



REFERENCES TO SOME OF THE EVIDENCE BEARING ON THE ASSESSMENT TOOLS IN USE ACROSS RGP SPECIALIZED GERIATRIC SERVICES TEAMS IN ONTARIO

Prepared by Dr. David Ryan and Ms. Judy Kirst, Regional Geriatric Program of Toronto– April 2005

CAREGIVER ASSESSMENT

1. Berg-Weger, M. R., DM; Tebb,SS (2000). "The Caregiver Well-Being Scale revisited." *Health and Social Work* 25(4): 255-63.

The Caregiver Well-Being Scale measures caregiver well-being from a strengths-based perspective by assessing caregivers' human needs and satisfaction with activities of daily living. This article revisits the scale to examine further the scale's psychometric properties using a caregiver-only sample. Reliability is determined through internal consistency. Construct validity is supported through factorial validity with factor analysis. Criterion-related validity is established by examining the concurrent validity of the Well-Being Scale with a measure of depression. Using a sample of family caregivers, results suggest that the Well-Being Scale is a valid and reliable measure.

COGNITIVE ASSESSMENT TOOLS

2. Helmes, E. M., H; Hachinski,VC; Wands,K. (1992). "An examination of psychometric properties of the Extended Scale for Dementia in three different populations." *Alzheimer's Disease and Associated Disorders* 6(4): 236-46.

The Extended Scale for Dementia (ESD), a development of the Mattis Dementia Rating Scale, has been used in the evaluation of dementia and aging and has shown substantial clinical utility. We report on analyses of its properties and internal structure in three samples of older people: 153 normals, 101 psychiatric hospital residents, and 114 patients with Alzheimer's disease. The results showed good internal consistency in the two clinical samples, with much lower reliability in the normals, for whom the test was too easy. A review of the item statistics led to the use of 17 of the 23 ESD items in item component analyses in the three samples. Use of Horn's parallel analysis criterions led to the retention of three components in the normal group and one in both the hospital and Alzheimer group. The results are compared with other work and are in accordance with the view that cognitive structure becomes more simple with increasing dementia.

- 3 Folstein, M. F., SE; McHugh,PR (1975). ""Mini-Mental State". A Practical Method for Grading the Cognitive State of Patients for the Clinician." *Journal of Psychiatric Research* 12(3): 189-198.

Full article at www.minimental.com/article.html

4. Hopkins, R. K., LA; Day,DJ; Rows,CP; Hamilton,PF (2003). "The Revised Kingston Standardized Cognitive Assessment." *International Journal of Geriatric Psychiatry* 19(4): 320-326.

The original Kingston Standardized Cognitive Assessment (KSCA) was designed to assess cognitive functioning in the elderly with suspected organic brain damage (i.e. dementia). It was specifically designed to be a relatively quickly administered assessment tool available to mental health professionals who were not trained in specialized cognitive assessment techniques. It was introduced over a decade ago to bridge a gap between brief, narrowly focused rating scales, and intensive, expensive, full neuropsychological assessment. Recently, a revision of the KSCA was completed. This revision includes the addition of a word-list memory task with immediate recall, delayed recall and recognition formats, as well as new norms for patients with Alzheimer's disease (AD). The updated norms reflect the abilities of higher-functioning (community-dwelling) patients. In order to facilitate the Revised KSCAs use we have developed a new scoring and analysis form as well as a more comprehensive scoring and administration manual. These changes have resulted in better detection of earlier Alzheimer's disease and use of comparison groups that reflect the changing referral base. The structure of the revised scale and updated normative data are described. An illustrative clinical case example is also provided.

5. Rodenburg, M. H., RW; Hamilton,PF; Ginsburg,L; Nashed,Y; Minde,N. (1991). "The Kingston Standardized Cognitive Assessment." *International Journal of Geriatric Psychiatry* 6(12): 867-874.

In spite of the fact that demographic trends show an increase in the number of elderly persons, and hence more diagnosed as suffering from senile dementia, the cognitive assessment of the organically brain damaged elderly person still too often depends on the use of brief and incomplete rating scales, and is thus often poorly performed. The simple dementia rating scales in current use may fail to provide the range or depth needed for an adequate evaluation of the individual patient. The Kingston Standardized Cognitive Assessment is offered as an improved method of assessment. This instrument is described here in terms of its structure, function, statistical properties and normative data. We also provide examples of its clinical use.

6. Ronnberg, L. E., K. (1994). "Reliability and validity of the Hierarchic Dementia Scale." *International Psychogeriatrics* 6(1): 87-94.

The aim of the study was to apply a method for measuring cognitive function in severely demented patients. Reliability and validity of the Hierarchic Dementia Scale (HDS) were tested. Fifty patients with dementia of the Alzheimer's type (DAT), multi-infarct dementia (MID), and dementia of mixed type (MIX) were studied. The interrater reliability was satisfactory as estimated by means of a kappa coefficient. The test retest reliability was $r_s=.96$. The concurrent validity of the HDS, as measured by the correlation between the HDS and the MMSE was $r_s=.86$, and between the HDS and the CDR was $r_s=-.71$. The results indicate that the HDS is a useful and valid instrument for determination of the heterogeneous cognitive deficits in severe dementia.

7. Shulman, K. (2000). "Clock-Drawing: is it the ideal cognitive screening test?" *International Journal of Geriatric Psychiatry* 15(6): 548-61.

The clock-drawing test has achieved widespread clinical use in recent years as a cognitive screening instrument and a significant amount of literature relates to its psychometric properties and clinical utility. This review aims to synthesize the available evidence and assess the value of this screening test according to well-defined criteria. DESIGN: A Medline and Psycho-info literature search of all languages was done from 1983 to 1998 including manual cross-referencing of bibliographies. A brief summary of all original scoring systems is provided as well as a review of replication studies. Psychometric data including correlations with other cognitive tests were recorded. Qualitative aspects of the test are also described. RESULTS: Among published studies, the mean sensitivity (85%) and specificity (85%) of the clock-drawing test are impressive. Correlations with the Mini-Mental State Examination and other cognitive tests was high, generally greater than $r=0.5$. High levels of inter-rater and test-retest reliability and positive predictive values are recorded and despite significant variability in the scoring systems, all report similar psychometric properties. The clock test also shows a sensitivity to cognitive change with good predictive validity. CONCLUSIONS: The clock-drawing test meets defined criteria for a cognitive screening instrument. It taps into a wide range of cognitive abilities including executive functions, is quick and easy to administer and score with excellent acceptability by subjects. Together with informant reports, the clock-drawing test is complementary to the widely used and validated Mini-Mental State Examination and should provide a significant advance in the early detection of dementia and in monitoring cognitive change. A simple scoring system with emphasis on the qualitative aspects of clock-drawing should maximize its utility.

COMPARISONS OF COGNITIVE ASSESSMENT TOOLS

8. Grace, J. N., JD; White, DA; Guilmette, TJ; Giuliano,AJ; Monsch,AU; Snow,MG (1995). "Folstein vs modified Mini-Mental State Examination in geriatric stroke. Stability, validity and screening utility." *Archives of Neurology* 52(2): 477-84.

Two studies were conducted with the Folstein Mini-Mental State Examination (MMS) and the Modified Mini-Mental State Examination (3MS) to examine whether the expanded version is a more useful screening tool in stroke populations. DESIGN: Clinical utility of screening tests (MMS and 3MS) was evaluated in reference to neuropsychological performance and functional outcome in rehabilitation. SETTING: Medical rehabilitation unit of university-affiliated hospital. PATIENTS: Two groups ($n=77$), and ($n=70$) of patients who were admitted consecutively. MAIN OUTCOME MEASURES: Neuropsychological performance and functional outcome (functional independence measure). RESULTS: The reliability, stability, and validity of the 3MS were established in the stroke population. Classification accuracy did not differ between the MMS and the 3MS, although the 3MS had higher sensitivity. In detecting cognitive impairment when compared with the extended neuropsychological battery, both instruments were adequate in patients with left-sided cerebrovascular accidents and were ineffective in patients with right-sided cerebrovascular accidents. The 3MS correlated with some cognitive domains missed by the MMS, thus adding useful clinical information. Finally, the 3MS was a significantly better predictor of functional outcome than the MMS. CONCLUSIONS: The 3MS was found to be a reliable, valid, and stable cognitive screening instrument in the stroke population. Classification accuracy indicates that both screening instruments are not strong in right-sided cerebrovascular accidents. The 3MS does have some advantages over the MMS; the expanded version of the screen not only provided additional cognitive information but also allows for better predicting of functional outcome.

CONFUSION AND DELIRIUM ASSESSMENT TOOLS

9. Inouye, S. v., CH; Alessi,CA; Balkin,S; Siegal,AP; Horwitz,RI (1990). "Clarifying confusion: the Confusion Assessment Method. A new method for detection of delirium." *Annals of Internal Medicine* 113(12): 941-8.

OBJECTIVE: To develop and validate a new standardized confusion assessment method (CAM) that enables nonpsychiatric clinicians to detect delirium quickly in high-risk settings. DESIGN: Prospective validation study. SETTING: Conducted in general medicine wards and in an outpatient geriatric assessment center in Yale University (site 1) and in general medicine wards at the University of Chicago (site 2). PATIENTS: The study included 56 subjects, ranging in age from 65 to 98 years. At site 1, 10 patients

with and 20 patients without delirium participated; at site 2, 16 patients with and 10 without delirium participated. MEASUREMENTS AND MAIN RESULTS: An expert panel developed the CAM through a consensus building process. The CAM instrument, which can be completed in less than 5 minutes, consists of nine operationalized criteria from the Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R). An a priori hypothesis was established for the diagnostic value of four criteria: acute onset and fluctuating course, inattention, disorganized thinking, and altered level of consciousness. The CAM algorithm for diagnosis of delirium required the presence of both the first and second criteria and of either the third or fourth criterion. At both sites, the diagnoses made by the CAM were concurrently validated against the diagnoses made by psychiatrists. At sites 1 and 2 values for sensitivity were 100% and 94%, respectively; values for specificity were 95% and 90%; values for positive predictive accuracy were 91% and 94%; and values for negative accuracy were 100% and 90%. The CAM algorithm had the highest predictive accuracy for all possible combinations of the nine features of delirium. The CAM was shown to have convergent agreement with four other mental status tests, including the Mini-Mental State Examination. The interobserver reliability of the CAM was high ($\kappa=0.81 - 1.0$). CONCLUSIONS: The CAM is sensitive, specific, reliable, and easy to use for identification of delirium.

10. Schuurmans, M. S.-B., LM; Duursma, SA (2003). "The Delirium Observation Screening Scale: a screening instrument for delirium." Research & Theory for Nursing Practice 17(1): 31-50.

The Delirium Observation Screening (DOS) scale, a 25-item scale, was developed to facilitate early recognition of delirium, according to the Diagnostic and Statistical Manual-IV criteria, based on nurses' observations during regular care. The scale was tested for content validity by a group of seven experts in the field of delirium. Internal consistency, predictive validity, and concurrent and construct validity were tested in two prospective studies with high risk groups of patients: geriatric medicine patients and elderly hip fracture patients. Among the patients admitted to a geriatric department (N=82), 4 became delirious; among the elderly hip fracture patients (N=92), 18 became delirious. The DOS scale was determined to be content valid and showed high internal consistency, $\alpha=0.93$ and $\alpha=0.96$. Predictive validity against the Diagnostic and Statistical Manual-IV diagnosis of delirium made by a geriatrician was good in both studies. Correlations of the DOS scale with the Mini Mental State Examination (MMSE) were $R_s=0.79$ ($p<or=0.001$) in the hip fracture patients and $R_s=0.66$ ($p<or=0.001$) in the geriatric medicine patients. Concurrent validity, as tested by comparison of the research nurse's ratings of the DOS scale and the Confusion Assessment Method (CAM), for the group of hip fracture patients was 0.63 ($p<or=0.001$). Construct validity of the DOS was tested against the Informant Questionnaire of Cognitive Decline in Elderly (IQCODE), a preexisting psychiatric diagnosis and the Barthel Index. Correlation with the IQCODE was 0.74 ($p<or=0.001$) in the study with the hip fracture patients and 0.33 ($p<or=0.05$) in the study with the geriatric medicine patients. Correlation with the Barthel Index was -0.26 ($p<or=0.05$) in the geriatric medicine patients and -0.55 ($p<or=0.001$) in the hip fracture patients. The overall conclusion of these studies is that the DOS scale shows satisfactory validity and reliability, to guide early recognition of delirium by nurses' observation.

DEPRESSION Assessment

11. Kurlowicz, L. E., LK; Strumpf,NE; Maislin,G (2002). "A psychometric evaluation of the Cornell Scale for Depression in Dementia in a frail, nursing home population." American Journal of Geriatric Psychiatry 10(5): 600-8.

OBJECTIVE: The authors conducted a psychometric evaluation of the Cornell Scale for Depression in Dementia (CSDD) through factor analysis and assessment of criterion validity in an older, frail nursing home population, with a secondary analysis of pre-intervention data from a longitudinal clinical trial aimed at reducing restraints in nursing homes. METHODS: The sample for the present study was 642 nursing home residents (mean [SD] age 84.3 [7.6] years; range: 61-105; 82% women) with completed CSDD scores, who were interviewed immediately before the intervention. Nursing home residents' scores from the 19-item CSDD were subjected to exploratory factor analysis and criterion-validity analysis. RESULTS: The factor analysis resulted in four distinct clinically interpretable domains: Depression, Somatic/Vegetative, Disturbed Sleep, and Anxiety. Sixteen items were retained in these domains, and summated score indices and a global score were constructed. The global score and the four indices demonstrated adequate internal consistency and reliability. The indices generated by the factor analysis correlated as expected with criterion variables. CONCLUSION: Results suggested that in frail, institutionalized older adults with high rates of dementia, medical illness, and functional disability, depression measurement methods that are less dependent on items highly sensitive to comorbid conditions and not necessarily associated with depression may be more appropriate. Authors recommend further validity testing of the CSDD with similar populations of frail, institutionalized older adults.

12. Sutcliffe, C. C., L; Burns,A; Mozley,CG; Bagley,H; Huxley,P; Challis,D (2000). "A new version of the Geriatric Depression Scale for nursing and residential populations: the geriatric depression scale (residential) (GDS-12R)." International Psychogeriatrics 12(2): 173-81.

The objective was to develop a new short-form Geriatric Depression Scale (GDS-12R) suitable for older people living in nursing and residential care settings, including those persons with significant cognitive impairment. A total of 308 newly admitted residents of 30 nursing and residential homes in northwest England were interviewed using the Geriatric Depression Scale (GDS-15), the Mini-Mental State Examination, and the Affect Balance Scale (ABS). A 12-item version of the GDS was shown to have greater internal reliability than the 15-item version, because of the context-dependent nature of the deleted items. There was close agreement between the GDS-12R items and another indicator of depressed mood (a single item from the ABS). Furthermore, moderate to high levels of cognitive impairment did not affect the performance of the new version of the scale. The GDS-12R provides researchers and clinicians with a brief, easy-to-administer depression scale that is relevant to residential and nursing home populations.

13. Weyerer, S. K., U; Ames,D;Allen,N. (1999). "The Even Briefer Assessment Scale for Depression (EBAS DEP): its suitability for the elderly in geriatric care in English -and German- speaking countries." International Journal of Geriatric Psychiatry 14(6): 473-80.

To determine the psychometric properties of the Even Briefer Assessment Scale for Depression (EBAS DEP) developed by Allen et al. (1994) among samples of the German elderly and compare the results with those from English-speaking countries. DESIGN: Depression scale scores from elderly persons in residential and gerontopsychiatric care were assessed for internal consistency using Cronbach's alpha and validity (Feighner Criteria of Depression and Center for Epidemiologic Studies Depression Scale.) SUBJECTS: Eight hundred and fifty-two elderly persons aged 65 years and over, in residential and gerontopsychiatric care. MATERIALS: The eight-item Even Briefer Assessment Scale for Depression (EBAS DEP) derived from the 21-item Brief Assessment Scale (BAS DEP). RESULTS: The analysis of the reliability of the German EBAS DEP yielded, as did that of the English version, a satisfactorily high internal consistency (0.73 and higher). Based on a subset of 71 subjects, the validity of the scale was tested by independent psychiatric experts using the Feighner Criteria of Depression. The EBAS DEP (cutoff 3/4) had a sensitivity and specificity for a diagnosis of depression of 93.3% and 85.3% respectively. Similar results were reported by Allen et al., but at a lower cutoff (2/3). In agreement with the English findings, the receiver operating curve (ROC) statistics revealed that the EBAS DEP is a screening instrument which is as efficient as the longer BAS DEP. CONCLUSION: The EBAS DEP is an instrument well suited for use in screening the depressed elderly in different settings in English- and German- speaking countries.

14. Yesavage, J. A. B., T.L. (1983). "Development and Validation of a Geriatric Depression Screening Scale: A Preliminary Report." J. psychiat. Res. 17(1): 37 - 49.

A new Geriatric Depression Scale (GDS) designed specifically for rating depression in the elderly was tested for reliability and validity and compared with the Hamilton Rating Scale for Depression (HRS-D) and the Zung Self-Rating Depression Scale (SDS). In constructing the GDS a 100-item questionnaire was administered to normal and severely depressed subjects. The 30 questions most highly correlated with the total scores were then selected and readministered to new groups of elderly subjects. The subjects were classified as normal, mildly depressed or severely depressed on the basis of Research Diagnostic Criteria (RDC) for depression. The GDS, HRS-D and SDS were all found to be internally consistent measures, and each of the scales was correlated with the subject's number of RDC symptoms. However, the GDS and the HRS-D were significantly better correlated with RDC symptoms than was the SDS. The authors suggest that the GDS represents a reliable and valid self-rating depression screening scale for the elderly populations.

FALLS, BALANCE & MOBILITY ASSESSMENT

16. Brosseau, L. L., L; Couroux,N; Marion,M; Theriault,J (1999). "Intra- and inter-rater reliability and factorial validity studies of the Physiotherapy Functional Mobility Profile (PFMP) in acute care patients." Physiotherapy Theory and Practice 15(3): 147-154.

In response to the increasing importance given to the use of outcome measures in physiotherapy, a new clinical tool known as the Physiotherapy Functional Mobility Profile (PFMP) was developed to assess functional mobility. Some reliability and validity studies have been conducted by this team with a chronic care population. The aims of this study were to examine intra- and inter-rater reliability and construct validity of the PFMP with an inpatient acute care population. Subject inpatients (mean age 68.6 years) from six different specialties (neurology, orthopaedics, surgery, general medicine, pulmonary medicine and rheumatology). The patients were videotaped during their initial physiotherapy assessment while performing items of the PFMP. For intra-rater reliability, the videotaped performance of the 10 patients was viewed and scored twice by two trained physiotherapists with a one week interval between viewings. Inter-rater reliability was examined by comparing the scoring of the two raters on the first viewing. All the data were used for factorial validity. Intra-rater and inter-rater reliability were examined with intra-class correlation coefficients (ICC) based on an ANOVA (random) procedure for repeated data. Kappa statistics were also used to assess reliability. The factorial validity of the PFMP was examined through a factor analysis. A.05 level of significance was chosen for each of the statistical tests. Intra-rater reliability was also excellent, with an ICC of.99 (P<.05). Inter-rater reliability was also impressive (ICC=.97). Factorial analysis revealed that there was one dimension present, which can be interpreted as disability. The preliminary results of this study showed that the PFMP appears to be a reliable and valid tool which can be applied to an acute care population, and that it is representative of the disability concept.

17. Hamilton, D. H., RG (2000). "Validity and reliability of the 6-minute walk test in a cardiac rehabilitation population." Journal of Cardiopulmonary Rehabilitation 20(3): 156-64.

PURPOSE: Although the 6-minute walk test is commonly used to assess the functional status of patients with severe cardiopulmonary disease, few studies have tested its value in a cardiac rehabilitation (CR) population with milder disease status. The purpose of this study was to examine the validity and reliability of the 6-minute walk test in a Phase II/III CR program. METHODS: Ninety-four patients (61 men, 33 women) aged 63+/- 10 years completed three 6-minute walks on nonconsecutive days. Patients also completed the Duke Activity Status Index (DASI) and the Short Form 36 Health Survey (SF-36). In addition, maximum metabolic equivalents (METs) from a symptom-limited graded exercise test were obtained from files. RESULTS: The 6-minute walk was linearly related to maximum METs ($r=0.687$, $P<0.001$), supporting the validity of the test. Patients walked significantly farther in each 6-minute walk ($F=19.83$, $P<0.001$), and strong test-retest reliability was demonstrated (intraclass correlation = 0.97). Distance walked decreased with older age ($F=19.49$, $P<0.001$), with men walking farther than women ($F=7.19$, $P<0.01$). The 6-minute walk was moderately correlated with scores from the DASI ($r=0.502$, $P<0.001$), and the Physical Function subscale of the SF-36 ($r=0.624$, $P<0.001$). CONCLUSIONS: The 6-minute walk test is a valid and reliable method of assessing functional ability in a Phase II/III CR population. A learning effect of 6% was observed over the three walks; however, it is unknown if this learning effect will be retained over time. This test may be particularly valuable to smaller CR centers that want to document functional improvements but do not have access to conventional treadmill tests.

19. Inness, E. H., J; Niechwiej,E; Jaglal,S; McIlroy,WE; Verrier,MC (2005). Measuring Balance after Traumatic Brain Injury: further validation of the Community Balance and Mobility Scale. Partners for Sustainability and Success - Rehabilitation:A Client and Family Affair, Toronto, Ontario.

PURPOSE: To examine the validity and sensitivity of the Community Balance and Mobility Scale (CB&M), developed for evaluation of individuals with Traumatic Brain Injury (TBI). **RELEVANCE:** Problems with balance are reported to be a longstanding concern post-TBI, possibly contributing to reduced community participation. Laboratory measures are able to detect persistent postural instability in ambulatory individuals. The CB&M, however, was developed as a clinical tool, and examined in this study for its ability to detect balance deficits and relevance to community integration following brain injury. **SUBJECTS:** A convenience sample (n=36) of ambulatory patients with TBI. **METHODS:** Sensitivity studies: Balance Scale (Berg) and CB&M were administered at admission and discharge (n=23). Validity studies: CB&M, laboratory assessments of postural sway and spatiotemporal variables of gait (n=27), and Community Integration Questionnaire (CIQ) were administered within a 5-day time frame. **ANALYSIS:** Standardized response means (SRM's) were calculated comparing change scores on the Balance Scale and CB&M. Correlations with CB&M and laboratory assessment, and CIQ were calculated. For the latter, pooled data across CB&M studies (n=47) were used. **RESULTS:** The SRM's for the Balance Scale and CB&M were 0.54 and 1.26, respectively. Correlations between CB&M and postural sway did not reach significance. Correlations with gait variables were highly significant: walking velocity (r=0.69), step length (r=0.75), step length variability (r=0.53) and step time variability (r=0.49). There was a significant correlation between CB&M and CIQ scores (r=0.54) **CONCLUSIONS:** The CB&M is more sensitive than the Balance Scale for higher functioning patients with TBI. The CB&M and CIQ correlation suggests that those with better balance skills are also those demonstrating greater community integration. The strong relationship of the CB&M with spatiotemporal measure of gait substantiates the utility of the CB&M - to evaluate the high level dynamic mobility required for the community environment.

20. Light, K. R., DK; Purser, JL. (1995). "The Functional Reach test for balance: criterion-related validity of clinical observations." Issues on Aging 18(2): 5-9.

Balance control assessment is a routine component of geriatric assessment. One of the better researched clinical assessments for balance disability is the Functional Reach (FR) Test. The FR Test has met many criteria of reliability and validity, however, the validity of clinical observation for measuring FR has not been tested against a simultaneously performed criterion standard. Clinical observations sometimes suffer from significant reading errors. The purpose of this study was to examine the criterion-related validity of clinical observation of the FR Test to those of videotape analysis. The results of this study indicate that carefully trained clinicians are capable of reading the FR measurement on a yardstick to 1/2 inch and that these readings correlated highly to those of videotape analysis (ICC=.86) These findings on the criterion-related validity of FR clinical observation adds to the standardization of the FR Test as a useful measure of balance control.

21. Morris, S. M., ME; Iansek,R (2001). "Reliability of measurements obtained with the Timed "Up & Go" test in people with Parkinson disease." Physical Therapy 81(2): 810-8.

The Timed "Up & Go" Test (TUG) is used to measure the ability of patients to perform sequential locomotor tasks that incorporate walking and turning. This study investigated the retest reliability, interrater reliability, and sensitivity of scores obtained with the TUG in detecting changes in mobility in subjects with idiopathic Parkinson disease (PD). **SUBJECTS:** The performance of 12 people with PD was compared with that of 12 age-matched subjects without PD. **METHODS:** The subjects with PD completed 5 trials of the TUG after withdrawal of levodopa for 12 hours ("off" phase of the medication cycle) as well as additional 5 trials 1 hour after levodopa was administered ("on" phase of the medication cycle). They were scored on the Modified Webster Scale at both sessions. The comparison subjects also performed 5 TUG trials. All trials were videotaped and timed by 2 experienced raters. The videotape was later rated by 3 experienced clinicians and 3 inexperienced clinicians. **RESULTS:** For the subjects with PD, within-session performance was highly consistent, with correlations (r) ranging from .80 to .98 for the "off" phase and from .73 to .99 for the "on" phase. The performance of the comparison subjects across the 5 trials was also highly consistent (r=.90-.97). Comparisons showed differences between trials 1 and 2 on the TUG for both groups. Removal of data for trial 1 (the practice trial) further enhanced retest reliability. There was close agreement in TUG scores among raters despite different levels of experience (intraclass correlation coefficient [3,1]=.87-.99). Mean TUG scores were different between the "on" and "off" phases of the levodopa cycle and between subjects with PD and comparison subjects during the "on" phase. **CONCLUSIONS AND DISCUSSION:** Retest reliability and interrater reliability of the TUG measurements were high, and the measurements reflected changes in performance according to levodopa use. The TUG can also be used to detect differences in performance between people with PD and elderly people without PD.

22. Myers, A. F., PC; Myers,AH; Sherk,W (1998). "Discriminative and evaluative properties of the activities-specific balance confidence (ABC) scale." Journals of Gerontology Series A: Biological Sciences and Medical Sciences 53(4): M287-M294.

BACKGROUND: Balance confidence is an important indicator of functional mobility and independence in older adults. Preliminary psychometric evidence for the Activities-specific Balance Confidence (ABC) Scale is promising with a series of four studies adding information on the discriminative and evaluative properties of this tool. **METHODS:** The original validation sample was reinterviewed one year later. In the second study, the ABC was administered to 475 older adults ranging from home care clients to highly functioning individuals in community exercise programs. The third study compared 31 residents of retirement homes given a 10-week balance control exercise program and fall education with 32 residents who received only fall education. The fourth study examined balance confidence preoperatively and postoperatively for 27 patients undergoing hip and knee replacement. **RESULTS:** ABC scores remained stable over 12 months in higher functioning elders, but deteriorated in retirement home residents over 26 weeks. Ten weeks of balance training significantly improved balance confidence, as did hip or knee replacement with standard physical therapy. ABC scores lower than 50 indicated a low level of physical functioning characteristic of home care clients. ABC scores above 50 and lower than 80 indicated a moderate level of functioning characteristic of elders in retirement homes and persons with chronic health conditions. ABC scores above 80 are indicative of highly functioning, usually physically active older adults, and are achievable through exercise and rehabilitative therapies. **CONCLUSIONS:** Balance confidence is amenable to

change and able to distinguish between elders at various levels of functional mobility. These results provide comparative benchmarks for researchers and clinicians working with different groups of older adults.

23. Newton, R. (2001). "Validity of the Multi-Directional Reach Test: a practical measure for limits of stability in older adults." *Journal of Gerontology Series A-Biological Sciences & Medical Sciences* 56(4): M248-52.

Falls occur not only in the forward direction, but also to the side and backward. The purpose of this study was to develop a portable and valid tool to measure limits of stability in the antero-posterior and medial-lateral directions. METHODS: Two hundred fifty-four community-dwelling older persons were administered the Berg Balance Test (BBT), the Timed Up & Go Test (TUG), and the Multi-Directional Reach Test (MDRT). For the MDRT, subjects performed maximal reaches with the outstretched arm forward (FR), the right (RR), to the left (LR), and leaning backward (BR), with feet flat on the floor. Reach was measured by the subject's total hand excursion along a yardstick affixed to a telescoping tripod. RESULTS: Mean scores of the MDRT were FR=8.89+/- 3.4 in., BR=4.64+/-3.07 in., RR=6.15+/-2.99 in., and LR=6.61+/-2.88 in. Interclass correlation (ICC2,1) for the reaches were greater than .92. Reliability analysis (Cronbach's Alpha .82) demonstrated that directional reaches measure similar but unique aspects of the MDRT. The MDRT demonstrated significant correlation with the BBT sum and significant inverse relationship with the scores of the TUG. Regression analysis revealed that activity level contributed to scores in the forward, right and left direction and that fear of falling contributes to scores in the backward direction. CONCLUSION: The Multi-Directional Reach Test is an inexpensive, reliable and valid tool for measuring the limits of stability as derived by reach in four directions. Values obtained on relatively healthy community-dwelling older adults serve as norms for screening populations

24. Cipriany-Dacko, L. I., D; Johannsen, J; Rude, V (1997). "Interrater reliability of the Tinetti Balance Scores in novice and experienced physical therapy clinicians." *Archives of Physical Medicine and Rehabilitation* 78(10): 1160-4.

To examine interrater agreement of scores by physical therapy novices and experienced clinicians on videotaped and live performances of the balance portion of Tinetti's Performance Oriented Mobility Assessment (BPOMA). DESIGN: A reliability design was used to assess the interrater agreement and consistency of the BPOMA score in an elderly population. SETTING: General community hospital and skilled nursing facility. PATIENTS: Twenty-six residents of a skilled nursing home, ranging in age from 66 to 99 yrs (mean=80.4, SD=6.8), participated in Phase 1. Twenty-four hospital inpatients and five residents of a skilled nursing home, ranging in age from 60 to 92 yrs (mean=74.7, SD=7.9), participated in Phase 2. RATERS: Three student physical therapists scored the patients in Phase 1. One student was designated the administering rater (AR). The AR instructed, guarded, and scored the subjects. The other two students were the observing raters (ORs), whose role was to observe and score the subject's performances. Nine physical therapy clinicians, ranging from 0 to 6 years experience, rated the subjects in Phase 2. MAIN OUTCOME MEASURES: Consistency and agreement of BPOMA scores were compared between clinicians with varying levels of experience. In Phase 1, BPOMA was scored on-site by three student physical therapists. In Phase 2, videotaped performances were scored by five physical therapists, one physical therapist assistant, and three student physical therapists. RESULTS: Phase 1 demonstrated fair to excellent kappa coefficients (40-1.00) in all maneuvers across all raters. The ORs had higher agreement compared with the AR, ranging from good to excellent (75-1.00). Phase 2 demonstrated fair to good kappa coefficients (40-.75) in 5 of 8 maneuvers across all nine raters. When comparing proportion of observed agreement to evaluate the years of experience on rater agreement, there was no significant difference between clinician groups. CONCLUSIONS: Fair to good reliability of BPOMA scores occurred across many rates of varied experience with a small amount of training.

25. Papaioannou, A. P., William; Cook, Richard; Ferko, Nicole; Coker, Esther; Adachi, Jonathan (2004). "Prediction of Falls using a risk assessment tool in the acute care setting." *BMC Medicine* 2(1).

The British STRATIFY tool was previously developed to predict falls in hospital. Although the tool has several strengths, certain limitations exist which may not allow generalizability to a Canadian setting. Thus, we tested the STRATIFY tool with some modification and re-weighting of items in Canadian hospitals. Methods: This was a prospective validation cohort study in four acute care medical units of two teaching hospitals in Hamilton, Ontario. In total, 620 patients over the age of 65 years admitted during a 6-month period. Five patient characteristics found to be risk factors for falls in the British STRATIFY study were tested for predicative validity. The characteristics included history of falls, mental impairment, visual impairment, toileting, and dependency in transfers and mobility. Multivariate logistic regression was used to obtain optimal weights to the construction of a risk score. A receiver-operating characteristic curve was generated to show sensitivities and specificities for predicting falls based on different threshold scores for considering patients at high risk. Results: Inter-rater reliability for the weighted risk score indicated very good agreement (inter-class correlation coefficient = 0.78). History of falls, mental impairment, toileting difficulties, and dependency in transfer/mobility significantly predicted fallers. In the multivariate model, mental status was a significant predictor ($P < 0.001$) while history of falls and transfer/mobility difficulties approached significance ($P = 0.089$ and $P = 0.077$ respectively). The logistic regression model led to weights for a risk score on a 30-point scale. A risk score of 9 or more gave a sensitivity of 91% and specificity of 60% for predicting who would fall. Conclusion: Good predicative validity for identifying fallers was achieved in a Canadian setting using a simple-to-obtain risk score that can easily be incorporated into practice.

26. Potter, J. E., AL; Duncan, G (1995). "Gait Speed and activities of daily living function in geriatric patients." *Archives of Physical Medicine and Rehabilitation* 76(11): 997-9.

To establish the relationship between gait speed (GS) and functional independence in elderly people. DESIGN: GS is suggested as being a criterion standard in rehabilitation reflecting muscle strength. This study assessed the relationship between gait speed and functional independence in Activities of Daily Living (ADL). GS was measured by portable accelerometer over 2 meters. The mean of 3 attempts was taken. ADL function was measured by an occupational therapist using the modified Barthel ADL Index. The relationship between these measurements was assessed by a statistician. SETTING: A geriatric unit in a hospital in Scotland. PATIENTS: One hundred sixty-one inpatients and outpatients were selected at random from the patients of a geriatric unit over a 3-

month period. Patients were eligible if they were mobile with or without a walking aid. INTERVENTIONS: GS was measured by a portable ultrasonic accelerometer. Patients were reviewed by an occupational therapist, blinded to their GS, who recorded functional capacity. Case sheet review provided diagnostic details and cognitive function. The type of floor surface was recorded. MAIN OUTCOME MEASURES: GS (m/sec), and Barthel score. RESULTS: Patients with GS of <.25m/sec were more likely to be dependent in one or more ADL function, $p<.01$. Those with a GS between .35 and .55m/sec were more likely to be independent in all ADL functions, $p<.001$. Patients whose GS was >.55m/sec did not maintain this independence. There was no relationship between GS and floor surface or cognitive function. CONCLUSIONS: GS is a useful indicator of ADL function in geriatric patients.

27. Prosser, L. C., A (1997). "Further validation of the Elderly Mobility Scale for measurement of mobility of hospitalized elderly people." *Clinical Rehabilitation* 11(4): 338-43.

To further assess the validity and inter-rater reliability of the Elderly Mobility Scale (EMS). Also whether the scale reflects elderly people's perceptions regarding their mobility and whether it can predict discharge destination, or likelihood of falling. DESIGN: Questionnaire-based study completed on admission and weekly after this on all patients referred to physiotherapy for mobility problems over the course of one month. SETTING: Care of the elderly wards in the Bristol General Hospital. SUBJECTS: Sixty-six patients (ages 66-69 years, 66% female) were included in the validity study. Nineteen patients (ages 71-95 years, 47% female) were included in the inter-rater reliability study. INTERVENTIONS: EMS, Barthel and patient's perceptions of mobility were tested with routine physiotherapy treatment carried out as necessary. MAIN OUTCOME MEASURES: Concurrent validity was assessed by correlating EMS scores with Barthel scores using Spearman's test. Inter-rater reliability was also tested using a Spearman's correlation. EMS scores of patients were also evaluated in conjunction with whether or not they fell and their destination on discharge. RESULTS: A significant correlation between EMS and Barthel scores indicated concurrent validity. Inter-rater reliability was demonstrated on 19 patients with significant correlation between scores. No predictive validity could be ascribed to EMS in terms of discharge destination or likelihood of falling. Results do indicate a possible predictive validity of the functional reach component of the EMS regarding the risk of future falls. CONCLUSIONS: The EMS was found to be a valid scale with good inter-rater reliability that could be readily applied during daily clinical work. However, it was found to have no predictive validity in terms of falling or discharge destination.

28. Read, A. C., J; Hubley-Kozey, C.L. (1998). "A Reliability Study of The Individual Phases of the Timed Up and Go Test for Older Adults." *American Society of Biomechanics* (<http://asb-biomech.org/onlineabs/NACOB98?172/>).

Accurate, reliable and valid assessments of the functional mobility of older adults is important to the physiotherapy management of older adults who are at risk for declines in function. The timed "Up and Go" test (TUGT) has been described as a screening tool and assessment tool for the functional mobility of the older adult (McCrae et al. 1996, Podsiadlo and Richardson, 1991). The TUGT was considered to have content validity as a measure of functional mobility because the components of the test are tasks considered important in activities of daily living (Podsiadlo and Richardson, 1991). The time to complete the TUGT, which includes a sit to stand task, a short walk, a turn and a return to a seated position has been correlated with measures of balance, walking speed and strength, although there is disagreement as to which physical factors the test reflects (Berg et al., 1989, Cressler et al., 1994, Podsiadlo, 1991). The individual tasks of the TUGT have been studied separately, and to some extent in combination to determine physical factors that affect performance (Hughes et al, 1994, Schenkman et al., 1996, Skeleton et al., 1996). It is hypothesized that the variability in the performance of the individual components of the test explains some of the controversy in the literature. A simple system was developed to measure the time required to complete each phase of this multiphase task. This study determined the reliability of the individual phases measured between repeated trials and between days for older adults performing the TUGT.

29. Rikli, R. J., C.J. (1998). "The reliability and validity of a 6-minute walk test as a measure of physical endurance in older adults." *Journal of Aging and Physical Activity* 6(4): 363-75.

The purpose of this study was to assess the reliability and validity of a 6-minute walk test as a measure of physical endurance in older adults. Seventy-seven subjects, ages 60-87, performed three separate 6-minute walk tests and a treadmill test and completed questionnaire items assessing physical activity level and functional status. The 6-minute walk had good test-retest reliability ($88<r<.94$), particularly when a practice trial preceded the test trial. Convergent validity of the 6-min walk was demonstrated by its moderate correlation ($71<r<.82$) with treadmill performances. Construct validity was assessed by determining the ability of the test to detect differences between different age and activity level groups. As expected, walking scores decreased significantly across decades and were significantly lower for low-active subjects compared to high-active subjects. There was a moderate relationship between 6-min walk scores and self-reported functional ability. It was concluded that the 6-min walk can be used to obtain reasonably reliable and valid measures of physical endurance in older adults and that it moderately reflects overall physical functional performance.

31. Rolland, Y. C., M; Miller, ME; Penninx, BW; Atkinson, HH; Pahor, M (2004). "Reliability of the 400-m usual-pace walk test as an assessment of mobility limitation in older adults." *Journal of the American Geriatrics Society* 52(6): 972 - 6.

To assess the test-retest reliability of the 400-m usual-pace test (400-MWT), and to determine whether the 4-m walk test predicts inability to walk 400 m. DESIGN: Observational. SETTING: Community, 20-m tract course. PARTICIPANTS: Sixty study participants (aged ≥ 65) were enrolled from the community and met the following eligibility criteria: self-reported difficulty in two or more of four functional domains (mobility and exercise tolerance, upper extremity function, basic self-care, higher functional tasks of independent living) and a score of 18 or higher on the Mini-Mental State Examination. METHODS: The 400-MWT and the 4-m walk test were each repeated within 7 days. RESULTS: The mean age \pm standard deviation of the study population was 84.3 ± 6.3 ; 88.3% were women. Nineteen participants (31.7%) failed both 400-MWTs, and 41 successfully completed both tests ($\kappa=1$). Mean walking speed for the 4-m test was 0.87 ± 0.18 m/s for those who completed the 400-MWT and 0.53 ± 0.17 m/s for those who failed ($P<.001$). The Spearman correlation coefficient between 4-m and 400-m walking speeds was 0.93. The estimated area under the receiver operating characteristic curve between 4-m walking speed and the ability to perform the 400-MWT was 0.91. The 4-m gait speed averaged less than 0.6 m/s in 80% of subjects who failed the 400-MWT. CONCLUSION: The test-retest

reliability for inability to complete the 400-MWT is high. Four-m walking speed is highly predictive of ability to perform the 400-MWT. These findings may prove useful to future clinical trials and observational studies that involve assessment of mobility limitation in older adults.

32. Shumway-Cook, A. B., S; Woolcott, M (2000). "Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test." *Physical Therapy* 80(9): 896-903.

This study examined the sensitivity and specificity of the Timed Up & Go Test (TUG) under single-task versus dual-task conditions for identifying elderly individuals who are prone to falling. SUBJECTS: Fifteen older adults with no history of falls (mean age=78 years, SD=6, range=65-85) and 15 older adults with a history of 2 or more falls in the previous 6 months (mean age=86.3 years, SD=6, range=76-95) participated. METHODS: Time taken to complete the TUG under 3 conditions (TUG, TUG with a subtraction task [TUGcognitive], and TUG while carrying a full cup of water [TUGmanual] was measured. A multivariate analysis of variance and discriminant function and logistic regression analyses were performed. RESULTS: The TUG was found to be a sensitive (sensitivity=87%) and specific (specificity=87%) measure for identifying elderly individuals who are prone to falls. For both groups of older adults, simultaneous performance of an individual task increased the time taken to complete the TUG, with the greatest effect in the older adults with a history of falls. The TUG scores with or without an additional task (cognitive or manual) were equivalent with respect to identifying fallers and nonfallers. CONCLUSION AND DISCUSSION: The results suggest that the TUG is a sensitive and specific measure for identifying community-Dwelling adults who are at risk for falls. The ability to predict falls is not enhanced by adding a secondary task when performing the TUG.

33. Smith, R. (1994). "Validation and reliability of the Elderly Mobility Scale." *Physiotherapy* 80(11): 744-7.

This study established the concurrent validity and inter-rater reliability of a simple mobility scale for frail elderly people. Concurrent validity was established by correlating the Elderly Mobility Scale (EMS) score with Barthel and functional independence measure (FIM) scores for 36 patients, age 70-93 years. Spearman's rho was 0.962 with Barthel and 0.948 with FIM -- highly significant correlations. Inter-rater reliability was established on 15 patients who were assessed independently by two physiotherapists. There was no significant difference between scores. The EMS provides physiotherapists with a standardized validated scale for assessment of frail elderly people and was designed to be used in conjunction with the assessment package recommended by the Royal College of Physicians and British Geriatric Society (1992). Further evaluation of the scale is recommended to assess its validity in other units, and to assess its predictive validity for discharge outcome and physiotherapy outcome.

34. Thomas, M. J., J; Suteerwattananon, M; Wankadia, S; Caroline, KS; Vuong, KD; Protas, E (2004). "Clinical gait and balance scale (GABS): validation and utilization." *Journal of neurological sciences* 217(1): 89-99.

Gait and Balance Scale (GABS) consists of historical information and examination of 14 different gait and balance parameters designed to assess the severity of these functional domains. Thirty-five patients with Parkinson's disease (PD), Hoehn and Yahr stages 1-3, were tested during their "off" period. GABS items were compared to qualitative data from two computerized gait analysis instruments, GAITrite and Pro Balance Master. Intra-class correlation coefficients were calculated to establish reliability. Intra-rater test re-test reliability was determined using Cohen's Kappa statistic. Concurrent validity was derived using Spearman's rho test with the items from GABS, GAITrite and Balance Master. Intra-rater reliability was high with $k > 0.41$ (k =kappa statistic) for 17 items, 6 had $k > 0.61$. When performing validity measurements, a number of items on the GABS had a correlation coefficient significant at $p < 0.01$ (2-tailed). Posture, pull test, balance during stance, single limb stance, tandem stance, turning, toe walking and functional reach had significant correlation with the Balance Master data ($R = 0.46-1$). Gait, arm swing, gait speed, steps/5 m, 'up-and-go test', modified performance oriented assessment of gait scale and provocative testing had significant correlation with the GAITrite items ($R = 0.51-0.83$). GABS is an easy-to-use comprehensive clinical scale with high intra-rater and internal item reliability. We have shown concurrent validity with two computerized gait analysis instruments. We expect GABS to have a particular utility in clinical trials designed to modify functional impairment associated with abnormalities in gait and balance.

COMPARISONS OF BALANCE AND MOBILITY ASSESSMENT TOOLS

35. Bennie, S. B., K; Dizon, A; Fritz, H; Goodman, B; Peterson, S (2003). "Measurements of Balance: comparison of the Timed "Up and Go" test and Functional Reach test with the Berg Balance Scale." *Journal of Physical Therapy Science* 15(2): 93-7.

To compare the Timed "Up and Go" and Functional Reach tests to the Berg Balance Scale for concurrent validity in a broad adult population. DESIGN: A prospective study of individuals with balance deficits. SETTING: Neurological rehabilitation, skilled nursing, and acute care facilities. SUBJECTS: Twenty subjects: 12 females and eight males, aged 38 to 86 years ($u = 68$, $SD = 14.5$). INTERVENTION: The order in which the three tests were performed on each subject was randomized and the scores from the three tests were analyzed using correlation coefficients. MAIN OUTCOME MEASURES: Balance and correlations were based on scores from each of the three balance tests performed. RESULTS: There was a significant correlation between the Berg Balance Scale and Timed "Up and Go" test ($r = 0.47$, $p = 0.04$) but no significant correlation between the Berg Balance Scale and the Functional Reach test ($r = 0.42$, $p = 0.06$). Pairing the Timed "Up and Go" and the Functional Reach tests however, revealed a significant correlation ($R = 0.56$, $p = 0.04$). CONCLUSION: This study suggests that the Timed "Up and Go" test alone or a combination of the Timed "Up and Go" and Functional Reach test can be used as a simple measure of balance comparable to the Berg Balance Scale.

36. Berg, K. M., BE; Williams, JI; Holliday, PJ; Wood-Dauphinee, SL (1992). "Clinical and laboratory measures of postural balance in an elderly population." *Archives of Physical Medicine and Rehabilitation* 73(11): 1073-80.

The objective of this cross-sectional study was to compare scores on the Berg Balance Scale with laboratory measures of postural

sway and other clinical measures of balance and mobility. Thirty-one elderly subjects were assessed on the clinical measures and the laboratory tests of postural sway while standing still and in response to pseudorandom movements of the platform. The average correlation between the Balance Scale and the spontaneous sway measures was -.55. It was slightly lower ($r=.38$) for the same parameters measured during the pseudorandom tests. There were high correlations between the Balance Scale and the Balance Sub-Scale developed by Tinetti ($r=.91$), Barthel Mobility to sub-scale ($r=.67$), and Timed "Up and Go" ($r=.76$). The Balance Scale was the most efficient measure (effect size >1) to statistically discriminate between subjects according to their use of each type of mobility aide (walker, cane, no aids). These data contribute to existing information on the performance of the Balance Scale and supports the validity of the Balance Scale in this geriatric population.

37. Steffen, T. H., TA; Mollinger, L (2002). "Age- and gender-related test performance in community-dwelling elderly people: Six-minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds." *Physical Therapy* 82(2): 128-37.

The interpretation of patient scores on clinical tests of physical mobility is limited by a lack of data describing the range of performance among people without disabilities. The purpose of this study was to provide data for 4 common clinical tests in a sample of community-dwelling older adults. SUBJECT: Ninety-six community-dwelling elderly people (61-89 years of age) with independent functioning performed 4 clinical tests. METHODS: Data were collected on the Six-Minute Walk Test (6MW), Berg Balance Scale (BBS), and Timed Up & Go Test (TUG) and during comfortable - and fast-speed walking (CGS and FGS). Intraclass correlation coefficients (ICCs) were used to determine the test-retest reliability for the 6MW, TUG, CGS, and FGS measurements. Data were analyzed by gender and age (60-69, 70-79), and 80-89 years) cohorts, similar to previous studies. Means, standard deviations, and 95% confidence intervals for each measurement were calculated for each cohort. RESULTS: The 6MW, TUG, CGS, and FGS measurements showed high test-retest reliability (ICC[2,1]=.95-.97). Mean test scores showed a trend of age-related declines for the 6MW, BBS, TUG, CGS, and FGS for both male and female subjects. DISCUSSION AND CONCLUSION: Preliminary descriptive data suggest that physical therapists should use age-related data when interpreting patient data obtained for the 6MW, BBS, TUG, CGS, and FGS. Further data on these clinical tests with larger sample sizes are needed to serve as a reference for patient comparisons.

38. Whitney, S. P., JL; Cass, SP (1998). "A review of Balance instruments for older adults." *American Journal of Occupational Therapy* 52(8): 666-71.

The purpose of this article is to review balance instruments developed within the past 10 years that can be used in the clinic or home environment. The use of such instruments may assist in identifying older adults who are at risk for falling, a major problem that can result in impaired function and loss of independence. METHOD: Six instruments were reviewed: the Berg Balance Scale (Berg), the Clinical Test of Sensory Interaction and Balance (CTSIB), the Functional Reach Test, the Tinetti Balance Test of the Performance-Oriented Assessment of Mobility Problems (Tinetti), the Timed "Up and Go" Test (TUG), and the Physical Performance Test (PPT). Considered were what aspects of balance are assessed, time needed to administer the instrument, tools or equipment needed, evidence of reliability and validity, advantages and disadvantages, and the target population. RESULTS: The Berg, Tinetti, and PPT measure a variety of aspects of balance, whereas the Functional Reach, TUG, and CTSIB measure narrow aspects of balance. All six instruments have been used with older adults and do not require much equipment. The instruments differ in their reliability and validity. CONCLUSION: Familiarity with balance instruments can be helpful in selecting the one most appropriate for clinical setting and clients in order to institute appropriate prevention programs, such as environmental modifications and lifestyle adaptations.

39. Whitney, S. W., D; Furman, J (2003). "Concurrent validity of the Berg Balance Scale and the Dynamic Gait Index in people with vestibular dysfunction." *Physiotherapy Research International* 8(4): 178-86.

The Berg Balance Scale is a reliable and valid measure that is used to assess characteristics of balance. The Dynamic Gait Index is a relatively new measure that has been used to record dynamic gait tasks in people with vestibular dysfunction. The purpose of the present study was to determine the concurrent validity of the Dynamic Gait Index with the Berg Balance Scale in people with vestibular disorders. METHOD: A retrospective review of the charts of people who met the criteria of having completed both the Berg Balance Scale and the Dynamic Gait Index during their first physiotherapy visit. Seventy patients (19 male, 51 female) were identified through retrospective review of the charts of people referred to vestibular rehabilitation with varying diagnoses of vestibular and balance dysfunction. All were seen at a tertiary medical centre in an outpatient physiotherapy setting. Their age range was from 14 to 88 years (mean 65 years). RESULTS: Correlation between the scores on the Dynamic Gait Index and the Berg Balance Scale was moderate but significant by use of the Spearman rank order correlation ($r=0.71$; $p<0.01$). No difference was found between scores on the Dynamic Gait Index or Berg Balance Scale based on gender or diagnosis. A significant difference was identified on the Berg Balance Scale between older and younger people with vestibular disorders. Using previously established criteria to determine increased risk of falling, the Berg Balance Scale and the Dynamic Gait Index agreed 63% of the time. CONCLUSIONS: The moderate correlation between the Dynamic Gait Index and the Berg Balance Scale establishes the concurrent validity of the Dynamic Gait Index in people with vestibular dysfunction. Both these measures provide valuable information to clinicians about patients' functional balance capabilities. However, the lack of perfect correlation indicates that the tests measure different aspects of balance. The Dynamic Gait Index appears to be a more sensitive assessment tool in identifying people with vestibular disorders who are at increased risk for falling, based on currently published criteria.

FUNCTIONAL ASSESSMENT TOOLS

40. Edwards, M. (1990). "The reliability and validity of self-report activities of daily living scales (Physical Self-Maintenance Scale and Instrumental Activities of Daily Living Scale)." *Canadian Journal of Occupational Therapy* 57(5): 273-8.

A measurement evaluation study was conducted to examine the reliability and validity of the Physical Self-Maintenance Scale and Instrumental Activities of Daily Living Scale (Lawton&Brody, 1969) as a discriminative measure to assess the independent living

status of elderly hospitalized patients. A sample of thirty patients from two geriatric specialty units was used for the study. Inter-observer reliability was shown to be excellent (ADL $r=.96$, IADL $r=.99$) and test-retest reliability was considered to be good (ADL $r=.59$, IADL $r=.93$). The results comparing patient self-report of ability and ratings based on direct observation of performance indicated that patients consistently over-rated their ability in both ADL and IADL tasks. This provides support for the use of direct observation of performance to obtain valid data on an individual's ability. There was a significant difference in scores between those patients able to be discharged home compared to those requiring institutionalization; however, a score predictive of discharge location could not be identified with the study data. The study provided support for the use of this measurement tool for assessing the independent living status of elderly hospitalized patients. It provides the therapist with a measure that has had some validation and has the potential for clinical use in describing patients' ability and identifying areas requiring treatment intervention.

41. Hobart, J. T., AJ (2001). "The five item Barthel Index." Journal of Neurology, Neurosurgery & Psychiatry 71(2): 225-30.

OBJECTIVES: Routine data collection is now considered mandatory. Therefore, staff rated clinical scales that consist of multiple items should have the minimum number of items necessary for rigorous measurement. This study explores the possibility of developing a short form Barthel index, suitable for use in clinical trials, epidemiological studies, and audit, that satisfies criteria for rigorous measurement and is psychometrically equivalent to the 10 item instrument. **METHODS:** Data were analyzed from 844 consecutive admissions to a neurological rehabilitation unit in London. Random half samples were generated. Short forms were developed in one sample ($n=419$), by selecting items with the best measurement properties, and tested in the other ($n=418$). For each of the 10 items of the BI, item total correlations and effect sizes were computed and rank ordered. The best items were defined as those with the lowest cross product of these rank orderings. The acceptability, reliability, validity, and responsiveness of three short form BIs (five, four and three item) were determined and compared with the 10 item BI. Agreement between scores generated by short forms and 10 item BI was determined using intraclass correlation coefficients and the method of Bland and Altman. **RESULTS:** The five best items in this sample were transfers, bathing, toilet use, stairs, and mobility. Of the three short forms examined, the five item BI had the best measurement properties and was psychometrically equivalent to the 10 item BI. Agreement between scores generated by the two measures for individual patients was excellent (ICC=0.90) but not identical (limits of agreement=1.84+/-3.84) **CONCLUSIONS:** The five item short form BI may be a suitable outcome measure for group comparison studies in comparable samples. Further evaluations are needed.

42. Ottenbacher, K. J. P. H., Yungwen MS; Granger, Carl V. MD; Fiedler, Roger C. PhD (1996). "The Reliability of the Functional Independence Measure: A Quantitative Review." Archives of Physical Medicine & Rehabilitation 77: 1226-1232.

Objective: The reliability of the Functional Independence Measure (FIM) for adults was examined using procedures of meta-analysis. **Data Sources:** Eleven published studies reporting estimates of reliability for the FIM were located using computer searches of Index Medicus, Psychological Abstracts, the Functional Assessment Information Service, and citation tracking. **Study Selection:** Studies were identified and coded based on type of reliability (interrater, test-retest, or equivalence), method of data analysis, size of sample, and training or experience of raters. **Data Extraction:** Information from the articles was coded by two independent raters. Interrater reliability for coding all elements included in the analysis ranged from .89 to 1.00. **Data Synthesis:** The 11 investigations included a total of 1,568 patients and produced 221 reliability coefficients. The majority of the reliability values (81%) were from interrater reliability studies, and the intraclass correlation coefficient (ICC) was the most commonly used statistical procedure to compute reliability. The reported reliability values were converted to a common correlation metric and aggregated across the 11 studies. The results revealed a median interrater reliability for the total FIM of .95 and .92 respectively. The median reliability values for the six FIM subscales ranged from .95 for Self-Care to .78 for Social Cognition. For the individual FIM items, median reliability values varied from .90 for Toilet Transfer to .61 for Comprehension. Median and mean reliability coefficients for FIM motor items were generally higher than for items in the cognitive and communication subscales. **Conclusions:** Based on the 11 studies examined in this review the FIM demonstrated acceptable reliability across a wide variety of settings, raters, and patients

43. Shelkey, M. R., MSN, PhD(cand); Wallace, Meredith PhD, RN, MSN, CS (1998). "Katz Index of Independence in Activities of Daily Living (ADL)." Best Practices in Nursing Care to Older Adults - The Hartford Institute for Geriatric Nursing(2).

The Katz ADL tool assesses basic activities of daily living. It does not assess more advanced activities of daily living. Katz developed another scale for instrumental activities of daily living such as heavy housework, shopping, managing finances and telephoning. Although the Katz ADL index is sensitive to changes in declining health status, the tool is limited in its ability to measure small increments of change seen in the rehabilitation of older adults. A full comprehensive geriatric assessment should follow when appropriate. The Katz inventory is very useful in creating a common language about patient function for all practitioners involved in overall care planning and discharge planning.

44. Wade, D. C., C (1988). "The Barthel ADL Index: a standard measure of physical disability?" International Disability Studies 10(2): 64-7.

There is no agreed single measure of physical disability for use either clinically or in research. It is argued that acceptance of a single standard measure of activities of daily living (ADL) might increase awareness of disability, improve clinical management of disabled patients, and might even increase acceptance of published research. The Barthel ADL Index is proposed as the standard index for clinical and research purposes. Its validity, reliability, sensitivity, and utility are discussed. The Barthel Index is as good as any other single index, and should be adopted as the standard against which future indices are compared. The temptation to use variations on the standard Barthel Index should be resisted.

COMPARISONS OF FUNCTIONAL ASSESSMENT TOOLS

45. Fricke, J. U., CA (1997). "Inter-rater reliability of the original and modified Barthel Index, and a comparison with the Functional Independence Measure." Australian Occupational Therapy Journal 44(1422-9): March 1997.

The Barthel Index (BI), the Modified Barthel Index (MBI) and the Functional Independence Measure (FIM) are all widely used by occupational therapists as assessment tools for clinical decision-making and outcome measurement. All of these tools have demonstrated validity and the BI and the FIM have demonstrated inter-rater reliability. The MBI has been modified to increase sensitivity; however, there have been no publications on the inter-rater reliability of this tool following changes. The purpose of this research was to examine the inter-rater reliability of two versions of the Barthel Index, and draw some comparisons between this assessment tool and the FIM. Twenty-five patients with neurological and orthopaedic conditions were assessed by three occupational therapists using the three tools. The method of analysis selected was percentage agreement and intraclass correlation coefficient. The results indicated that both the original and modified versions of the Barthel Index possess good inter-rater reliability. As all three tools have demonstrated adequate reliability and validity, it is suggested that clinicians select the most sensitive tool that best meets their clinical needs, and use this assessment tool in its standard format.

46. Gosman-Hedstrom, G. S., E (2000). "Parallel reliability of the Functional Independence Measure and the Barthel ADL Index." Disability & Rehabilitation 22(16): 702-15.

The aim of this study was to evaluate the concordance between the assessment of ADL according to the Functional Independence Measure (FIM) and the Barthel ADL Index (BI) by means of a rank-invariant statistical method. METHOD: The construct validity, also called parallel reliability, of FIM and BI was assessed on the item level. Two different approaches to condensing the FIM assessments into 2-4 scale steps on the item level in order to calibrate the two ADL instruments were compared. One was determined by the theoretical operational definitions and the other defined by empirical definitions. The 204 assessments of elderly persons were made three months after stroke by means of interviews. RESULTS: The parallel reliability of the FIM and the BI on the item level was strong, both according to the theoretical cut-off levels defined by the operational definitions and the empirical cut-off levels defined by the marginal distributions. CONCLUSIONS: The concordance between the FIM and the BI was high. There was a slight difference in favour of the operational definitions. The clinical key elements are to have a critical attitude towards ADL instruments; how they are constructed and operationalized. These are important elements in the quality assurance in the everyday work.

NUTRITIONAL ASSESSMENT TOOLS

47. Group, M. A. (2003). "A consistent and reliable tool for malnutrition screening." Nursing Times 99(46): 26-7.

Malnutrition is both preventable and treatable and yet it continues to undermine the health of a significant proportion of the UK population. The Malnutrition Advisory Group (MAG) has launched a screening tool for use by all health professionals to screen patients with any disease or condition. The malnutrition universal screening tool (MUST) is endorsed by the British Dietetic Association, the RCN and the Registered Nursing Homes Association. These groups are now actively working with MAG to develop appropriate professional training opportunities.

48. Lawrence, J. A., EJ (2004). "Assessing Nutrition in older Adults -MNA." Best Practices in Nursing Care to Older Adults - The Hartford Institute for Geriatric Nursing(9).

While poor nutrition is not a natural concomitant of aging, older adults who experience several concurrent diseases are at higher risk for under- or malnutrition. Persons who are underweight (Body Mass Index<19) and those who are overweight (Body Mass Index >25) often have loss of muscle mass, a compromised immune system and have increased complications and premature death. The progression of malnutrition is often insidious, and is often undetected. The Mini Nutritional Assessment (MNA) is both a screening and assessment tool for the identification of malnutrition in the older age. This tool eliminates the need for invasive tests such as blood sampling. The MNA has been validated in many research studies in older adults throughout the world in hospital, nursing home and ambulatory care patients and in community screening. Internal consistency, inter-observer reliability and validity were shown to be acceptable (beck, Overson, & Schroll, 2001; Bleda, Bolibar, Pares, & Salva, 2002)

49. Russell, L. T., J; Brewitt,J; Ireland,M; Reynolds,T (1998). "Development and validation of the Burton Score: a tool for nutritional assessment." Journal of Tissue Viability 8(4): 16-22.

This paper describes the development and validation of the Burton Score, a nutritional assessment tool based on the Waterloo score, with the rationale that since nurses already collect data for one score, it would only lead to unnecessary duplication of effort if a totally different scoring scheme were to be used for nutritional assessment. Initial cut off's were determined by a pilot study of 26 patients on an elderly care ward and validated by comparing the nutritional status of 263 patients estimated by the Burton Score with a dietitian's assessment of nutrition. The validation study showed that although there was significant correlation between the Burton and Waterloo scores the Burton score correlated more closely with the dietitian's assessment. The King's Fund report of 1992 stated that all patients should have assessment of nutritional status on admission to hospital; we believe the Burton score could provide a simple tool to achieve this goal

PAIN ASSESSMENT TOOLS

50. Cleeland, C. R., KM (1994). "Pain assessment: global use if the Brief Pain Inventory." Annals of the Academy of Medicine Singapore 23(2): 129-38.

Poorly controlled cancer pain is a significant public health problem throughout the world. There are many barriers that lead to undertreatment of cancer pain. One important barrier is inadequate measurement and assessment of pain. To address this problem, the Pain Research Group of the WHO Collaborating Centre for Symptom Evaluation in Cancer Care has developed the Brief Pain Inventory (BPI), a pain assessment tool for use with cancer patients. The BPI measures both the intensity of pain (sensory dimension) and interference of pain in the patient's life (reactive dimension). It also queries the patient about pain relief, pain quality, and patient perception of the cause of pain. This paper describes the development of the Brief Pain Inventory and the various applications to which the BPI is suited. The BPI is a powerful tool and, having demonstrated both reliability and validity across cultures and languages, is being adopted in many countries for clinical pain assessment, epidemiological studies, and in studies of the effectiveness of pain treatment.

51. Ferrell, B. S., WM; Beck, JC (2000). "The Geriatric Pain Measure: validity, reliability and factor analysis." *Journal of American Geriatrics Society* 48(12): 1669-73.

Pain is a multidimensional experience that should be evaluated beyond as estimate of intensity. A multidimensional pain measure has not been developed for older persons undergoing comprehensive geriatric assessment. **OBJECTIVE:** To develop and evaluate validity and reliability of a multidimensional pain assessment instrument for older persons. **RESEARCH DESIGN:** A series of steps in instrument development and evaluation. **SUBJECTS:** A total of 176 subjects (mean age 84 \pm 6.0 years) in ambulatory geriatric clinics; 64% were women, and 73% had a history of chronic pain. **MEASUREMENTS:** Measurement included the Geriatric Pain Measurement (GPM), the McGill Pain Questionnaire, Yesavage GDS, Katz ADLs, Lawton IADLs, Tinetti Gait and Balance, Folstein MMSE, and other demographic and clinical characteristics from interview and chart review. **RESULTS:** The GPM demonstrated a standardized alpha = 0.9445, homogeneity ratio = 0.457, and average inter-item correlation = 0.415. A subgroup of 50 subjects demonstrated concurrent validity of the GPM in comparison with the McGill Pain Questionnaire (Pearson's r correlation 0.6296 ($P < .0000$)). Test-retest reliability was demonstrated in another subgroup of 50 subjects who repeated the GPM within 48 to 72 hours (Pearson's $r = 0.9018$; $P < .0000$). Factor analysis revealed five clusters of components: Pain Intensity, Disengagement, Pain with Ambulation, Pain with Strenuous Activities, and Pain with Other Activities. **CONCLUSIONS:** The GPM is a 24-item questionnaire that is easy to administer and has significant validity and reliability in older persons with multiple medical problems. The GPM may be a useful addition to the multidimensional geriatric assessment process.

52. Herr, K. M., PR; Kohout, FJ; Wagenaar, D (1998). "Evaluation of the Faces Pain Scale for use with the elderly." *Clinical Journal of Pain* 14(1): 29-38.

The specific objective for this research was to determine initial psychometric properties of the Faces Pain Scale (FPS) as a measure of pain intensity for use with the elderly. **DESIGN:** The study was descriptive correlational in nature, with nonrandom sampling. A total sample of 168 community subjects (30-121, depending on task completed), aged 65 or older, participated in the research protocol. To determine validity, reliability, and scaling properties of the FPS, rating and ranking procedures, placement tasks, and test-retest methods were used. **RESULTS:** Response to six Likert-type items indicated that subjects agreed that the FPS represents pain; however, it is clear that the perception of the meaning of the faces can be influenced by the context in which they are presented. Rank ordering tasks for the individual faces demonstrated near-perfect agreement between the actual expected ranking and the ranking produced by the subjects (Kendall's $W = .97$, $p = .00$). When subjects placed individual faces along a 1-m-long red wedge indicating the amount of pain represented by each face, statistically significant separation of the faces in the anticipated equal interval position was demonstrated by the lack of overlap of the 95% confidence intervals when all faces were viewed and positioned simultaneously. However, when subjects placed faces independent of others, the expected placement fell outside the 95% confidence limit for three of the five faces placed. In addition, the actual intervals between the five faces placed by subjects demonstrated substantial variances from the 167 mm expected in several instances. Rating a vividly remembered painful experience about the degree of pain perceived using the FPS initially and again 2 weeks later, the FPS demonstrated strong reproducibility over time with a Spearman rho correlation coefficient of .94 ($p = .01$). **CONCLUSION:** These results provide preliminary support for the construct validity, strong ordinal properties, and strong test-retest reliability of the FPS with a sample of elderly individuals. The equality of intervals in the FPS has not been fully supported in the older adult, but given the complexity of the task used, the results should not be considered to be refuted. Further evaluation of the FPS with experimental and clinical pain conditions and comparison with other standard pain assessment instruments in the elderly population are warranted.

53. Keller, S. B., CM; Dodd, SL; Mendoza, TR; Cleeland, CS (2004). "Validity of The Brief Pain Inventory for use in documenting the outcomes of patients with noncancer pain." *Clinical Journal of Pain* 20(5): 309-18.

The Brief Pain Inventory (BPI) is a short, self-administered questionnaire that was developed for use in cancer patients. While most empirical research with the BPI has been in pain of that etiology, the questionnaire is increasingly evident in published studies of patients with non-cancer pain. The current research addressed the need for formal evaluation of the reliability and validity of the BPI for use in non-cancer patients. **METHODS:** Approximately 250 patients with arthritis or low back pain (LBP) self-administered a number of generic and condition-specific health status measures (including the BPI) in the clinic of their primary care provider at 2 time points: the initial clinic visit and the first visit following treatment. **RESULTS:** The reliability of the BPI data collected from non-cancer pain patients was comparable to that reported in the literature for cancer patients and sufficient for group-level analyses (coefficient alphas were greater than 0.70). The factor structure of the BPI was replicated in this sample and the relationship of the BPI to generic measures of pain was strong. The BPI exhibited similar relationships to general and condition-specific measures of health as did a generic pain scale (SF-36 Bodily Pain). Finally, the BPI discriminated among levels of condition severity and was sensitive to change in condition over time in arthritis and LBP patients. **DISCUSSION:** Results support the validity of the BPI as a measure of pain in patients without cancer and, in particular, as a measure of pain for arthritis and LBP patients.

54. Tan, G. J., MP; Thornby, JI; Shanti, BF (2004). "Validation of the Brief Pain Inventory for chronic nonmalignant pain." *Journal of Pain* 5(2): 133-7.

The Brief Pain Inventory (BPI; Cleeland and associates) has been used primarily to assess patients with cancer-related pain. Although it has been validated in many languages and is widely used, there has not yet been research published to validate its use for patients with chronic nonmalignant pain as the primary presenting problem. This study was designed to fill this gap by examining the psychometric properties of the BPI in 440 patients with chronic intractable pain referred to a chronic pain clinic at a metropolitan tertiary-care teaching hospital. Results indicated acceptable internal consistency (Cronbach alpha coefficients were .85 for the intensity items and .88 for the interference items). A factor analysis resulted in 2 distinct and independent factors, supporting the validity of the 2-factor structure of the BPI. Zero-order correlations indicated that the association with a measure of disability (the Roland-Morris Disability Questionnaire [RMDQ]) was significantly higher for BPI interference ($r=0.57$) than for BPI intensity ($r=0.40$, $t=5.71$, $P<.01$) and that the correlation with BPI interference was not more than 0.80, supporting the conclusion that these scales assess related, but also distinct, dimensions. Finally, the finding that both BPI scales showed statistically significant improvement with treatment confirms the responsiveness of BPI in detecting and reflecting improvement in pain over time. PERSPECTIVE: This paper validated the psychometric properties of a pain Assessment instrument (The Brief Pain Inventory) originally developed to assess cancer pain and extended its use for the chronic nonmalignant pain population. This provides an important and widely used diagnostic tool for the clinician treating chronic pain.

COMPARISONS OF PAIN ASSESSMENT TOOLS

55. Kaasalainen, S. C., J (2003). "A comparison of pain-assessment tools for use with elderly long-term-care residents." Canadian Journal of Nursing Research 35(4): 58-71.

The purpose of this study was to examine the psychometric properties (test-retest and interrater reliability, criterion concurrent validity) of 3 verbal pain-assessment tools (Faces Pain Scale, Numerical Rating Scale, Present Pain Intensity Scale) and a behavioral pain-assessment scale for use with an elderly population. The study used a repeated-measures design to examine the reliability and validity of the tools across 4 groups of participants with varying levels of cognitive impairment using a non-random stratified sample of 130 elderly long-term-care residents. The findings support the test-retest and interrater reliability of the behavioral pain-assessment tool across all levels of cognitive impairment, whereas the same measure of reliability for the verbal-report tools decreased with increasing cognitive impairment; however, the majority of elderly with mild to moderate cognitive impairment were able to complete at least 1 of these tools. The findings are discussed in relation to their clinical and research implications.

PERCEPTUAL ASSESSMENT TOOLS

56. Boys, M. F., P; Holzberg, C; Reid, DW (1988). "The OSOT Perceptual Evaluation: a research perspective." American Journal of Occupational Therapy 42(2): 92-8.

Although the Ontario Society of Occupational Therapists (OSOT) Perceptual Evaluation has been widely used, it has never been standardized. A study was undertaken to examine the validity of the battery for differentiating neurologically normal persons from those who have been independently diagnosed as neurologically impaired. A group of 80 brain-damaged patients was compared with a matched group of 70 neurologically normal persons. Comparison scores for the two groups supports the validity of the instrument for differentiating the neurologically normal from the perceptually impaired person. The distribution of scores suggests that the degree of impairment can be classified as mild, moderate, or severe. Finally, the OSOT Perceptual Evaluation is found to be a reliable procedure for the assessment of perceptual dysfunction.

SKIN ASSESSMENT

57. Halfens, R. V. A., T; Bal, RM. (2000). "Validity and reliability of the Braden Scale and the influence of other risk factors: a multi-centre prospective study." International Journal of Nursing Studies 37(4): 313-9.

The Braden scale is one of the most intensively studied risk assessment scales used in identifying the risk of developing pressure sores. However, not all studies show that the sensitivity and specificity of this scale is sufficient. This study, therefore, investigated whether adding new risk factors can enhance the sensitivity and specificity of the Braden scale. The Braden scale was tested in a prospective multi-centre design. The nurses of 11 wards filled in the Braden scale every 5 days for each patient who was admitted without pressure sores and who had a probable stay of at least 10 days. Based on a literature study and in-depth interviews with experts, the Braden scale was extended by the risk factor blood circulation. In addition, other risk factors, which are more or less stable patient characteristics, were measured during the admission of the patient. Independent research assistants measured the presence of pressure sores twice a week. As the external criterion for the risk of developing pressure sores, the presence of pressure sores and/or the use of preventive activities was used. Results showed that the original Braden scale was a reliable instrument and that the sensitivity and specificity was sufficient. However, reformulating the factors moisture and nutrition, and adding the risk factor age could enhance the sensitivity and specificity. Furthermore, results showed that the factors sensory perception, and friction and shear were especially important risk factors for the Braden scale. In fact, using only the factors sensory perception, friction and shear, moisture (a reformulated factor) and age give the highest explained variance of the risk of developing pressure sores. The added risk factor blood circulation, did not enhance the sensitivity and specificity of the original Braden scale. Suggestions are given on how to use risk assessment scales in practice.

SWALLOWING Assessment

58. Lambert HC, G. E., Groher ME, Wood-Dauphinee S. (2003). "McGill Ingestive Skills Assessment (MISA): development and first field test of an evaluation of functional ingestive skills of elderly persons." Dysphagia 18(2): 101-13.

There is a lack of reliable and valid clinical assessment tools for individuals with loss of ingestive skills. The McGill Ingestive Skills Assessment (MISA) was developed to facilitate the reliable and valid bedside assessment of elderly persons with feeding difficulties. Items were generated by a literature review and selected with the collaboration of a multidisciplinary team. The first version of the MISA comprised 190 items in 7 scales, covering the domains of medical history, mealtime environment, physical characteristics of the patient, food textures consumed, solid ingestion, liquid ingestion, and behaviors related to self-feeding. The first field test for item selection included 50 individuals, aged 60 years and older, living in the community, supervised housing, and long-term care centers. After field testing, 134 items were eliminated due to poor face validity, redundancy, or poor psychometric performance. The remaining 56 items were provided with 4 response categories and were reorganized into 5 scales. The revised version was field tested to determine its preliminary psychometric properties on 33 individuals, 60 years of age and older, residing in a long-term care center. Six items were eliminated due to redundancy after the second field test. Analyses of the revised version resulted in the elimination of another 6 items that were redundant or that demonstrated poor reliability. Internal consistency of all scales is ≥ 0.86 and interrater agreement is ≥ 0.92 . These analyses suggest that the psychometric properties of the MISA are adequate for diagnosis and treatment planning. This supports its readiness for clinical use following further reliability and validity testing with a larger sample.

OTHER Assessments

59. Lane, R. G., WM; Hansen, TE; Berman, WH; Kramer, SI (1985). "Assessment of tardive dyskinesia using the Abnormal Involuntary Movement Scale." *Journal of Nervous and Mental Disorders* 173(6): 353-7.

Over a 10 month period, 33 patients with tardive dyskinesia (TD) were evaluated with the Abnormal Involuntary Movement Scale (AIMS) simultaneously and independently by two experienced and two inexperienced raters. The experienced raters generally had higher levels of agreement and their scores were more consistent over time. It is concluded that experience with TD influences AIMS inter-rater reliability and that it is useful to differentiate TD movements into the dimensions of quality, frequency and amplitude, dimensions not currently used in the AIMS. The usefulness and difficulty of developing more specific guidelines for AIMS ratings are discussed.

60. Lanska, D. G., CG. (2000). "Romberg's sign: development, adoption, and adaptation in the 19th century." *Neurology* 55(8): 1201-6.

In the first half of the 19th century, European physicians - including Marshall Hall, Moritz Romberg and Bernardus Brach - described loss of postural control in darkness of patients with severely compromised proprioception. Romberg and Brach emphasized the relationship between the sign and tabes dorsalis. Later, other neurologists evaluated the phenomenon in a broader range of neurological disorders using a variety of simple but increasingly precise and sensitive clinical tests. Although now known as Romberg's sign, among neurologists in the late 19th century this phenomenon was sometimes credited to Romberg, sometimes to both Brach and Romberg, and sometimes discussed without attribution.