Routine functional assessment for hip fracture patients

Are there sufficient predictive properties for subgroup identification in treatment and rehabilitation?

Tonny J PEDERSEN [1,2] and Jens M LAURITSEN [3,4]

[1] Rehabilitation Department, OUH Odense University Hospital, Svendborg Hospital; [2] Institute of Public Health, University of Southern Denmark; [3] Orthopaedic Department, OUH Odense University Hospital; [4] Institute of Clinical Research, University of Southern Denmark, Odense, Denmark.
Correspondence: tonny.jaeger.pedersen@rsyd.dk
Submitted 2016-01-06. Accepted 2016-03-03.

Background and purpose — Pre-fracture functional level has been shown to be a consistent predictor of rehabilitation outcomes in older hip fracture patients. We validated 4 overall pre-fracture functional level assessment instruments in patients aged 65 or more, used the prediction of outcome at 4 months post-fracture, and assessed cutoff values for decision making in treatment and rehabilitation.

Patients and methods — 165 consecutive patients with acute primary hip fracture were prospectively included in the study. Pre-fracture Barthel-20, Barthel-100, cumulated ambulation score, and new mobility score were scored immediately after admission. Outcome defined as mortality, residential status, and independent walking ability was assessed at 4 months.

Results — 3 of the assessment instruments, namely Barthel-20, Barthel-100, and new mobility score, correlated with outcome at 4 months post-fracture and were valid predictors. Thresholds were estimated. We found no evidence that Barthel-100, with its finer granularity, performs better than Barthel-20 as a predictor.

Interpretation — Our findings indicate that pre-fracture scores of Barthel-20 and new mobility score have predictive ability, and further investigation of usage for guidance of clinical and rehabilitation decisions concerning hip fracture patients is warranted.

The elderly hip fracture population represents the whole spectrum of functional levels, from those confined to bed to those who are active and living independently. Targeted treatment and rehabilitation regimens must therefore be based on assessment of the functional level of each individual.

There are 2 strands of legal responsibility for rehabilitation in Denmark. The local health region (the hospital) is responsible for the first, short inpatient period (mean 8 days, according to Statistics Denmark for 2014). A report, including a brief summary of status at discharge, is electronically sent to the municipality for initiation of a targeted, longer period of general rehabilitation. The length of municipality-based rehabilitation depends on the specific municipal decision, and could for example be 2–3 months twice a week, often administered as a combination of individual and group training sessions (Kronborg et al. 2015). Due to the divided responsibility for rehabilitation, it is crucial to establish efficient cross-sectoral collaboration and communication.

To develop closer collaboration between hospitals and municipalities for hip fracture patients and older people in general in our county, a measurement system was established in the late 1990s. It defined 3 modes of assessment to be performed: (1) patient-reported quality of life assessment (based on EuroQuol EQ-5D (Wittrup-Jensen et al. 2009)); (2) overall assessment of independent function based on all available information from patients, relatives, and other professionals—and scored in structured functional indices (Barthel index (Mahoney and Barthel 1965)); and (3) reproducible performance-based tests of physical function, e.g. 30-second chair-stand test (Jones et al. 1999), and of cognitive level, e.g. orientation memory concentration test (hereafter called OMC) (Wade and Vergis 1999). Written manuals and a DVD with video instruction material were developed (Lauritsen 2007), and staff were trained to understand the paradigm and perform the performance-based tests.

The specific appropriateness of such instruments for assessment has been the subject of much debate in the Danish Hip Fracture Quality Assurance Program (Danish_Interdisciplinary_Register_for_Hip_Fracture 2015). The program initially included the new mobility score (hereafter called NMS) (Parker and Palmer 1993) and the cumulated ambulation score (hereafter called CAS) was added later (Foss et al. 2006).

DOI 10.1080/17453674.2016.1197534

The outcome of medical treatment and rehabilitation is influenced by—and varies with—factors such as age, sex, fracture type, co-morbidity, and (in particular) pre-fracture functional level (Cree and Nade 1999, Thorngren et al. 2005, Kristensen et al. 2010, Kristensen 2011). It therefore makes sense to use assessments of pre-fracture functional level at an early point during hospitalization (Krishnan et al. 2014). Because a substantial proportion of patients have cognitive problems (Jones et al. 2015), it is important to supplement information received from the patient with information from relatives and caregivers.

If we can document reliable assessment of functional level and prediction of outcome based on validated tests and indices at group level in the routine setting, it might be feasible to establish pathways for treatment and rehabilitation of individual patients based on pre-fracture functional level—classified as "higher independent function" and "lower function; dependent on assistance".

We therefore assessed the reliability, validity, and predictive capacity of 2 versions of the Barthel Index, namely Barthel-20 (Collin et al. 1988) and Barthel-100 (Shah et al. 1989), and of 2 short overall assessment instruments, NMS and CAS, in hip fracture patients in the routine setting.

Furthermore, our objective was to define thresholds for identification of patients with higher and lower levels of function in the establishment of focused clinical and rehabilitation pathways, based on a formal statistical analysis.

Patients and methods

165 consecutive patients aged 65 or older from 4 local municipalities, who were acutely admitted with a primary hip fracture to Odense University Hospital (OUH), Svendborg, between August 2012 and April 2013, were prospectively included in the study cohort.

The patients were identified from hospital inpatient records based on primary ICD-10 diagnosis. Completeness of the consecutive patient series was ensured by checking of surgical plans, perusal of the surgical procedures, and overview screens on a daily basis plus follow-up, comparing with monthly patient registry extracts. The hospital is the only one in the area, and all hospital service costs and municipal rehabilitation costs are publicly funded and provided free to the individual patient.

The treatment principles used were based on established clinical pathways that had been developed over a decade and implemented as a collaboration between orthopedic surgeons, geriatricians, nursing staff, physiotherapists, and occupational therapists. Mobilization and other rehabilitation efforts take place immediately after surgery and continue until discharge. Apart from the assessments conducted as part of the project, all the patients received standard treatment and follow-up.

Table 1. Overview, items, and scoring principle for the 4 assessment instruments

Item	Barthel-20	Barthel-100	CAS	NMS
Feeding	0–2	0–10	–	–
Chair-bed transfers	0–3	0–15	0–2	–
Chair-standing transfers	–		0–2	–
Grooming	0–1	0–5	–	–
Toilet	0–2	0–10	–	–
Bathing	0–1	0–5	–	–
Walking inside	0–3	0–15	0–2	0–3
Wheelchair [a]	0–1	0–5	–	–
Walking outside	–	–	–	0–3
Walking to go shopping	–	–	–	0–3
Stairs	0–2	0–10	–	–
Dressing	0–2	0–10	–	–
Bowel	0–2	0–10	–	–
Bladder	0–2	0–10	–	–
Sum-score	0–20	0–100	0–6	0–9

[a] Used for patients confined to a wheelchair, for mobility (as a replacement subset of walking inside).
CAS: Cumulated ambulatory score
NMS: New mobility score

Ethics

As recommended by the local ethics committee, all patients were given information about the project, and participation was on a voluntary basis. The study was approved by the Danish Data Protection Agency (2008-58-0035).

Functional level assessment instruments (Table 1)

The Barthel index was originally developed by Mahoney and Barthel (1965) as a 10-item, weighted measurement instrument for assessment of dependency in function. Collin et al. (1988) adopted a non-weighted 0- to 20-point scale (hereafter called Barthel-20), and Shah et al. (1989) refined it and produced the 100-point weighted scale (hereafter called Barthel-100).

CAS is a Danish language-specific hip fracture measurement instrument that is now mandatory for measurement of pre-fracture functional mobility level (Kristensen et al. 2012). It has been validated as a short-term predictor on the basis of measurements taken over the first 3 postoperative days in 1 study (Foss et al. 2006), but not as an assessment of pre-fracture functional level. CAS consists of 3 items, 2 of which are also found in the Barthel index.

NMS is also a hip fracture-specific instrument. It has been shown to facilitate short-term prediction of inpatient outcome (Kristensen et al. 2010). NMS consists of 3 items concerning walking ability in different contexts. Of the 3 items, only 1 is shared with the other instruments.

The actual formulation of items follows the Danish versions, which are obtainable from (http://www.ouh.dk/uag) for the Barthel scorings, and the original publications for NMS and CAS.

Assessment of functional status

Pre-fracture functional status was prospectively assessed in the days following surgery, based on available information in hospital patient records, in which the entire hospital stay—including accident and emergency room, orthopedic and orthogeriatric department records—is documented by the staff as a basis for treatment and decision making. If a patient had difficulty in giving relevant responses, close relatives or community caregivers were contacted to supplement the patient's information.

All supplementary project-specific assessments were done by the first author (TJP), with no immediate comparison to similar registrations conducted by the regular staff (nurses, occupational therapists, or physiotherapists).

Definition of outcome

All the patients who survived were visited in their homes by TJP 4 months after the fracture, and outcome was assessed. This was assessed using the following definitions.

A. Survival: alive at 4 months.

B. Residential status:
- maintained. Living in own home or in sheltered accommodation before fracture.
- not maintained. Moved to sheltered accommodation or nursing home.

C. Independent ability to walk:
- maintained. Able to walk independently as before (with aids if necessary).
- not maintained. Support from another person or use of a wheelchair required at 4 months.

For analysis of change, all patients were included in A. Only those not living in a nursing home before the fracture were included in B. Only those patients with independent walking ability before the fracture were included in C.

Analysis

Reliability—defined as intraclass correlation coefficient (ICC), changes in mean (95% CI), limits of agreement (LOA), standard error of measurement (SEM), and smallest detectable change (SDC)—was estimated as previously recommended (Lexell and Downham 2005). (STATA version 14 commands: "icc, mean, and baplot". SEM calculation: SD $\sqrt{1 - ICC}$. SDC calculation: SEM × 1.96 × $\sqrt{2}$)

Prediction, defined as non-parametric correlation, was estimated using gamma coefficients (Epidata Analysis version 2.2.2.178), including 95% CI and p-values.

Analyses of sensitivity, specificity, and positive and negative predictive values (PV+ and PV−) with CI were conducted using Wilson's method (Machin et al. 2000). Likelihood ratios (LR+ and LR−) and ROC areas were calculated using the traditional method. (Stata version 14 command: "roctab, binomial").

Optimal cutoff analysis was based on Zhou's optimal decision thresholds on the ROC curve (Zhou et al. 2011), defined as the point c* that satisfies max[Sens(c*) + Spec(c*) − 1], where max means "the maximum of".

Table 2. Characteristics of the participants

Number of participants	165
Female sex	117 (73%)
Age, median (range)	84 (66–102)
Dwelling before	
Own home	107 (65%)
Sheltered home	20 (12%)
Nursing home	38 (23%)
Living alone	119 (72%)
Receiving home assistance	110 (67%)
Walking aids before	
Not used	64 (39%)
Used	92 (56%)
Wheelchair	9 (5%)
Fracture type	
Medial (S72.0)	98 (59%)
Per-trochanteric (S72.1)	59 (36%)
Sub-trochanteric (S72.2)	8 (5%)
Type of surgery	
Hemiarthroplasty	58 (35%)
Osteosynthesis	107 (65%)
Length of stay, median (range)	9 (2–35)

Results

All 165 patients were included in the assessment of survival. 127 were included for residential status, and 156 for walking ability. Mean age was 84 (65–101) years; 117 (73%) were females. Cognitive level OMC 18+ (n = 77) (OMC was missing for 49 patients). The number of patients living in their own home was 107 (65%) (Table 2).

Inter-tester reliability

Inter-tester agreement in the scorings of pre-fracture functional level was rather high, with ICC varying from 0.64 to 0.73. SDC found in the Barthel-20 was 3.4, and in the Barthel-100 it was 15.8 (Table 3, see Supplementary data).

Predictive ability

Correlations between pre-fracture functional level and the 3 outcome variables at 4 months post-fracture were moderate to high (from 0.32 to 0.71) and were statistically significant (Table 4, see Supplementary data).

Predictive validity using optimal thresholds of functional levels

Cumulated ambulatory score (CAS) generally had lower areas under the ROC curve and lower decision thresholds for all 3 outcomes (Table 5).

The 2 versions of the Barthel index and NMS had good predictive values and areas under the ROC curve, which were generally better than CAS, as exemplified by the prediction of survival at 4 months (Figure 1).

Table 5. Analysis of optimal thresholds based on sensitivity, specificity, and ROC analysis regarding prediction of selected outcomes at 4 months

	ROC-area	(95% CI)	Optimal threshold	Lower scores of functioning	Frequency dead	Higher scores of functioning	Frequency dead	Frequency ceiling
Survival at 4 months (n = 165) [a]								
Barthel-20	0.76	(0.70–0.92)	0.48	0–15	20/48	16–20	9/117	61/165
Barthel-100	0.76	(0.69–0.83)	0.45	0–84	20/52	85–100	9/113	59/165
CAS	0.65	(0.58–0.73)	0.28	0–5	12/30	6	17/135	135/165
NMS	0.77	(0.69–0.83)	0.43	0–4	20/55	5–9	9/110	41/165
Maintained residence status (n = 127) [b]								
Barthel-20	0.73	(0.65–0.81)	0.24	0–15	13/17	16–20	32/110	58/127
Barthel-100	0.74	(0.65–0.81)	0.18	0–84	12/19	85–100	33/108	57/127
CAS	0.53	(0.44–0.62)	0.05	0–5	4/7	6	41/120	120/127
NMS	0.74	(0.65–0.81)	0.30	0–4	17/23	5–9	28/104	39/127
Independent walking ability (n = 156) [c]								
Barthel-20	0.65	(0.57–0.72)	0.31	0–15	31/39	16–20	44/117	61/156
Barthel-100	0.66	(0.58–0.73)	0.27	0–84	31/43	85–100	44/113	59/156
CAS	0.60	(0.52–0.68)	0.20	0–5	18/21	6	57/135	135/156
NMS	0.64	(0.56–0.72)	0.25	0–4	32/46	5–9	43/110	41/156

[a–c] See Table 4

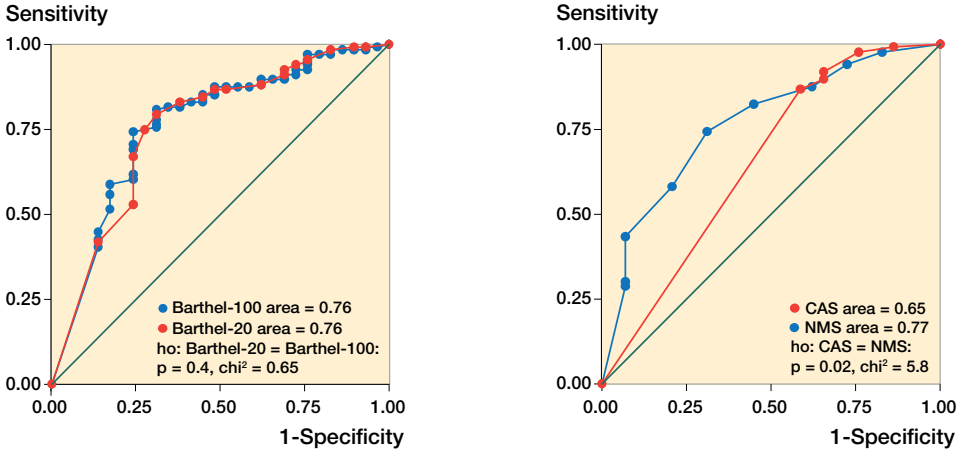

Figure 1. Comparisons of ROC curves for the ability of the 4 instruments to predict survival at 4 months.

No statistically significant differences were seen between the Barthel-20 and Barthel-100 scores. On the other hand, in the case of the 2 shorter, specific hip fracture measurement instruments, NMS was significantly better than CAS (Figure 1).

For prediction of mortality and gait function at 4 months, acceptable values for optimal decision thresholds were found for the 2 versions of the Barthel index and NMS (Table 5).

CAS had a lower value for an optimal decision threshold, and also showed a markedly higher ceiling effect. Due to the low number of patients in the lower functional group, who were dependent on assistance, we were unable to establish similar cut-points for these patients.

Discussion

Outcome at 4 months—defined as survival, individual walking ability, and unchanged residence—was highly correlated to pre-fracture functional level, as measured using the Barthel-20, Barthel-100, NMS, and CAS instruments. This is a first requirement in establishing structured clinical and rehabilitation pathways based on the pre-fracture functional level locally.

With the established levels of SDC, we do see the relevance of suggesting cut-points for higher level of function, which also showed a clear picture in relation to outcomes at 4 months. However, it must also be stated that these results are at the group level; for clinical decision making at the individual level, one must always consider all the information that is available. When looking critically at the analysis, there is no doubt that it is much easier to show convincing prediction ability (non-parametric correlation of prediction at 4 months) than to show classification as indicated with the ROC analysis, where areas are not always convincing.

A point of discussion at the initiation of the Danish National Hip fracture quality assurance program was whether to use NMS or Barthel, and if using the latter, whether to choose Barthel-20 or Barthel-100. Later, CAS was proposed as a replacement for NMS. CAS has only been documented by the originators (Foss et al. 2006). We chose, therefore, to evaluate 2 longer instruments (Barthel-20 and Barthel-100) and 2 shorter instruments (CAS and NMS) with the aim of optimizing evidence for choice of instruments, which will be important in Denmark as well as in other countries (Bryant et al. 2009, Hutchings et al. 2011)

The 4-month outcome point was chosen because most community hip fracture rehabilitation courses are of 3–4 months' duration (Kronborg et al. 2015). In spite of this, about 1 in every 6 patients who received municipality rehabilitation in our study had rehabilitation courses longer than 4 months, and it has been shown that extended outpatient rehabilitation for 6 months improves physical function and quality of life (Binder et al. 2004). With this in mind, it can be discussed whether we have provided enough time to achieve the goals of rehabilitation within 4 months. Anyway, we consider that the goal can be expected either to have been almost achieved or unlikely to be achieved after 4 months.

We see it as a strength of the study that we wanted to enhance the evidence base for cross-sectoral hospital-municipal collaborative efforts using internationally accepted instruments. We therefore followed a complete consecutive patient series from fracture to after rehabilitation, with participation of hospital and municipal staff—which is logistically much more complicated than just working within the hospital. It is common in studies to exclude those who are cognitively impaired, but we chose outcomes that could be estimated regardless of cognitive status.

The relatively small number of participants (165) could be considered a weakness of the study, as could the fact that for some patients there was incomplete assessment in the patient records in the routine setting. This could be because of a lack of acceptance of standardized tools by staff, such as assessment of cognitive levels using the OMC instrument.

In order to classify patients and make decisions at the individual level, there should be evidence that the level of precision for given instruments is comparable to the width of the classes. We found the smallest detectable changes (SDCs) for Barthel-20 (3.4) and Barthel-100 (15.8). It has been suggested that the SDC of Barthel-100 for hip fracture patients is 7.1 points (de Morton et al. 2013). We have not found any studies that have suggested the SDC for Barthel-20, but in stroke patients it was suggested that it is 1.85 (Hsieh et al. 2007). Collin et al. (1988) suggested 2 points as a significant difference in Barthel-20. This all points to the relevance of the suggested cutoff limits of 16+ for Barthel-20 and 85+ for Barthel-100 in the indication of independent function. Further studies must address the clinical relevance and possible consequences in terms of resource allocation, effect, and patient adherence based on such a limit.

In Danish geriatric settings, Barthel-100 has been preferred because of an undocumented assumption that the finer granularity of Barthel-100 as compared to Barthel-20 would give a more precise description of development of function. We could not see such a difference in our study. The ROC analysis indicated that there was no difference between the 2 variants in prediction. From our point of view, the scoring of the Barthel-20 is simpler because it consists of a smaller number of categories.

Of the 2 short hip fracture-specific mobility tests used, the validity of NMS was superior to that of CAS in terms of prediction of outcome at 4 months. The initial developers of NMS studied a combination of NMS and a cognitive score, where a cutoff point at < 5 was proposed for prediction of 1-year mortality (Parker and Palmer 1993), which is the same cutoff for dichotomizing NMS scores as proposed in the current study. Although CAS was shown to be a valid instrument in geriatric settings in 1 study (Kristensen et al. 2012), we found a major ceiling effect (135 of 165 pre-fracture scores were maximum). Of the 2 short instruments, NMS appears to be preferable when all factors are taken into account.

Outcomes best suited for evaluation of prediction are open to debate. Although inclusion of survival and maintained walking ability are prerequisites for independent living, further outcomes regarding body function, activities, and participation (ICF-levels) (WHO 2002) could be considered for a complete evaluation.

In summary, we found promising possibilities for the use of assessment of pre-fracture levels in decision making in individual clinical pathways. Further research on the actual outcomes in further clinical consecutive series should be undertaken. The effect of cognitive level and how it is addressed should be included in future studies, as should the impact of waiting time for rehabilitation on functional outcome. A closer evaluation of agreement and scale composition for the 2 versions of the Barthel index should also be conducted.

Supplementary data

Tables 3 and 4 are available on the Acta Orthopaedica website (www. Actaorthop.org), identification number 9884.

Both authors contributed equally to the work.

No competing interests declared.

Binder E F, Brown M, Sinacore D R, Steger-May K, Yarasheski K E, Schechtman K B. Effects of extended outpatient rehabilitation after hip fracture: a randomized controlled trial. JAMA 2004; 292: 837-46.

Bryant D M, Sanders D W, Coles C P, Petrisor B A, Jeray K J, Laflamme G Y. Selection of outcome measures for patients with hip fracture. J Orthop Trauma 2009; 23: 434-41.

Collin C, Wade D T, Davies S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud 1988; 10: 61-3.

Cree A K, Nade S. How to predict return to the community after fractured proximal femur in the elderly. Aust N Z J Surg 1999; 69: 723-5.

Danish_Interdisciplinary_Register_for_Hip_Fracture. 2015. National annual report 2014 [Online]. Available: https://www.sundhed.dk/content/cms/62/4662_hoftelprrapport2014endelig-kommenteret24-april-2015.pdf.

de Morton N A, Harding K E, Taylor N F, Harrison G. Validity of the de Morton Mobility Index (DEMMI) for measuring the mobility of patients with hip fracture during rehabilitation. Disabil Rehabil 2013; 35: 325-33.

Foss N B, Kristensen M T, Kehlet H. Prediction of postoperative morbidity, mortality and rehabilitation in hip fracture patients: the cumulated ambulation score. Clin Rehabil 2006; 20: 701-8.

Hsieh Y W, Wang C H, Wu S C, Chen P C, Sheu C F, Hsieh C L. Establishing the minimal clinically important difference of the Barthel Index in stroke patients. Neurorehabil Neural Repair 2007; 21: 233-8.

Hutchings L, Fox R, Chesser T. Proximal femoral fractures in the elderly: how are we measuring outcome? Injury 2011; 42: 1205-13.

Jones C J, Rikli R E, Beam W C. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999; 70: 113-9.

Jones C A, Jhangri G S, Feeny D H, Beaupre L A. Cognitive status at hospital admission: postoperative trajectory of functional recovery for hip fracture. J Gerontol A Biol Sci Med Sci 2015 Aug 21. pii: glv138. [Epub ahead of print]

Krishnan M, Beck S, Havelock W, Eeles E, Hubbard R E, Johansen A. Predicting outcome after hip fracture: using a frailty index to integrate comprehensive geriatric assessment results. Age Ageing 2014; 43: 122-6.

Kristensen M T. Factors affecting functional prognosis of patients with hip fracture. Eur J Phys Rehabil Med 2011; 47:257-64.

Kristensen M T, Foss N B, Ekdahl C, Kehlet H. Prefracture functional level evaluated by the New Mobility Score predicts in-hospital outcome after hip fracture surgery. Acta Orthop 2010; 81: 296-302.

Kristensen M T, Jakobsen T L, Nielsen J W, Jorgensen L M, Nienhuis R J, Jonsson L R. Cumulated Ambulation Score to evaluate mobility is feasible in geriatric patients and in patients with hip fracture. Dan Med J 2012; 59: A4464.

Kronborg L, Bandholm T, Kehlet H, Kristensen M T. Municipality-based physical rehabilitation after acute hip fracture surgery in Denmark. Dan Med J 2015; 62: A5023.

Lauritsen J. Simpel funktionsmåling - instruktionsmateriale (DVD og pjece). ISBN 978-87-992241-0-4 [Online]. Available: http://www.ouh.dk/uag. 2007.

Lexell J E, Downham D Y. How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil 2005; 84: 719-23.

Machin D, Altman, D G, Bryant T N, Gardner M. Statistics with confidence: confidence intervals and statistical guidelines, BMJ Books. London, 2000.

Mahoney F I, Barthel D W. Functional evaluation: The Barthel Index. Md State Med J 1965; 14: 61-5.

Parker M J, Palmer C R. A new mobility score for predicting mortality after hip fracture. J Bone Joint Surg Br 1993; 75: 797-8.

Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989; 42: 703-9.

Thorngren K G, Norrman P O, Hommel A, Cedervall M, Thorngren J, Wingstrand H. Influence of age, sex, fracture type and pre-fracture living on rehabilitation pattern after hip fracture in the elderly. Disabil Rehabil 2005; 27: 1091-7.

Wade D T, Vergis E. The Short Orientation-Memory-Concentration Test: a study of its reliability and validity. Clin Rehabil 1999; 13: 164-70.

Wittrup-Jensen K U, Lauridsen J, Gudex C, Pedersen K M. Generation of a Danish TTO value set for EQ-5D health states. Scand J Public Health 2009; 37: 459-66.

WHO. International classification of functioning, disability and health: ICF. Kbh.: Danmarks Blindebibliotek. 2002.

Zhou, X-H, McClish D K, Obuchowski N A. Statistical methods in diagnostic medicine. 2 ed. NJ: Wiley. Hoboken, 2011.