Original Investigation

Use of the Barthel Index, mini mental status examination and discharge status to evaluate a special dementia care unit

A. Brühl[1], J. Hoffmann[2], U. Sottong[3], M. Isfort[4], D. Tucman[4]

[1] Department of Statistics and Standardised Methods in Nursing Science, Vallendar University of Philosophy and Theology (PTHV), Faculty of Nursing Science, Germany; [2] Malteser Hospital, St. Hildegard, Cologne, Germany; [3] Special Dementia Care Unit, Malteser Hospital, St. Hildegard, Cologne, Germany; [4] German Institute of Applied Nursing Research, Cologne, Germany

Background: when dementia patients are grouped according to their ordinal Barthel Index and Mini Mental Status Examination sum scores, it is not clear which portions of the data should be used when valid classifications are developed. Criteria used for classification of patients must be stochastically independent.
Objectives: the relationship between Barthel-Index, Mini-Mental State Examination sum scores and discharge status was investigated to develop subgroups of dementia patients. The developed classification uses stochastically independent information and can be used to evaluate special care units.
Methods: we used an unrestricted partial credit model to assess the possible scores on the Barthel-Index. We investigated the individual effects of items on the Barthel sum score by using non-parametric conditional-inference-regression trees. The relationships between Barthel score, Mini Mental Status score, and discharge status, in terms of classifying the dementia patients into subgroups, were investigated using a latent class analysis.
Results: an interval scale Barthel-Index did not yield a significant improvement versus the ordinal Barthel-Index sum score. Differences in Barthel-Index were meaningful only in the context of three groups using four items. A classification of dementia patients in latent classes could be developed using three Barthel – and Mini Mental Status-Groups and the discharge status of patients who were living at home before admission. Three Barthel – and Mini Mental Status Groups can be combined with the discharge status of those patients who live at home before admission. A combination of a high Barthel – with a low Mini-Mental-Status-group has the highest probability to live no longer at home after discharge.
Discussion: the relative frequency of living at home after discharge in different Barthel and Mini Mental Status subgroups could be compared between different acute hospitals as an indicator of service quality for dementia patients. A high risk group is identified by a combination of a high Barthel- and a low Mini Mental Status Examination Group.

Key words: Special Dementia Care Unit, Evaluation, Classification of Dementia Patients, Measurement, Barthel-Index, MMSE

INTRODUCTION

Several studies have demonstrated that the Barthel Index (BI) sum score is not reliable, because it represents neither interval nor uni-dimensional data [1 2 3]. Furthermore, the simple 'raw data' of each BI item differs markedly in terms of weighting. Therefore, as highlighted by current research, use of the ordinal total BI sum score is problematic.

Barer et al. [4] reduced the number of answer options for each BI item to two, resulting in a clear hierarchy among the 10 items. Sum scores can then be interpreted as

Received: February 23, 2017 - Accepted: June 13, 2017

Correspondence: Albert Brühl, PTHV, Department of Statistics and Standardised Methods in Nursing Science, Vallendar University of Philosophy and Theology (PTHV), Faculty of Nursing Science, University of Cologne, Pallottistrasse 3, 56179 Vallendar – Tel.: +49 2616402503 - Fax: +49 2616402120 - E-mail: abruehl@pthv.de

representative of a certain pattern of indicators. Several researchers have attempted to develop different versions of the BI with fewer items [5 6]. Others have focused on differences in BI item weights [7].

We changed the weighting of the items as per Table I, and then examined the overall fit of a partial credit model. The maximum modified raw BI score is 20.

MATERIAL AND METHODS

We investigated whether BI scores could be combined with Mini-Mental State Examination (MMSE) scores and discharge status to categorise patients in a latent class model. By using these data in combination, a hospital may easily ascertain the percentage of patients who no longer live alone after discharge, even if they lived alone pre-admission; in combination with being in the highest BI and low MMSE groups, such dementia patients might be at risk of poor hospital service. Variability in discharge status was compared among different subgroups of patients (classified according to BI and MMSE scores).

In total, 214 patients in the special dementia care unit of Malteser Hospital, St. Hildegarde, Cologne, Germany between 2014 and 2015 were included in the study. These patients had a primary or secondary diagnosis of dementia and an MMSE sum score below 28. Patients without dementia and those with a MMSE sum score above 27 on discharge were excluded. In terms of patient demographics, 25% of the patients were male and 75% were female; 25% were aged 64-78 years, 25% were 79-84 years, 25% were 85-88 years, and 25% were 89-98 years.

The BI was administered by nursing staff upon both admission and discharge. We used discharge BI scores in analyses.

Rasch analysis

We used a partial credit model to analyse data related to items with more than two answer options. This model extends on the probabilistic one-parameter logistic Rasch model discussed by Rost [8] and Strobl [9], in which the probability of providing a certain answer to a given item depends on the relationship between the item's difficulty and the ability of the responder.

In a Rasch model, the probability that a person with a certain level of ability will score positively on an ADL item is a logistic function of the difference between that person's ability (i.e. their BI sum score) and the difficulty of the ADL to which the item pertains (e.g., the 'toilet' ADL). If the ability of a person is equal to the difficulty of an ADL, the probability of success is 0.5. When the person's ability exceeds the difficulty of the ADL, the probability that the ADL will be completed successfully increases.

A good global Rasch model fit is indicated by a total-item chi-square probability greater than 0.05.

The Rasch model requires that all BI items are measured using a single underlying construct. All items that are summarised by a sum value are assumed to be one-dimensional. Another assumption is that the items are locally independent; i.e. the answer to one item should not determine the answer to any other item. We tested whether the model residuals were associated, where local dependence is indicated by residuals being highly positively correlated. Furthermore, we analysed the loadings of items in a principal components analysis of the residuals.

We used a pairwise algorithm to estimate Rasch model parameters using the RUMM2020 and RUMM2030 software packages (RUMM Laboratory Pty Ltd., Duncraig, Australia).

The overall model fit was evaluated using the chi-square test of model fit (item-trait interaction). Individual item fit was tested using the item residuals, calculated as the difference between the observed and the expected values for each item. According to Andrich [10], item residuals should be between -2 and 2. We used the person separation index to assess the extent to which the model distinguished between respondents according to their ability.

We also assessed differences in item functioning (DIF) for subgroups defined according to cognitive status (MMSE scores below 20 *versus* MMSE scores 20 or higher). An additional consideration was the answer option thresholds: if ordinal items are changed to an interval scale, the correct answer option thresholds must also be preserved.

To calculate sample size, we referred to Linacre [11], who proposed a minimum sample size of 108 cases for precision to 0.5 logits, and a confidence interval of 99%, for tests with less than 30 items during Rasch model testing. We also used the RUMM2030 software for the power calculation.

Regression trees

Regression tree algorithms are commonly used and have been applied in recursive partitioning by Strobl, Malley, and Tutz [12]. This technique involves dividing a sample such that the resulting subgroups, with respect to the value of the dependent variable, exhibit maximal in-group homogeneity and maximal between-group heterogeneity. The conditional inference trees of Hothorn, Hornik, and Zeileis [13] represent an alternative to the least squares method, and involve significance tests based on permutations of the items. An exhaustive search was conducted during partitioning of the

Table I. Scoring of Barthel Index items: original versus modified version. All values represent points scored.

Barthel item	Unable to perform	Able to perform with assistance	Able to perform independently
1. Feeding (original)	0	5	10
Feeding (modified)	0	1	2
2. Transfer (chair/bed) (original)	0	5-10	15
Transfer (chair/bed) (modified)	0	1-2	3
3. Grooming (original)	0	0	5
Grooming (modified)	0	0	1
4. Toilet (original)	0	5	10
Toilet (modified)	0	1	2
5. Bathing (original)	0	0	5
Bathing (modified)	0	0	1
6. Walking (original)	0	5-10*	15
Walking (modified)	0	1-2	3
7. Stairs (original)	0	5	10
Stairs (modified)	0	1	2
8. Dressing (original)	0	5	10
Dressing (modified)	0	1	2
9. Bowels (original)	0	5	10
Bowels (modified)	0	1	2
10. Bladder (original)	0	5	10
Bladder (modified)	0	1	2

* With walker/wheelchair

sample, and the selection for the next split was based on items with the highest significance.

The C-Tree algorithm avoids overfitting of the model to the data. We used those BI items that were detected in a regression analysis of the original BI sum score and the BI groups classified by the regression trees. The R Party software package was used for the analysis [14].

Latent Class Analysis (LCA)

The latent class model is one example of mixed distribution or 'mixture' models, which are designed to detect 'unobserved heterogeneity' [15] within a population and classify the data according to a latent variable to derive meaningful groups. It is important to note that this latent variable is not used to structure the data; rather it is constructed only during the process of evaluation [8]. The LCA was used to test whether BI and MMSE scores, as well as discharge status, are useful for differentiating dementia patients into subgroups.

The LCA operates according to the conditional assumption that a person belongs to a certain subgroup when a specific pattern of criteria are met. For the model calculation, four assumptions of the LCA are important [8]: 1. The probability of meeting a criterion is constant among all patients within a group; 2. There is local stochastic independence of criteria within a group; 3. The groups are disjoint and exhaustive; and 4. Homogeneity prevails within a group item.

We used likelihood ratio tests to compare the true likelihood of a particular response pattern with that estimated by the model. Ideally, the ratio of these two values is 1. The double-negative logarithm of the likelihood ratio is chi-square distributed, and can therefore be tested for significance. After applying the model, the chi-square test should be associated with a high P-value. We also used the dissimilarity index (DI), which has a value between 0 and 1 and describes the proportion of patients that must be re-classified to perfectly reproduce the actually observed frequencies. The DI should be as close to 0 as possible. We also used the bivariate residuals to test whether the criteria used to generate the latent classes were stochastically independent. We used Latent GOLD software (ver. 4.5; Statistical Innovations, Belmont, MA, USA) to perform this analysis [16].

RESULTS

Testing the Partial Credit Model

The chi-square statistic of the degree of model fit (item-trait interaction) was not significant (27.8, 20 degrees of freedom; model probability, 11%). The person separation index for the BI items was high (0.87), demonstrating that the model had excellent reliability.

All BI items exhibited a good fit with the partial credit model, with fit residuals of between -2.0 and 2.5. The chi-square statistics of item fit were not significant (all $p > 0.05$). Items were ordered by threshold.

No differential item functioning (DIF) was observed for the subgroups defined according to the MMSE.

Although a partial credit model could be fitted to the data, the mean sum score of the recalibrated BI, of 54 points or -0.323 logits, could be produced by different scoring patterns across the 10 items (where 0 = *unable to perform activity*; 1/2 = *can perform with assistance*, and 2/3 = *can perform independently*.

Table II shows that there is a high degree of variance in dependency within the mean BI sum score.

There were three clearly dissociable BI-groups (Fig. 1):

- a small 'low' scoring group composed of patients who cannot use the toilet, nor perform any transfer actions, without assistance;
- a large 'middle' scoring group characterised by six different patterns of response across the 'toilet', 'transfer' and 'dressing' items; and
- a small 'high' scoring group composed of patients who can use the toilet, dress themselves, and climb stairs assisted or unassisted.

By combining the BI groups and MMSE groups (MMSE: 'low' = score of 0-16, 'middle' = score of 17-21; and 'high' = score > 21), it could be seen that patients with scores in the 'high' category on both the BI and the MMSE had a 93% probability of being able to live alone after discharge. Among patients scoring in the 'middle' category on both instruments, that probability was reduced to 50%.

The three of BI- and MMSE-groups, in addition to the discharge status of patients who were admitted to the hospital after living alone in their own homes, are stochastically independent criteria that can inform subgroups of dementia patients in acute care hospitals. The LCA, which included the three groups of BI and MMSE as well as the discharge status, had a DI of 0.06 and a chi-square probability value of 0.76; furthermore, none of the bivariate residual values, for any combination of the three criteria, exceeded 1.0. Therefore, the derived latent classes represent a good classification scheme; such classification is not possible with the original (ordinal) or revised (interval scale) BI sum scores (Fig. 2).

Table II. Response patterns resulting in the average Barthel Index (BI) sum score: in total, 82% of the change in the revised BI was explained by a combination of up to four variables.
The change in BI groups between admission and discharge was mainly caused by changes from the middle BI group to the high BI group. The three groups differed significantly in both their ordinal and interval BI sum scores.

Av-erage Barthel-logit-score	Transfer	Feeding	Bowels	Toilet	Mobility	Bladder	Groo-ming	Dressing	Stairs	Bathing
-0.323	2	1	2	1	1	2	1	1	0	0
-0.323	2	2	2	1	1	1	1	1	0	0
-0.323	2	2	1	1	2	1	1	1	0	0
-0.323	2	2	1	1	2	1	1	1	0	0
-0.323	2	1	2	1	2	1	1	1	0	0
-0.323	2	2	2	1	1	2	0	1	0	0
-0.323	2	2	2	1	1	2	0	1	0	0
-0.323	2	1	2	1	2	1	1	1	0	0
-0.323	2	2	0	1	3	0	1	2	0	0
-0.323	2	1	1	1	2	2	0	1	1	0
-0.323	2	2	2	1	1	2	0	1	0	0
-0.323	2	1	2	1	2	1	1	1	0	0
-0.323	2	2	1	1	2	1	1	1	0	0
-0.323	3	2	1	1	1	1	1	1	0	0
-0.323	2	1	2	1	2	0	1	1	1	0
-0.323	2	1	2	1	1	1	1	1	1	0
-0.323	3	2	1	1	2	0	1	1	0	0
-0.323	3	1	2	1	2	1	0	1	0	0
-0.323	2	1	2	1	2	1	0	1	1	0
-0.323	2	1	2	1	2	1	1	1	0	0
-0.323	3	1	1	1	3	1	0	1	0	0

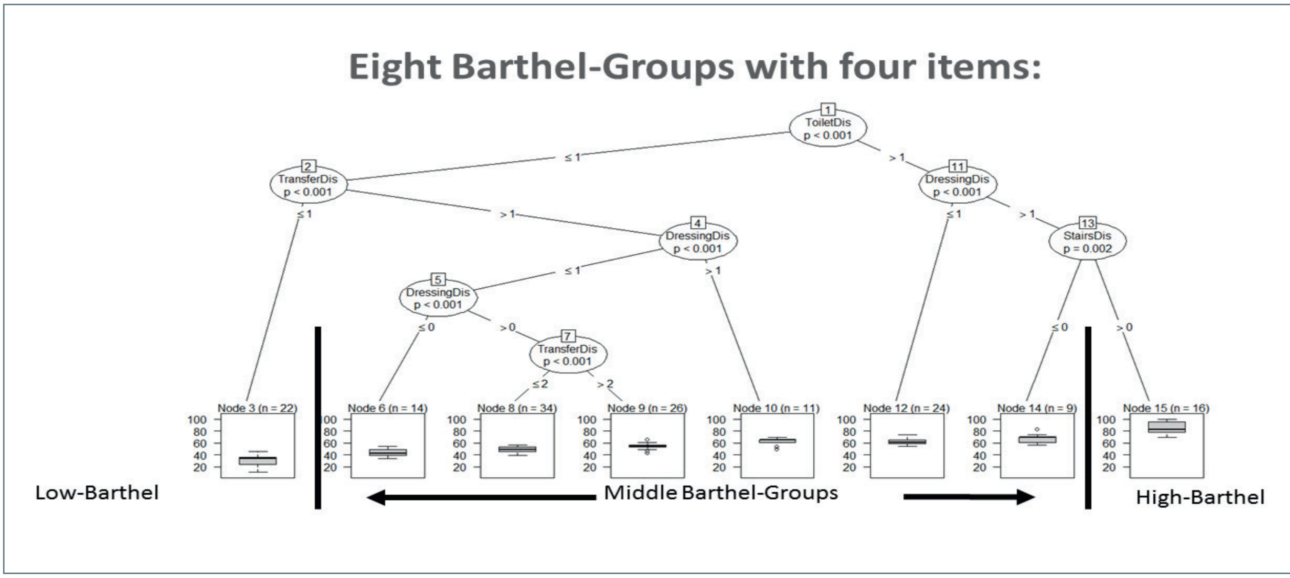

Figure 1. Dementia patients grouped according to scores on the four significant BI items in a regression tree.

April 2016, N= 197		Cluster 1	Cluster 2	Cluster 3
Cluster Size		0.5623	0,3327	0,1051
MMSE				
	low	0.1265	0.7053	0.4923
	middle	0.3778	0.2543	0.3792
	high	0.4958	0.0403	0.1285
Three Barthel-Groups				
	Barthel low	0.0807	0.0611	0.0002
	Barthel middle	0.7694	0.9351	0.5657
	Barthel high	0.2299	0.0039	0.4341
Discharge of patients living at home before admission				
	home	0.9979	0.8224	0.4395
	Assisted living	0.0021	0.1285	0.1738
	Short term care	0.0000	0.0075	0.0256
	Long term care	0.0000	0.0417	0.3811

Figure 2.

DISCUSSION

The results of the Rasch model revealed that the revised (interval scale) BI sum score did not lead to much improvement over the original (ordinal) BI sum score. Changes in BI sum scores among dementia patients during a stay in an acute care hospital depended mainly only on four items: 'toilet', 'transfer', 'dressing,' and 'stairs'.

Patients who were living at home pre-admission, and had 'high' scores on both the BI and MMSE, had a probability of being able to live alone again after discharge of nearly 1. This probability fell to 0.5 in patients being a member of the high Barthel and the low MMSE group.

Classification of patients according to original (ordinal) BI score, revised (interval scale) BI score, MMSE score, or discharge status was not possible in an LCA because these measures are stochastically dependent. The LCA demonstrated that use of these BI groups based on combinations of individual items represents a good classification scheme for dementia patients in acute hospitals. Furthermore, these data are routinely collected during normal practice.

The rate of those patients admitted from home and living alone after discharge who are members of the high Barthel-Group (independent toilet, transfer, dressing and partly stairs) and of the low MMSE-group (< 17 points) is a quality criterion for special care units. The rate of those patients living at home who were admitted from home who are members of the low Barthel and the high MMSE Group should be compared between different hopitals.

Grouping of patients with four Barthel-Items and three MMSE score groups identifies a high-risk group for losing the opportunity of living at home. By identifying these subgroup of patients a quality indicator for special care units could be developed, when the rates of those living at home after discharge who were admitted from home are cpmpared.

Acknowledgements

The study was financed by the porticus foundation, Düsseldorf.

References

[1] Sainsbury A, Seebass G, Bansal A, et al. *Reliability of the Barthel Index when used with older people.* Age Ageing 2005;34:228-32.

[2] Tennant A, Geddes JML, Chamberlain MA. *The Barthel Index: an ordinal score or interval level measure?* Clinical Rehab 1996;10:301-8.

[3] de Morton N, Keating J, Davidson M. *Rasch analysis of the Barthel Index in the assessment of hospitalized older patients after admission for an acute medical condition.* Arc Phys Med ad Rehab 2008;89:641-7.

[4] Barer DH, Murphy JJ. *Scaling the Barthel: a 10 point hierarchical version of the ADL index for use with stroke patients*. Clinical Rehab 1993;7:271-7.

[5] Hobart J, Thompson A. *The five item Barthel Index*. J Neurol Neurosurg Psychiatry 2001;71:225-30.

[6] Ellul J, Watkins C, Barer D. *Estimating total Barthel scores from just three items: the European Stroke database 'minimum dataset' for assessing functional status at discharge from hospital.* Age Ageing 1998;27:115-22.

[7] van Hartingsveld F, Lucas C, Kwakkel G, et al. *Improved interpretation of stroke trial results using empirical Barthel item weights.* Stroke 2006;37:162-6.

[8] Rost J. *Testheorie und Testkonstruktion*. 2nd ed. Bern: Huber 2004.

[9] Strobl C. *Das Rasch-Modell*. München: Rainer Hampp 2010.

[10] Andrich D, Sheridan B, Luo G. *Rasch Unidimensional Measurement Models (RUMM). Interpreting RUMM2020 Part I Dichotomous Data. Version 2020*: Rumm Laboratory Pty Ltd, 2004.

[11] Linacre JM. *Sample size and item calibration stability*. Rasch Meas Trans 1994;7:S328.

[12] Strobl C, Malley J, Tutz G. *An introduction to recursive partitioning*. Technical Rep N. 55. Department of Statistics, University of Munich, 2009.

[13] Hothorn T, Hornik K, Zeileis A. *Unbiased recursive partitioning: a conditional inference framework.* J Computat Graph Statistics 2006;15:651-74.

[14] Torsten H, Hornik K, Strobl C, et al. *A Laboratory for Recursive Partitioning*. Package Party 2013 (http://cran.r-project.org/web/packages/party/party.pdf).

[15] Nylund K, Asparouhov T, Muthén B. *Deciding on the number of classes in latent class analysis and growth mixture modeling*. *A Monte Carlo Simulation study*. Structural Equation Modeling 2007;14:535-69.

[16] Magidson J, Vermunt JK. *Latent GOLD 4.5.* 2009 (http://www.statisticalinnovations.com/latent-gold-5-1/, zuletzt geprüft am 05.08.2016).