is forgone. *Arch Intern Med*. 2005;165(15):1729-1735.
23. Van der Steen JT, Ribbe MW, Deliens L, Gutschow G, Onwuteaka-Philipsen BD. Retrospective and prospective data collection compared in the dutch end of life in dementia (DEOLD) study [published online April 28, 2013]. *Alzheimer Dis Assoc Disord*. 2013.
24. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the national institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):263-269
25. Roman GC, Tatemichi TK, Erkinjunti T, et al. Vascular dementia. diagnostic criteria for research studies: report of the NINDS-AIREN international workshop. *Neurology*. 1993;43(2):250-260.
26. Mckeith IG, Galasko D, Kosaka K, et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): Report of the consortium on DLB international workshop. *Neurology*. 1996;47(5):1113-1124.
27. Van der Ark LA. Mokken scale analysis in R. *J Stat Software*. 2007;20(11):1-19.
28. Sijtsma K, Molenaar I. *Introduction to Nonparametric Item Response Theory*. Thousand Oaks, CA: Sage; 2002.
29. Sijsma K, Emons W, Bouwmeester, Nyklícek I, Roorda LD. Nonparametric IRT analysis of quality-of-life scales and its application to the World Health Organization quality-of-life scale (WHOQOL-Bref). *Qual Life Res*. 2008;17(2):275-290.
30. Van der Ark LA. Stochastic ordering of the latent trait by the sum score under various polytomous IRT models. *Psychometrika*. 2005;70(2):283-304.
31. Molenaar I, Sijtsma K. *MSP5 for Windows. A Program for Mokken Scale Analysis for Polytomous Items*. Groningen: ProGamma; 2000.
32. Van der Ark LA. New developments in Mokken scale analysis. *J Stat Software*. 2013;48(5):1-27.
33. Zhou QP, Jung L, Richards KC. The management of sleep and circadian disturbance in patients with dementia. *Curr Neurol Neurosci Rep*. 2012;12(2):193-204.
34. Duffy JR, Josephs KA. The diagnosis and understanding of apraxia of speech: why including neurodegenerative etiologies may be important. *J Speech Lang Hear Res*. 2012;55(5):S1518-S1522.
35. Amici S, Gorno-Tempini ML, Ogar JM, Dronkers NF, Miller BL. An overview on primary progressive aphasia and its variants. *Behav Neurol*. 2006;17(2):77-87.
36. Kirshner HS. Primary progressive aphasia and Alzheimer's disease: brief history, recent evidence. *Curr Neurol Neurosci Rep*. 2012;12(6):709-714.
37. Nelissen N, Dupont P, Vandenbulcke M, Tousseyn T, Peeters R, Vandenberghe R. Right hemisphere recruitment during language processing in frontotemporal lobar degeneration and Alheimer's disease. *J Mol Neurosci*. 2011;45(3):637-674.
38. Cera ML, Ortiz KZ, Betolucci PH, Minett TS. Speech and orofacial apraxias in Alzheimer's disease. *Int Psychogeriatr*. 2013; 25(10):1679-1685.
39. Emery VO. Language impairment in dementia of the Alzheimer type: a hierarchical decline? *Int J Psychiatry Med*. 2000;30(2): 145-164.
40. Vuorinen E, Laine M, Rinne J. Common pattern of language impairment in vascular dementia and in Alzheimer disease. *Alzheimer Dis Assoc Disord*. 2000;14(2):81-86.
41. Harper DG, Stopa EG, Kuo-Leblanc V, et al. Dorsomedial SCN neuronal subpopulations subserve different functions in human dementia. *Brain*. 2008;131(pt 6):1609-1617.