1-1-2000

A core group of standardized tests to fulfill the clinical needs of entry-level physical therapy students

Dustin Joder

NORTHERN ILLINOIS UNIVERSITY

A Core Group of Standardized Tests to Fulfill the Clinical Needs of Entry-Level Physical Therapy Students

A Thesis Submitted to the

University Honors Program

In Partial Fulfillment of the

Requirements of the Baccalaureate Degree

With Upper Division Honors

Department of Allied Health

By Dustin Joder

Dekalb, Illinois

May 13,2000

University Honors Program

Capstone Approval Page

Capstone Title: (print or type):

A Core Group of Standardized Tests to Fulfill the Clinical Needs of Entry-Level Physical Therapy Students

Student Name (print or type): Philip Joder

Faculty Supervisor (print or type): Dr. Elizabeth Kay

Faculty Approval Signature: Elizabeth Kay PT, PhD

Department of (print or type): Physical Therapy

Date of Approval (print or type): 4-21-08

HONORS THESIS ABSTRACT
THESIS SUBMISSION FORM

AUTHOR: Dustin Joder

THESIS TITLE: A Core Group of Standardized Tests to Fulfill the Clinical Needs of Entry-Level Physical Therapy Students

ADVISOR: Elizabeth Kay, PT, PhD ADVISOR DEPT: AHPT

DISCIPLINE: Physical Therapy YEAR: 2000

PAGE LENGTH: 14 BffiLIOGRAPHY: 5 pages ILLUSTRATED: No

PUBLISHED: No

COPIES AVAILABLE: Hard Copy

ABSTRACT:

During the past decade, there has been a growing emphasis on, and demand for, outcome information in medicine. A dedicated effort to utilize standardized outcome measures would have multiple benefits for the physical therapy profession and the patients who seek physical therapy treatment. In order for the physical therapy profession to make a dedicated effort to utilize standardized outcome measures, physical therapy students need to learn the rationale, administration, and interpretation of a core group of standardized tests that they can apply to meet a broad range of clinical needs. The purpose of this paper is to suggest a core group of standardized tests that would offer entry-level physical therapy students the tools necessary to obtain objective measures for a wide range of patients and clinical purposes. A core group of twenty-four tests are suggested, and although the tests may not be definitively superior measures, the core group of tests meet the clinical needs of entry-level physical therapy students and are an important step towards increasing utilization of standardized tests in the physical therapy profession.

Definition of Standardized Tests

A standardized test is defined as a published measurement tool designed for a specific purpose in a given population with detailed instructions provided as to when and how it is to be administered and scored, interpretation of the scores, and results of investigations of reliability and validity.[1] If the physical therapy profession is to make a dedicated effort to utilize standardized outcome measures, physical therapy students need to learn the rationale, administration, and interpretation of a core group of standardized tests that they can apply to meet a broad range of clinical needs. The purpose of this paper is to suggest a core group of standardized tests that would offer entry-level physical therapy students the tools necessary to obtain objective measures for a wide range of patients and clinical purposes.

Uses of Standardized Tests

Over the past decade, there has been a growing emphasis on, and demand for, outcome information in medicine.[2] Therefore, a dedicated effort to utilize standardized outcome measures would have multiple benefits for the physical therapy profession and the patients who seek physical therapy treatment. Standardized tests can be used to identify impairments, functional limitations, and disabilities, to measure wellness, to predict the outcome of an intervention, to measure change over time, to measure the efficacy and efficiency of interventions, to discriminate normal from abnormal development, to do quality assessments of a physical therapist or physical therapy department, and to justify reimbursement.

Identifying impairments, functional limitations, and disabilities is useful for helping physical therapists to determine and/or quantify deficits so that a suitable intervention can be planned and monitored. Scores from tests making measurements at different levels of the disablement model (i.e., pathology, impairment, functional limitation, and disability) can be used to gather a detailed

picture of each patient's deficits and illustrate the impact, if any, of impairments upon the more significant issues of function and health. The ability to identify and quantify deficits at a variety of levels could adequately support the use of standardized tests, but other advantages make their use even more compelling. Tests that use a questionnaire format in simple direct language can be mailed to the patient or administered over the phone so that the therapist has access to data that will help him/her to focus the physical therapy examination on key problem areas, or even determine if the person would benefit from a more detailed examination. This ability to screen new patients, or to do a follow-up examination at low cost and over great distances, has multiple implications for dealing with people who have limited access to health care as well as for gathering long-term outcome data for research or service planning purposes. These standardized test results can then be used to monitor the intervention's effect.

Another use for standardized tests is that some are designed to provide therapists with predictions of functional outcomes, which are useful for rehabilitation team planning and patient/family education. Rehabilitation teams use the information for setting discharge time frames, determining discharge destinations, and determining the level of assistance that the patient will require after discharge. In addition, test results can be used to educate the patient and family about the patient's future needs as well as any progress or decline that may occur over time. When therapists recommend that patients and their families make life altering decisions, such as placement in an extended care facility or investment of substantial financial resources in equipment or environmental adaptations, presenting objective data about the patients current functional level and expected functional outcome should help make the therapist's recommendation more informative and objective to the patient, family members, physicians, and third party payers.

Utilizing standardized tests to objectively measure change over time allows clinicians to measure the efficacy of an intervention for an individual. This is an efficient and objective way to measure if a program is effecting the desired change or if improvement is even possible. When these data are pooled for a group then this same information can be used to measure the efficiency, and even the efficacy, of physical therapy practices. Physical Therapy departments may use the measures to evaluate the cost of resources (e.g., treatment units, visits, length of stay) needed to achieve an outcome, and/or the quality of care delivered by individual clinicians or the entire physical therapy department. . Quantifying the quality of care provided by an individual therapist can be helpful when doing job performance appraisals, while measuring a department's quality can be useful for marketing, increasing referrals, and increasing the use of "best practices".

Measures of treatment efficacy can also help improve the quality of care by identifying superior practices. Distinct practice variations occur throughout the physical therapy profession, and this phenomenon can be blamed at least in part on the absence of a large enough body of objective data to support or refute the efficacy of our practices. If standardized tests were used more systematically, then physical therapists could help build the body of objective data that is necessary to identify superior practices and decrease practices that do not benefit patients.

Reimbursement for physical therapy services, like all medical services, has been in decline.[3] However, utilizing standardized tests should give clinicians the objective data they need to justify reimbursement for services that improve function. As utilization of standardized tests increases, clinicians would be able to support the efficacy of selected interventions and provide third party payers with objective measures of improved function that should support reimbursement. . Until the members of the physical therapy profession increase their utilization of

standardized tests, therapists will continue to find themselves having to react to reimbursement changes rather than armed to influence, or even stop, changes in reimbursement policies. [2]

Standardized tests can also be helpful in discriminating normal from abnormal development and measuring wellness. Discriminating normal from abnormal development is especially critical for identifying children under seven years of age whom would benefit from special education services, which would include physical therapy. Developmental tests can also be used to monitor an intervention for these infants and children. Since the promotion of wellness is an important goal for all healthcare professionals, an objective measure of wellness is also beneficial. Quantifying wellness can help measure decline in people with degenerative diseases or determine a person's attitude toward disability. For example some patients have been found to believe that the presence of impairment means that they are "sick", resulting in their tendency to avoid activity. [9, 10] While measuring a decline in health could justify a need for professional intervention such as medical care or further functional training, quantifying a low wellness score in a patient expected to have better health can justify interventions to improve patient education regarding wellness interventions that could improve their functional level and/or reduce their risk of future health problems that would impair their function. For example, someone with a BMI >30, who according to the NIlI has a greater risk for cardiovascular disease, may be more motivated to follow a program of healthier eating and exercise if their risk is quantified than if he/she is merely directed to lose a few pounds. [11] Standardized test scores can aid patient education by quantifying risk or performance.

Criteria for Selection

Several factors need to be taken into consideration when selecting a core group of standardized tests that will most effectively fulfill the clinical needs of entry-level physical therapy students. To assure all of the important factors are considered when selecting the core group of tests, the selected tests when taken together should:

1. demonstrate acceptable reliability and validity
2. require reasonable time and practice to learn
3. be cost effective to administer (e.g., time and equipment)
4. be sensitive to change, with minimal ceiling or basement effects
5. have been standardized on diverse socioeconomic and racial populations
6. be applicable to people with a broad range of diagnoses
7. allow measurement in people from a variety of age groups
8. allow measurement at all levels of the disablement model
9. include tests used for purposes other than measurement at different levels of the disablement model (e.g., developmental assessment)
10. include tests that are useful for predicting expected functional outcomes
11. measure progress and decline
12. be commonly used in the university's clinical sites and geographical region

Many of these criteria are self-explanatory, but a bit of justification may clarify the importance of each. While all standardized tests must by definition have some degree of reliability and validity, the criteria that was adopted for the purposes of developing the core group of tests to be taught to entry-level PT students was that the tests have interrater and intrarater reliability of ICC ~ .95, r ~ .95, or P 5.05, and have established construct validity. These statistical standards are demanding, but because of the impact test results can have, the selected tests should meet the standards to be most valuable for clinical application. Although not every test will meet these standards, or will have reported psychometric properties, these tests may still be included on the basis of other strengths.

In order to smoothly integrate the core group of tests into the curriculum, recommended tests should only require classroom instruction and minimal practice in the clinic to learn. Since the

students will be learning a comprehensive group of tests. one test should not require excessive time and practice to learn as this would make fitting the other core tests into the curriculum difficult.. In addition to requiring minimal time to learn. the tests should allow administration in less than sixty minutes to allow multiple tests to be administered. to allow time for explaining the results of the examination to the family, and to beginning the intervention. Also, the tests should not require extensive resources as the cost and time limit would limit the usefulness of these tests in all but the most specialized settings. Tests that require less time and cost to administer are more suitable for today's health care in which financial issues and limited care continue to push clinicians to get more results with less resources.

Standardized tests that are sensitive to change, yet have minimal ceiling and basement effects would be the most useful because they would quantify change over the widest range of functional levels. Even if the tests have excellent psychometric properties, it would be important that the students be trained how to match appropriate patients and tests to ensure change is measured properly.

To allow students to utilize standardized tests in a variety of clinical situations. the core group of tests must be applicable to a broad range of diagnoses and situations. Therefore, the core group of tests must allow the students to measure impairments in the musculoskeletal, neuromuscular, cardiopulmonary, and integumentary systems as well as functional limitations or health problems related to pathology affecting many body systems. The effects of interventions aimed at improving impairments with the goal of restoring function and decreasing disability must be measurable at the impairment, functional limitation. and/or disability level.. To accomplish this, standardized tests that measure common impairments should be included to monitor change at the impairment level. and measures of functional limitation and disability

should be included to monitor change at the functional and disability level. Tests at the impairment level allows clinicians to specifically measure if their intervention is effective in minimizing the impairment they feel is resulting in a functional limitation. Measurement at the functional limitation level allows clinicians to measure the effect of their intervention on function, and tests at the disability level allows clinicians to measure changes in the patient's ability to function in their social roles.

Physical therapist must be able to utilize standardized tests on patients who not only have a variety of diagnoses but also have different demographic characteristics. Patients with different demographics often participate in different daily functional tasks, live in different environmental conditions, and have different interpretations of test questions. As a result, differences in demographic characteristics such as race and socioeconomic status have been documented to effect the results and interpretation of standardized tests. [12,13,14] To help neutralize these differences, selected tests should be standardized on a variety of racial and socioeconomic populations so the tests are applicable to patients with various demographic characteristics. For example, most developmental assessments were standardized on White, middle-class children. As a result, children from different backgrounds scored differently. The Denver Developmental Screening Test is widely used in over a dozen countries, but secondary to concerns about the appropriateness of the test for various ethnic subgroups, a major revision of the test occurred culminating in the publication of the Denver IT which is more sensitive to demographic variables. [12]

Age is another factor that requires consideration when selecting a battery of tests that will meet the diverse needs of a generalist. Age accounts for differences in physiology, social, vocational, and family roles, resulting in different functional expectations. Therefore, the core

group of tests needs to be divided up into a pediatric, an adult, and a geriatric section with appropriate tests selected to meet the needs of each age group.

In addition to measurement at the levels of the disablement model, standardized tests were also selected to meet "specialized" needs such as measuring pediatric development, frailty among older adults, and wellness. Entry-level physical therapy students would benefit from being able to measure in each of these areas. Developmental tests are a necessity for therapists working with the pediatric population. Tests can be used to discriminate normal from abnormal and to evaluate change in response to intervention. Measures of frailty can be used identify older adults with an increased risk of falling and/or functional deficits that would affect their ability to accomplish activities of daily living or to function in the community. The information can also be used to make living environment recommendations and to measure the effect of intervention aimed at improving physical functioning. Wellness is in important aspect of health care that should be encouraged by physical therapist.. Therefore, entry-level physical therapy students should learn a tool that can be used to offer patients objective data on their level of wellness. The information can be used to encourage healthy lifestyle changes and can help the patient monitor their progress towards goals. In addition, wellness measures can help the therapist determine if patients have learned to separate the concepts of impairment or functional loss from "sickness", a concept that can cause a patient to unnecessarily reduce their activity level and not use the function that they have mastered in physical therapy. [9, 10]

Simply learning the core group of tests and when to use them is not sufficient unless students are given opportunities to use them in clinical situations. While test selection should be done with the expectation that they will need to be introduced during the didactic portion of the curriculum, efforts should be made to make sure that the students get practice in the selected

tests during their clinical education. Therefore, the suggested tests should be utilized at the clinical sites the students will be attending to allow the students to practice the application of the tests. Unfortunately, this may not be possible in all cases. Clinical sites and clinics surrounding the university may not use the best available tests, and in this case, the best tests should still be taught and other means for the students to practice the tests such as a lab activities or working with clinical sites to implement the usage of the selected tests should be sought after.

Selected Tests

Based on the uses of standardized tests and the criteria for selection,, a core group of twenty-four tests that best fulfill the needs of entry-level physical therapy students were selected (Table 1). Table 2 reviews the psychometric properties of the selected tests as the properties were significant factors in the selection process. Table 3 summarizes other important characteristics of the tests that affected selection.

The superiority of the selected tests when compared to other measures can be debated, but after learning the core group of tests, entry-level physical therapy students will be able to collect objective data in a variety of clinical situations for a wide range of uses. After physical therapy programs implement the tests into their curriculums, valuable experience and feedback from faculty and students will make it possible to adjust the selections when indicated. The core group of tests will be continually refined as new tests are developed, as data on current tests is updated, and as experience suggests other tests should be added. The core group of tests are an important step towards increasing utilization of standardized tests in the physical therapy profession.

Table 1: CORE GROUP OF STANDARDIZED TESTS FOR ENTRY-LEVEL PHYSICAL THERAPY STUDENTS

	Pediatric Population	Adult Population	Geriatric Population
Disabilitt	Child Health Questionnaire	SF-36	SF-36
	Pediatric Pain Disability Index	Pain Disability Index	Pain Disability Index
		Sickness Impact Profile	Sickness Impact Profile
Functional Limitation	WeeFIM	FIM	FIM
	PEDI		Physical Performance Test (pPT - measure of frailty)
	GMFM		
Imuairment Balance	Functional Reach	Berg Balance Scale	Berg Balance Scale
Balance		Tinetti Gait and Balance	Tinetti Gait and Balance
Cognition	Glascow Coma Scale	Glascow Coma Scale	Glascow Coma Scale
Cognition		Mini Mental State Examination	Mini Mental State Examination
Endurance	BABI	BABI	BABI
Endurance		Perceived Exertion Scale	Perceived Exertion Scale
Endurance	Percent Max Heart Rate	Percent Max Heart Rate	Percent Max Heart Rate
Endurance		Six Minute Walk Test	Six Minute Walk Test
Pain	Visual Analog Scale	Visual Analog Scale	Visual Analog Scale
Risk of Cardiovascular Disease	Body Mass Index	Body Mass Index	Body Mass Index
Develoumental	Denver II		
	Peabody		
	TIMP		

Table 2:	PSYCOMETRIC PROPERTIES		
Standardized Test	Reliability	Validity	Other
Disability			
Child Health Questionnaire	Established in a representative sample in the US, refer to manual for details [15]	Correlation of CHQ subscales with subscales of Health Utilities Index = 0.38 - 0.60[16]	
Pain Disability Index	Internal consistency =0.871[17]	Construct: patients with high scores reported more psychological distress, more severe pain characteristics, and more restriction of activities than patients with low scores"	
Pediatric Pain Disability Index	Not reported, but pediatric version is fairly similar to adult version	Not reported, but pediatric version is fairly similar to adult version	
SF-36	Social Functioning subscale = 0.76[18] Other subscales ~ 0.80[18]	Predictive: linked to utilization of health care services, clinical course of depression, loss of job within one year, and 5-year survival[19,20] Construct: established"	Sensitivity of 74% and Specificity of 81% in detecting patients diagnosed with depression"
Sickness Impact Profile	Interrater = 0.92[22] Testretest = 0.88-0.92[22]	Correlation with self-assessment of functional limitation = 0.69[22] Construct: scores had positive correlation with "up time" and a negative correlation with "down time,,22	
Functional Limitation			
FIM	Interrater = 0.83 - 0.96[1]	Predictive of min. of assistance required for patient's care' Concurrent with the Modified Barthel = 0.83 - 0.89[1]	
GMFM	Interrater of total score = 0.9g23 Interrater of individual dimensions = 0.87 - 0.99[1] Intrarater = O.~	Concurrent established by correlating change with parent judgement of change r = 0.54, therapist judgement of change r = 0.65, and masked evaluation of videotape r = 0.82[23]	
PEDI	Interrater for subscales ~ 0.91 except for social function =0.30[1]	Concurrent with Battelle Developmental Inventory Screening Test =0.70- 0.80[24] Construct: supported by significant differences between disabled and nondisabled wOUp~4	
PPT (measure of frailty)	Interrater = 0.W[5]	Concurrent with self-reported measures of function[2']	Sensitivity 79.3%26 Specificity 71%26
WeeFIM	Interrater total score> 0.95[27] Equivalence = 0.93[28]	Concurrent with Battelle Developmental Inventory and Vineland Ad~tive Behavior Scales = 0.72 -0.94	

Table 2 Continued:	PSYCOMETRIC PROPERTIES		
Standardized Test	**Reliability**	**Validity**	**Other**
Imnairment BABI	Intrarater = 0.91[30] Interrater = 0.84[30]	Construct: measured task difficulties conformed to ordering N<Cane≤Crutch[3O]	
Berg Balance Scale	Interrater = 0.99[31] Testretest = 0.98[31]	Concurrent with Tinneti = 0.ge[1] Construct: tested on 60 acute CVA with correlation with Barthel = 0.80-0.94[31]	
Body Mass Index	Not Applicable	Cardiovascular risk increased with a BMI between 25 and 30 and greatly increased with a BMJ above 30[32]	
Functional Reach	Interrater = 0.98[31] Intrarater = 0.924[31]	Correlation with center of pressure exertion = 0.7tl[3]	
Glascow Coma Scale	Not reported,, but is the gold standard for measuring level of arousaJ[34]	Not reported,, but is the gold standard for measuring level of arousal[33]	
Mini Mental State Examination	Test-retest = 0.887[35] Interrater = 0.827[35]	Predictive validity: score < 20 indicative of dementia or delirium[33]	Sensitivity 800/0[36] Specificity 98%36
Perceived Exertion Scale	Reliability ranges from 0.75 to 0.82 and decreases as exercise intensity increases[33]	Correlation with heart rates = 0.80 - 0.90[35] Correlation with power output = 0.56 - 0.83[38]	
Percent Max Heart Rate	ACSM guideline for exercise testing[33]	Construct: linear relationship to oxygen consumption[33]	
Six Minute Walk Test	Test-retest = 0.95[41]	Construct: correlation with $VO_{2\ max}$ determined in laboratory = 0.897[1] Concurrent: correlation with oxygen cost diagram = 0.68[1]	
Tinetti Gait and Balance	Interrater items=O.85[31] Interrater total=O.90[31]	Concurrent with Berg = O.si[11] Predictive of falls in the community dwelling elderly [1]	
Visual Analogue Pain Rating Scale	Test retest = 0.94[35]	Concurrent with McGill Pain Questionnaire = 0.60 - 0.63[35] and with Numeric Pain Rating Scale = 0.80[1]	

Table 2 Continued:	PSYCOMETRIC PROPERTIES		
Standardized Test	Reliability	Validity	Other
Pediatric Developmental Assessment Denver II	Interrater = 0.99[42] Intrarater = 0.90[42]	Construct: not applicable[43] High degree of face validity as the test was standardized on more than 2000 children.[43]	Specificity 43%44 Sensitivity 83%44
Peabody	Interrater of gross motor scales = 0.97[45] Interrater of fine motor scales = 0.94[45]	Construct: significant incremental increase in scores was observed at each age level except that corresponding to 54 to 59 months, this age level did not differ significantly from preceding age level	
TIMP	Intrarater = 0.8946 Internal Consistency = 0.98[42]	Construct: correlation between postconceptional age and TIMP performance measures = 0.83[47]	

Table 3: *all data cited in Table 2 CHARACTERISTICS AFFECTING SELECTION

Standardized Test *references refer to original tests	Reliability	Concurrent Validity	Construct Validity	Time to Administer	Methodoc Administration	Other Considerations
Disability Child Health Questionnaire [48]	Refer to user manual	-	Refer to user manual	10 -IS min.	++	Used in large population studies internationally
Pain Disability Index [17]	++	--------	#	5 -10 min.	++	
Pediatric Pain Disability Index	not reported, but only minimal changes were made from the adult version			5-lOmin.	++	
SF_36[49]	+,++	--------	#	10 -IS min.	++	
Sickness Impact Profile [50]	++;+++	-	#	20 - 30 min.	++	
Functional Limiytion FIM31	++,+++	++		30 min.	+	Part of the UDS for Medical Rehabilitation
GMFM'z	+++	-,++		45 -60 min.	+	
PEDI[53]	+++	+,++	#	Depends on format	+,++	
PPT[54]	+++	--------			+	
WeeFIM[55]	+++	+,+++		30 min.	+,++	Part of the UDS for Medical Rehabilitation
Impairment BABIJ!	++;+++	-------	#	Depends on task	•	
Body Mass Index [56]	--------	--------		< 5 min.	•	
Berg Balance Scale [57]	+++	+++	#	IS -20 min.	+	
Functional Reach [33]	+++	+		< 5 min.	•	
Glascow Coma Scale [34]	NR	NR		5 -10min.	+	
Mini Mental State Examination [58]	++	-------		5-IOmin.	++	
Perceived Exertion Scale [59]	+,++	-,++,+++		< 5 min.	++	
Percent Max. Heart Rate [39]	--------	-------	#	< 5 min.	•	Used by ACSM as a guideline for exercise testing
Six Minute Walk Test [60]	+++	-	#	20-30 min.	•	
Tinetti Gait and Balance [61]	++;+++	+++		10-15 min.	+	
Visual Analog Scale [62]	+++	-,++		5 min.	++	
Develpmental Assessment Denver II [63]	+++	NR	NA	5-10 min.	+,++	Standardized on subjects with a wide variety of demographic characteristics
Peabody [64]	Secondary to resources and set up required, the test will only be introduced to the students, arrangements to learn the test in more detail can be made with clinical sites for students who have a strong desire to do so					Used by Schools in Ulinois
TIMP[65]	++;+++	-------	#	36 min.	+, •	

Key

UDS - Uniform Data Set
ACSM - American College of Sports Medicine

Reliability:
+++~0.90
++ 0.80 - 0.89
+ 0.70-0.79
- 0.69 or less
NR - not reported

Concurrent Validity:
+++~0.90
++ 0.80 - 0.89
+0.70-0.79
- 0.69 or less
NR - not reported

Construct Validity:
established
NA not applicable

Method of Administration:
++ Questionnaire
+ Observation
• other

References

1Cole B, Finch Elspeth,, Gowland C, Mayo N. *Physical Rehabilitation Outcome Measures.* Canadian Physiotherapy Association; 1994.

2Dobrzykowski E. The outcomes movement:. progress in a sea of change. *PT Priority.* 1999~15:5-7.

3Glascoe FP, et al. Accuracy of the Denver-II in developmental screening. *Pediatrics.* 1992;89: 1221-1225.

4Jenkinson C, Wright L, Coulter A. The Short-Form 36 Health Survey: evaluation of the UK
SF-36 in various groups and development of the UK summary scale scores. Health Services Research Unit Annual Report. Oxford,, England: Health Services Research Unit, Department of Public Health and Primary Care, University of Oxford,, 1995.

5Jenkinson C, Wright L, Coulter A. Quality oflife measurement in health care: a review of measures and population norms for the UK SF-36. Oxford,, England: Health Services Research Unit, Department of Public Health and Primary Care, University of Oxford,, 1993.

~cHorney CA, Kosinski M, Ware JE. Comparisons of the costs and quality of norms for the SF-36 Health Survey collected by mail versus telephone interview: results from a national survey. *Medical Care. 1994;32:551-567.*

7Jette DU, Jette AM. Physical therapy treatment choices for musculoskeletal impairments. *Physical Therapy. 1997;77:145-154.*

8Jette DU, Jette AM. Professional uncertainty and treatment choices by physical therapists. *Archives of Physical Medicine and Rehabilitation. 1997;78:1346-1351.*

~rooks WB, Jordan JS, Divine GW, Smith KS, Neelon FA. The impact of psychologic factors on measurement of functional status. *Medical Care. 1990;28:793-814.*

l<>:MulrowCD, Gerety MB, Cornell JE, Lawrence VA, Kanten DN. The relationship between disease and function and perceived health in very frail elders. *Journal of American Geriatrics Society. 1994;42:374-380.*

llLean ME, Han TS, Seidell JC. Impairment of health and quality of life using new US federal guidelines for the identification of obesity. *Archives of Internal Medicine.* 1999;159:837-843.

12Kerfeld CL, Guthrie MR, Stewart KB. Evaluation of the Denver II as applied to Alaska native children. *Pediatric Physical Therapy. 1997;9:23-31.*

13Newton RA. Balance screening of an inner city older adult population *Archives 0/ Physical Medicine and Rehabilitation. 1997;78:587-591.*.

14Cintas HL. Cross-cultural similarities and differences in development and the impact of parental expectations on motor behavior. *Pediatric Physical Therapy. 1995;7:103-111.*.

15CHQ: http://www.sf-36.com/general/chq.html ...

16Speechley NK, et al. Mutual validity of the Child Health Questionnaire and the Health Utilities Index: an explaoratory analysis using survivors of childhood cancer. *International Journal of Cancer.*. 1999;83(SI2):95-105.

17Tait RC, Pollard A, Margolis RB, Duckro PN, Krause SJ. The Pain Disability Index: pyschometric and validity data. *Archives 0/ Physical Medicine and Rehabilitation.* 1987;68:438-441.

18Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey Manual and Interpretation Guide. Boston, MA: New England Medical Center, The Health Institute, 1993.

19Ware JE, Keller S, Bentler PM, Sullivan M, Brazier J, Gandek B. Comparisons of health status measurement models and the validity of SF-36 in Great Britain, Sweden,, and the USA.. *Quality of Life Research. 1994;3:68.*

2<>WellKB, Rogers W, Hays RD, Camp P. The course of depression in adult outpatients: results from the Medical Outcomes Study. Arch Gen Psychiatry. 1992;49:788-794.

21SF_36: http://www.sf-36.com/general/sf36.html ...

22McDowell I, Newell C. *Measuring Health.. A Guide to Rating Scales and . Questionnaires.* Oxford, NY. Oxford University Press; 1996.

23Boyce et al. The Gross Motor Performance Measure: validity and responsiveness of a measure of quality of movement.. *Physical Therapy. 1995;75:603-613.*

24Feldman AB, Haley SM, Coryell J. Concurrent and construct validity of the Pediatric Evaluation of Disability Inventory. *Physical Therapy. 1990;70:602-610.*

25Reuben DB, Siu AL. An objective measure of physical function of elderly outpatients. The Physical Performance Test.. *Journal 0/ the American Geriatric Society.* 1990;38:1105-0012.

26VanSearingen JM, Paschal KA, Bonino P, Chen TW. Assessing recurrent fall risk of community-dwelling, frail older veterans using specific tests of mobility and the physical performance test of function. *Journal of Gerontology and Biology Science Medical Science.* 1998;53:M547-64.

270ttenbacher KJ, Msall ME, Lyon NR, Duffy LC, Granger CV, Braun S. Interrater agreement and stability of the Functional Independence Measure for Children (WeeFIM): use in children with developmental disabilities. *Archives of physical Medicine and Rehabilitation.* 1997;78:1309-1315.

28Sperle PA, Ottenbacher KJ, Braun SL, Lane SJ, Nochajskj LS. Equivalence reliability of the Functional Independence Measure for Children (WeeFIM) administration methods. *American Journal of Occupational Therapy.* 1997;51 :35-41.

290ttenbacher KJ, Msall ME, Lyon N, Duffy LC, Granger CV, Braun S. Measuring developmental and functional status in children with disabilities. Dev Med Child Neurol.. 1999;41: 186-194.

30pine ZM, Colbran E, Corplongo R. Reliability and validity of the Beats Above Baseline Index: a new measure of task difficulty. *Archives of physical Medicine and Rehabilitation. 1994;75:545-550.*

31Russo SG. Clinical balance measures: literature resources. *Neurology Report.* 1997;21 :29-36.

32Lean ME, Han TS, Seidell JC. Impairment of health and quality of life using new US federal guidelines for the identification of obesity. *Archives of Internal Medicine.* 1999;159:837-843.

33Duncan PW, Weiner DK, Chandler J, Studenski S. Functional reach: a new clinical measure of balance. *Journal of Gerontology.* 1990;45:M192-197.

34Wade DT.. *Measurement in Neurological Rehabilitation.* Oxford, NY. Oxford University Press; 1995.

35Lewis CB, McNerney T. *Clinical measures of functional outcomes. 'The functional outcomes tool box."* Washington, DC. Learn Publication Inc. 1994.

3~acKenzie *DM,* Copp P, Shaw RJ, Goodwin GM. Brief cognitive screening of the elderly: a comparison of the Mini-Mental State Examination (MMSE), Abbreviated Mental Test (AMT) and Mental Status Questionnaire (MSQ). *Psychological Medicine.* 1996;26:427-430.

37Lamb KL, Eston RG, Corns D. Reliability of ratings of perceived exertion during progressive treadmill exercise. *British Journal of Sports Medicine. 1999;33:326-329.*

38Robertson R, et. al. Validity of the Borg perceived exertion scale for the use in semirecumbant ergometry during immersion in water. *Percept Mot Skills. 1996;83:3-13.*

39Mahler DA. ACSM's Guidelines for Exercise Testing and Prescription, , 5th edition. Williams & Wilkins. 1995.

40pollock ML, et al. ACSM position stand on the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in adults. *Medicine & Science in Sports & Exercise.* 1998~30:975-991.

41Harada ND, Chiu V, Stewart AL. Mobility-related function in older adults: assessment with a 6-minute walk test. *Archives of physical Medicine and Rehabilitation.* 1999~80:837-341..

42Long TM, Cintas HL. *Handbook of Pediatric Physical Therapy.* Philiadelphia, PA. Williams and Wilkins~ 1995.

43Prankenburg WK., et al. The Denver II: a major revision and restandardization of the Denver Developmental Screening Test.. *Pediatrics. 1992;89:91-97.*

44Glascoe PP, Byrne MA, Ashford LG, Johnson KL, Chang B, Strickland B. Accuracy of the Denver-II in developmental screening. *Pediatrics. 1992;89:1221-1225.*

45Schmidt LS, Westcott SL, Crowe TK. Interrater reliability of the Gross Motor Scale of the Peabody Developmental Motor Scales with 4- and 5-year-old children. *Pediatric Physical Therapy. 1995;5:169-174.*

46CampbeU SK. Test-retest reliability of the Test of Infant Motor Performance. *Pediatric Physical Therapy. 1999;77:60-66.*

47Campbell SK, Kolobe TH, Osten ET, Lenke M, Girolami GL. Construct validity of the Test of Infant Motor Performance. *Physical Therapy. 1995;75:585-596.*

48Landgraf JM, Abetz L, Ware JE. The CHQ User's Manual.. First Edition. Boston, MA: The Health Institute, New England Medical Center.

49Ware JE. A New MOS Short-Form Health Survey (SP-36). Boston, MA: The Health Institute, New England Medical Center, 1988.

5≤namiano AM. The Sickness Impact Profile Users Manual and Interpretation Guide. Baltimore: Johns Hopkins University.

51Guidefor the Uniform Data Setfor Medical Rehabilitation (Adult FIM). Version 4.0. Buffalo, NY 14214: State University of New York at Buffalo; 1993.

52Russell D, Rosenbaum P, Gowland C, et al. *Gross Motor Function Measure (GMFM): A measure of gross motor function in cerebral palsy.* Hamilton, Ontario: McMaster University; 1990.

53Haley SM, Coster WJ, Ludlow LH, Haltiwanger JT, Andrellos PJ. *Pediatric Evaluation of Disability Inventory (PEDI). Development, Standardization and Administration Manual.* Boston: PEDI Research Group, New England Medical Center Hospitals; 1992.

54Reuben DB, Siu AL. An objective measure of physical function of elderly outpatients: The Physical Performance Test. *JAGS. 1990;38:1105-1112.*

55Guide for Uniform Data Set for Medical Rehabilitation for Children (WeeFimSM) Version 4.0 - Community Outpatient. Buffalo, NY: Uniform Data System for Medical Rehabilitation, State University of New York; 1994.

5~ody Mass Index: www.nhlbi.nhi.gov/guidelines/obesity/bmi_tbl.html

57Berg K, Wood-Dauphinee S, Williams Jl, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiotherapy Canada.* 1989;41 :304-311.

58Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975; 12:189-198.

5~org G. Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise. 1982;14:377-381.*

~utland RJA, Pang J, Gross ER, Woodcock AA, Geddes DM. Two-, six-, and 12-minute walking tests in respiratory disease. *British Medical Journal. 1982;284:1607-*1608.

61Gottschalk M. The performance-oriented assessment of mobility: Tinetti Balance and Gait Scale. *GeriNotes. 1999;5:16-17.*

62McCormack HM, deL Home DJ, Sheather S. Clinical applications of visual analogue scales: 2. A critical review. *Psychological Medicine. 1988;18:1007-1019.*

63Frankenburg WK, Dodds J, Archer P, et al. *Denver II Screening Manual.* Denver, CO: Developmental Materials, Inc; 1990.

64Folio R, Fewell RR. *Peabody Developmental Motor Scales and Activity Cards.* Hingham, MA: DL.M Teaching Resources 1983.

65CampbeU SK, Osten ET, Kolobe THA, Fisher AG. Development of the Test of Infant Motor Performance. Physical Medicine and Rehabilitation Clinics of North America. 1993;4:541-550.