Measures of Functional Status in Community-Dwelling Elders

Scott E. Sherman, MD, MPH, David Reuben, MD

OBJECTIVE: To evaluate two performance-based measures of functional status and assess their correlation with self-report measures.

DESIGN: Cross-sectional study.

PARTICIPANTS: Of the 363 community-dwelling elders enrolled in a trial of comprehensive geriatric assessment who participated, all had at least one of four target conditions (urinary incontinence, depression, impaired functional status, or history of falling).

MEASUREMENTS: Two performance-based measures, National Institute on Aging (NIA) Battery, and Physical Performance Test (PPT), and three self-report functional status measures, basic and intermediate activities of daily living and the Short-Form-36 (SF-36) physical functioning subscale, were used. Measures of restricted activity days, patient satisfaction and perceived efficacy were also used.

MAIN RESULTS: All measures were internally consistent. There was a high correlation between the NIA and PPT ($\kappa = 0.71$), while correlations between the performance-based and self-report measures ranged from 0.37 to 0.50. When patients with values above the median on the two performance-based measures were compared with those below, there were significant differences ($p \leq .0001$) for age, number of medications, and the physical function, pain, general health, and physical role function SF-36 subscales.

CONCLUSIONS: Performance-based measures correlated highly with each other and moderately with questionnaire-based measures. Performance-based measures also had construct validity and did not suffer from floor or ceiling effects.

KEY WORDS: performance-based measure; functional status; elderly.

J GEN INTERN MED 1998;13:817-823.

In the last 15 years, a tremendous amount of research on quantifying how well individuals function has been reported.¹ Valid, reliable instruments have been developed to measure impairment, screen for early disability, and measure change over time. Assessing function is particularly important in the elderly, as the prevalence of functional disability increases with age.² Most instruments have relied either on patient self-report,³ or on di-

rect observation of the individual performing a variety of tasks (performance-based measures).⁴

Although there is widespread agreement that screening for functional status in older persons is important,⁵ the preferred method is still uncertain. Moreover, it is unclear whether self-report and performance-based measures can be used interchangeably. Previous studies comparing different measures of functional status assessment have typically found a good but not great correlation. The reason for this less-than-perfect agreement is not clear. Some possibilities are that the scales are measuring different aspects of function, that function is too complex to measure precisely with questionnaires or short performance assessments, or that the method of assessment is responsible for the differences.⁶

When considering screening for functional impairment in clinical settings, several questions remain: Which instrument should be used? Is one method (self-report vs performance-based) preferable? Does the method of measurement affect the type of impairment detected? Using data obtained from community-dwelling elders, we addressed three study questions: (1) whether two commonly used performance-based measures, the National Institute on Aging (NIA) Battery of lower extremity function⁷ and the Physical Performance Test (PPT),⁸ are reliable and valid; (2) how well these two performance-based measures correlate with each other and with self-report measures of function; and (3) whether there are particular areas in which these two types of measures correlate highly.

METHODS

We used data from subjects participating in Project Safety Net, a randomized controlled trial of comprehensive geriatric assessment in community-dwelling older subjects^{9,10}. Briefly, Project Safety Net recruited English-speaking subjects aged 65 years or older who had a telephone and met the following criteria: did not live in a nursing home; had one of the following four conditions—urinary incontinence, depression, history of falling within the last 6 months, or impaired functional status; was not cognitively impaired (Mini Mental State Exam [MMSE]¹¹ score ≥ 24); and was willing to give informed consent. We recruited subjects through senior centers and retirement homes, predominantly in lower socioeconomic areas. After a brief description of the study, subjects completed an 11-minute, 30-item screening questionnaire that included three subscales of the Functional Status Questionnaire¹² (basic and intermediate activities of daily living, and social activities) and questions on depression,¹³ urinary incontinence,¹⁴ falls,¹⁵ and demographics. Subjects who answered positively to the question on depression also completed

Received from the Multicampus Program in Geriatric Medicine and Gerontology, UCLA School of Medicine, and the Department of Veterans Affairs Medical Center, Sepulveda, Calif., and West Los Angeles, Calif. (SES, DR).

Address correspondence and reprint requests to Dr. Sherman: VA Medical Center, PACE Center (00PG), 16111 Plummer St., Sepulveda, CA 91343.

the 30-Item Geriatric Depression Scale (GDS).¹⁶ Final inclusion criteria were falling within the warning zone on any of the Functional Status Questionnaire subscales, a positive answer to both questions on incontinence, score of 11 or higher on the GDS, or a positive answer to the falls screening question. Subjects who met one or more of these final criteria and had a primary care provider then underwent a 90-minute interviewer-administered survey and performance-based assessment of physical function. The questions included the Medical Outcome Study Short Form-36 (SF-36),^{17,18} the Patient Satisfaction Questionnaire,¹⁹ the Perceived Efficacy in Physician-Patient Interactions Scale,²⁰ the Illness Self-Mastery Scale,²¹ and questions on restricted activity days.²²

All subjects completed two tests of physical performance, the NIA Battery and the PPT. The NIA Battery assesses lower extremity function and tests balance, gait, strength and endurance.⁷ Subjects are rated on their standing balance, how long it takes to complete an 8-foot walk, and how long it takes to get up from a chair and sit down again five times. Each of the three items is scored from 0 to 4, giving a total possible score range of 0 to 12.

The PPT assesses both lower and upper extremity function.⁸ The version used in this study consists of seven items, each scored from 0 to 4, for a total range of 0 to 28. It includes writing a sentence, simulated eating, simulated dressing, picking up a penny, placing a heavy object on a shelf, turning 360°, and walking 50 feet.

The SF-36 from the Medical Outcome Study¹⁷ (also known as the RAND 36-Item Health Survey 1.0¹⁸) is a multidimensional survey that assesses patient function and well-being. It consists of eight subscales that have 2 to 10 items each. The subscales include physical functioning, role limitations due to physical problems, social functioning, bodily pain, general mental health, role limitations due to emotional problems, energy/fatigue, and general health perceptions. The Physical Function Subscale (PF-10) includes 10 items that span a range of tasks and activities, such as participation in moderate or vigorous activities, climbing one flight of stairs, walking, and bathing or dressing. Subjects report for a typical day the extent to which their health limits them in these activities, with choices being "not limited at all," "limited a little," or "limited a lot." The PF-10 and the other subscales of the SF-36 are each scored on a range of 0 to 100, with 100 being the best possible score.

Basic activities of daily living (BADL) and intermediate activities of daily living (IADL) are two of the subscales of the multidimensional Functional Status Questionnaire.¹² For each of these, subjects were asked to report during the past 4 weeks how much physical difficulty they have had with a specific item, such as moving in and out of bed, walking several blocks, errands such as grocery shopping, or driving a car. The five possible responses for each item are "usually done with no difficulty," "some difficulty," or "much difficulty"; "usually did not do because of health"; and "usually did not do for

other reasons." As with the PF-10, each of these two subscales has a range of 0 to 100, with 100 being the best possible score. Warning zones are considered to be below 90 for the 3-Item BADL and below 73 for the 6-Item IADL.

We measured the internal consistency of each scale using Cronbach's α .²³ We computed the Pearson product moment correlates between scales and among items within scales. Correlations between scales were also disattenuated. We assessed differences between correlations using Student's *t* tests for dependent correlations.²⁴ We assessed construct validity by looking for both convergent and discriminant validity.²⁵ Convergent validity determines whether the performance-based measures correlate highly with other measures of similar concepts, such as physical functioning. Discriminant validity determines whether the performance-based measures correlate poorly with concepts less closely related to function, such as patient satisfaction. Specifically, we anticipated that the performance-based measures would be significantly related to physical function, physical role function, general health, and pain subscales of the PF-10; age; and the number of medications. We anticipated that the measures would not be related to emotional well-being and emotional role function subscales of the PF-10; MMSE score; and measures of depression, patient satisfaction, perceived efficacy, and illness self-mastery. To test construct validity of the two performance-based measures, we split each distribution at the median and compared mean scores for the group that scored below the median with the group that scored above.

We used both a priori and a posteriori methods to see if there were areas of high and low concordance between the performance-based and self-report measures. We used a framework developed at a 1993 NIA-sponsored conference on assessing function (D. Reuben, unpublished data). This framework provides an overlay to fit performance-based measures into the model of function developed by Verbrugge and Jette,²⁶ as well as others.^{27,28} The model states that function and performance can be seen as a hierarchy of four progressively more integrated levels, ranging from basic components of physical function (such as strength or balance) to specific physical movements (such as sitting, standing, or reaching) to goal-directed activities (such as writing or eating) to personal role choices such as working or exercising. Using this scheme, we classified each item from the PPT, NIA, and PF-10 into one of three categories: a basic component of physical function, a specific movement, or a highly integrated activity (combining goal-directed activities and personal role choices). We hypothesized that there should be a higher correlation between items from different scales within the same category than between items that fall within different categories. For example, the PF-10 questions on specific movements should correlate more highly with the PPT items on specific physical movements than with the PPT items on more highly integrated activities.

For the a posteriori method, we assessed the correlation between each item of the NIA and PPT with the overall

score for the PF-10. Noting that the mean and median of these correlations were both 0.30, we split the items into those with higher ($r \geq .30$) and lower ($r < .30$) correlations. We then looked at the items with lower correlation and at those with higher correlation to see if there were discernible patterns. We also looked at seven pairs of items from different scales that we would expect to be highly correlated, such as bending/kneeling (PF-10) and picking up a penny (PPT), or bathing/dressing (PF-10) and putting on a jacket (PPT), to see if there was in fact a particularly high correlation. We used SPSS for Windows version 6.1 to perform all analyses.

RESULTS

The average age $\pm$ SD for the 363 subjects in our study was 75.9 ± 5.9 years. Subjects were taking an average of 2.5 ± 2.2 medications and had a mean MMSE score of 28.2 ± 1.4 . Table 1 shows the characteristics of the different measures of physical performance and functional status. There were essentially no floor or ceiling effects for the PPT, NIA, or the physical function subscale of the SF-36 (PF-10). There was a substantial ceiling effect for the two measurements of activities of daily living, particularly basic activities of daily living. Internal consistency was moderately high for each of the scales and very high for the PF-10.

The correlations between the scales are shown in Table 2. There was a very high correlation ($r = .71$) between the two performance-based measures, and a good correlation ($r = .37-.50$) between either of the two measures and the three self-report measures of functional status. The correlation between the NIA and PPT scores was significantly higher than the correlation of either performance-based measure with the three self-report measures ($p < .001$ for each comparison). Disattenuated correlations, shown above the diagonal, reveal the same pattern.

Our assessment of the construct validity of the two performance-based measures—whether they are related significantly with similar measurements and less strongly

related to dissimilar measurements—is shown in Table 3. Means are given for subgroups defined by dividing the PPT and NIA groups at the median of their respective distributions. The results were very similar for each of the two performance-based measures. In general, constructs associated with physical function such as general health, age, energy, and pain, did differ between the group below the median and the group above the median. Constructs that do not have as clear a logical association with physical function, such as emotional state, patient satisfaction, and perceived efficacy in communicating with the provider, showed no difference between the groups.

We split the PF-10 and PPT into specific movements (e.g., lifting) and more highly integrated activities (e.g., walking several blocks or writing a sentence). The NIA has only specific movements. Table 4 shows the correlation between the different parts of these three instruments. As shown in the bottom row, the NIA specific movements had a higher correlation with specific movements on the PF-10 and PPT than with the more highly integrated activities on the PF-10 ($p < .02$) and PPT ($p < .001$). Similarly, the specific movements on the PPT had a higher correlation with the specific movements than with the more highly integrated activities on the PF-10 ($p < .02$). However, the more highly integrated activities on the PPT correlated no better with the more highly integrated activities than with the specific movements on the PF-10 ($p = .71$).

When looking at specific items in the NIA and PPT, the total PF-10 score had correlation $\geq .30$ with the following performance-based items: putting on/or removing a jacket, picking up a penny, getting up from a chair, and walking either 8 or 50 feet. In contrast, the total PF-10 score had a lower correlation ($r < .30$) with the remaining performance-based items: writing a sentence, simulated eating, placing a book on a shelf, turning 360° , and standing balance.

Pairs of similar items from different scales were correlated, but the correlations were in general no better than pairs of dissimilar items from different scales. For example, the correlation between putting on a jacket (PPT) and bathing or dressing (PF-10) was .24, while its correlation

Table 1. Characteristics of Tests of Physical Performance and Functional Status Used in Project Safety Net*

Test†	Number of Items	Range of Possible Scores	Mean	Median	SD	% at Floor‡	% at Ceiling‡	Internal Consistency, Cronbach's α
Performance-based measures								
PPT	7	0–28	22.3	23	4.2	0	2.2	0.67
NIA	3	0–12	8.0	8	2.3	0	2.2	0.60
Self-report measures								
FSQ BADL	3	0–100	87.8	89	13.8	0	41.1	0.63
FSQ IADL	6	0–100	72.6	73	22.3	0.5	16.2	0.79
SF-36 PF-10	10	0–100	57.1	60	26.2	0.3	4.1	0.90

*These results are based on the 363 subjects enrolled in Project Safety Net, a randomized controlled trial of comprehensive geriatric assessment in community-dwelling elders.

†PPT indicates Physical Performance Test; NIA, National Institute on Aging Battery; FSQ, Functional Status Questionnaire; BADL, basic activities of daily living; IADL, intermediate activities of daily living; SF-36, 36-Item Short Form; PF-10, Physical Function Subscale.

‡Subjects are at floor if they have the lowest possible score for the test and at ceiling if they have the highest possible score.

Table 2. Correlation Between Tests of Physical Performance and Functional Status in Project Safety Net*

Test*	PPT	NIA	FSQ BADL	FSQ IADL	SF-36 PF-10
PPT	—	1.00	0.57	0.67	0.59
NIA	0.71 [†]	—	0.68	0.73	0.65
FSQ BADL	0.37 [†]	0.42 [†]	—	0.85	0.56
FSQ IADL	0.49 [†]	0.50 [†]	0.60 [†]	—	0.78
SF-36 PF-10	0.46 [†]	0.48 [†]	0.42 [†]	0.66 [†]	—

*PPT indicates Physical Performance Test; NIA, National Institute on Aging Battery; FSQ, Functional Status Questionnaire; BADL, basic activities of daily living; SF-36, 36-Item Short Form; PF-10, Physical Function Subscale.

[†]Disattenuated correlations are above the diagonal. For all other correlations, $p < .001$.

with other items on the PF-10 ranged from .16 to .30. Similarly, picking up a penny (PPT) correlated well with bending or stooping on the PF-10 ($r = .33$), but also with many of the other PF-10 items ($r = .17$ – $.32$, with more than half of the correlations between individual items exceeding $r = .25$).

DISCUSSION

Our first study objective was to determine whether two commonly used performance-based measures were reliable and valid. We found that in this study of community-dwelling older persons, the two performance-based measures studied—the NIA Battery⁷ and the PPT⁸—are potentially useful. Both are internally consistent, with Cronbach's α of 0.60 and 0.67, respectively. Neither suffers from either

floor or ceiling effects in this population of community-dwelling elders. Both appear to have construct validity.

Our second study objective was to determine how well these two performance-based measures correlated with each other and with other measures of functional status. The two performance-based measures were highly correlated with each other ($r = .71$). They also correlated, although to a lesser extent, with self-report measures of functional status ($r = .37$ – $.50$). The disattenuated correlation was 1.00 for the two performance measures and correlations exceeded 0.56 for other measures. This result suggests that any of these measures are potentially useful for screening in primary care.

Our final study objective was to determine if there are particular areas in which these two types of measures correlate highly, to help determine what kind of functional impairment each instrument detects. Our hypothesis was that performance-based measures and self-report measures are assessing aspects of function that are similar but not identical, like overlapping circles on a Venn diagram. The disattenuated correlations clearly support this hypothesis. To further test this, we used a framework developed at a 1993 NIA-sponsored conference on assessing function (D. Reuben, unpublished data). Our results support this hypothesis, as in most cases we found a significantly higher correlation across scales between the same level of function (e.g., specific movements) than between different levels of function (e.g., specific movements with more highly integrated activities). However, this is not the entire story, as similar items on different scales did not correlate as highly as expected. For example, putting on a jacket did not correlate better with bathing and dressing than with

Table 3. Assessment of Construct Validity, Based on Splitting the Performance-Based Measures at the Median

Test (Range or Unit)*	Physical Performance Test		National Institute on Aging Battery	
	Below Median (0–23)	Above Median (24–28)	Below Median (0–8)	Above Median (9–12)
SF-36 subscales (0–100)				
Physical function	48	67 [†]	46	69 [†]
Physical role function	47	65 [†]	44	67 [†]
General health	56	64 [†]	55	65 [†]
Energy/fatigue	48	58 [†]	47	58 [†]
Pain	57	68 [†]	56	70 [†]
Emotional well-being	69	72	69	72
Emotional role function	75	77	74	78
Number of days did less	3.7	2.2 [‡]	3.5	2.4
Age, years	77	75 [†]	77	74 [†]
Medications, n	2.9	2.1 [§]	2.7	2.3 [‡]
Mini-Mental State (0–30)	28.2	28.2	28.1	28.3
Depression (GDS) (0–30)	10.3	10.0	11.0	9.3
Patient satisfaction (1–5)	3.3	3.4	3.3	3.4
Perceived efficacy (1–5)	4.0	4.0	4.0	4.0
Illness self-mastery (1–5)	4.0	3.9	3.9	4.0

*SF-36 indicates 36-Item Short Form¹⁷; days did less, number of days per month patient reported restricting his or her activity because of health; GDS, Geriatric Depression Scale¹⁶.

[†] $p < .0001$ for comparison of the group below the median with the group above the median.

[‡] $p < .05$.

[§] $p < .001$.

Table 4. Correlation Between Different Parts of the PF-10, PPT, and NIA Subscales*

Subscale and Part	PF-10		PPT		NIA	
	Specific Movements [†]	More Highly Integrated [‡]	Specific Movements [§]	More Highly Integrated	Specific Movements [¶]	More Highly Integrated [#]
PF-10						
Specific movements [†]	—					
More highly integrated [‡]	0.59	—				
PPT						
Specific movements [§]	0.42	0.32	—			
More highly integrated	0.34	0.19	0.42	—		
NIA						
Specific movements [¶]	0.49	0.35	0.66	0.39	—	
More highly integrated [#]	—	—	—	—	—	—

*Values listed are correlations. All are statistically significant, with $p < .0005$. PF-10 indicates Physical Function Subscale; PPT, Physical Performance Test; NIA, National Institute on Aging Battery.

[†]Bending/kneeling.

[‡]Moderate and vigorous activities, lifting/carrying, walking ≥ 1 block or up to ≥ 1 flight, bathing/dressing.

[§]Lifting, picking up, turning 360°, walking 50 feet.

^{||}Write sentence, put on/remove jacket, eating.

[¶]Chair stands, walking 8 feet.

[#]The NIA Battery has no items that assess more highly integrated functions.

most of the other items on the self-report scale, the PF-10. This lack of a high correlation may be due to the fact that the PPT asks the patient to *do something*, while the PF-10 asks whether the patient *has been able to do something*. Alternatively, it may be an “instrument effect,” where items from a scale correlate more highly with each other than with items from outside the scale, owing in part to similar wording and structure.

We believe that for each of the four *levels* of the model described above (basic components, specific movements, goal-directed activities, and personal role choices), there are three *aspects*—capability, motivation, and perception. Thus, how subjects answer or perform is affected by what they *can* do (capability), what they *want* to do (motivation), and what they *feel* they can do (perception). Therefore, a self-report measure assessing more highly integrated activities (such as working or shopping) may not correlate highly with a performance-based measure assessing specific movements (such as picking up a penny) because they are measuring different levels of function. Similarly, two measures assessing the same level of function may still not correlate highly if one is more influenced by motivation or by perception. Our findings are consistent with this model, but other explanations are certainly possible. This may provide direction for future research, although assessing motivation and perception has been difficult. In a recent paper, Glass breaks function down into three “tenses”: the hypothetical (can do...), the experimental (could do...), and the enacted (do do...).⁶ Our categories are very similar to those of Glass, with the chief difference being that ours emphasize motivation a bit more than his do.

Primary care providers and administrators are faced with the question of whether to screen for functional impairment and, if so, how to do it. Because functional disability is common, serious, possible to diagnose, often over-

looked, and often modifiable, many have chosen to screen older patients for it. A wide range of questionnaire measures are available (e.g., Functional Status Questionnaire,² SF-36,^{17,18} Short Form-12²⁹) or performance-based measures (e.g., PPT,⁸ NIA Battery⁷). Our results indicate that either questionnaire measures or performance-based measures work reasonably well. The performance-based measures are oriented toward more basic levels of function, but our results showed that there is considerable overlap. More important in deciding how to screen will be the logistics involved, including who will administer the test, who will score, and how the clinician will be alerted about abnormal results. Self-administered questionnaires have less initial cost, but the data still must be transformed into a useful form. The performance-based measures require someone to administer them, but take only a few minutes and are easy to score. Two very simple and reliable options would be the Short Form-12 (a 12-item questionnaire) or the NIA Battery (a 3-item performance-based measure). Probably the most important issue for any office or system of care will be how to get the primary care provider to act on the results. Past studies indicate that simply giving the clinician the patient's score and the reference range has no effect on provider behavior.³⁰ Something more detailed, such as specific advice on how and where to refer, is generally needed to effect a change in provider behavior,³¹ and even that may not be sufficient.³²

Several other studies have looked at how well performance-based measures correlate with self-report questionnaires. There has consistently been a moderately good correlation between performance-based measures and individual self-reported items.³³⁻³⁵ The correlation has generally been better for lower extremity skills than for upper extremity skills. However, we did not find that the NIA Battery (which includes only lower extremity skills) correlated

any better with the self-reported measures of functional status than the PPT did (which also includes upper extremity skills). There also has been a moderately good correlation when comparing entire scales.³⁶⁻³⁸

Only one study to date has examined why performance-based measures and self-report measures of functional status are not highly correlated. Kempen et al. studied the relation between self-report measures and performance-based measures in a subsample of the population-based Groningen Longitudinal Aging Study.³⁹ They found that subjects with more depressive symptoms and those with lower levels of perceived physical competence were more likely to "underestimate" their self-reported functioning (in comparison with the performance-based assessment).

Our study has several limitations. First, our subjects lived in or around Los Angeles, and it is not clear whether our results apply to other geographic areas. Second, our subjects all agreed to participate in a randomized trial, which may make them somewhat different from subjects outside the trial. Finally, our subjects were community-dwelling elders who all had a primary care provider. There are certainly other groups (e.g., nursing home patients) to whom these results may not apply. Furthermore, the entry criteria were weighted toward those with some dysfunction, as subjects were required have depression, urinary incontinence, functional impairment, or a history of falls to enter the study.

Our study adds to the literature in several ways. First, it is the first study we know of that compares two different performance-based measures in detail. Second, it adds information on how and where the two performance-based measures correlate with some commonly used questionnaire measures of function. Unfortunately, our results only partially agreed with the model we had hypothesized. The model we outline presents a plausible explanation for why the agreement was less than we had expected. A challenge for future researchers will be to develop a reliable way to assess the motivational component of each of these measures.

Dr. Sherman has received a Career Development Award from the Claude Pepper Older Americans Independence Center, National Institute on Aging grant 5P60AG10415.

The authors thank Ron Hays, PhD, and Michael Rotblatt, MD, for very helpful comments on this manuscript and Audree Chapman for assistance in preparing the manuscript.

REFERENCES

- Applegate WB, Blass JP, Williams TF. Instruments for the functional assessment of older patients. *N Engl J Med*. 1990;322:1207-14.
- Jette AM, Pinsky JL, Branch LG, Wolf PA, Feinleib M. The Framingham disability study: physical disability among community-dwelling survivors of stroke. *J Clin Epidemiol*. 1988;41(8):719-26.
- Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med*. 1993;118:622-9.
- Guralnik JM, Branch LG, Cummings SR, Curb JD. Physical performance measures in aging research. *J Gerontol Med Sci*. 1989;44(5):M141-6.
- Hall WJ. Update in geriatrics. *Ann Intern Med*. 1997;127:557-64.
- Glass TA. Conjugating the tenses of function: discordance among hypothetical, experimental, and enacted function in older adults. *Gerontologist*. 1998;38:101-12.
- Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol Med Sci*. 1994;49(2):M85-94.
- Reuben DB, Siu AL. An objective measure of physical function of elderly outpatients: the physical performance test. *J Am Geriatr Soc*. 1990;38:1105-12.
- Reuben DB, Hirsch SH, Chernoff JC, et al. Project Safety Net: a health screening outreach and assessment program. *Gerontologist*. 1993;33(4):557-60.
- Reuben DB, Maly RC, Hirsch SH, et al. Physician implementation of and patient adherence to recommendations from comprehensive geriatric assessment. *Am J Med*. 1996;100:444-51.
- Folstein 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;12:189-98.
- Jette AM, Davies AR, Cleary PD, et al. The Functional Status Questionnaire: reliability and validity when used in primary care. *J Gen Intern Med*. 1986;1:143-9.
- Lachs MS, Feinstein AR, Cooney LM, et al. A simple procedure for general screening for functional disability in elderly patients. *Ann Intern Med*. 1990;112:699-706.
- Diokno AC, Brown MB, Brock BM, et al. Clinical and cystometric characteristics of continent and incontinent noninstitutionalized elderly. *J Urol*. 1988;140:567-71.
- Cummings SR, Nevitt MC, Kidd S. Forgetting falls: the limited accuracy of recall of falls in elderly. *J Am Geriatr Soc*. 1988;36:613-6.
- Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res*. 1983;17:37-49.
- Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36), 1: conceptual framework and item selection. *Med Care*. 1992;30:473-83.
- Hays RD, Sherbourne CD, Mazel RM. The RAND 36-item health survey 1.0. *Health Econ*. 1993;2:217-27.
- Sloss EM, Hosek SD. Evaluation of the CHAMPUS Reform Initiative: Beneficiary Access and Satisfaction. Santa Monica, Calif: RAND Corporation; 1993. R-4244/3-HA.
- Maly RC, Abrahamse AS, Hirsch SH, Frank JC, Reuben DB. What influences physicians' practice behavior? An interview study of physicians who received consultative geriatric assessment recommendations. *Arch Fam Med*. 1996;5:448-54.
- Marshall GN. A multidimensional analysis of internal health locus of control beliefs: separating the wheat from the chaff? *J Pers Soc Psychol*. 1991;61(3):483-91.
- National Center for Health Statistics. Current estimates from the National Health Interview Survey: United States 1986. *Vital Health Stat*. 1986;10.
- Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297-334.
- Heath RW, Downie NM. *Basic Statistical Methods*. 5th ed. New York, NY: Harper & Row; 1983:192-3.
- Stewart AL. Psychometric considerations in functional status instruments. In: World Organization of National Colleges, Academies and Academic Associations of General Practice/Family Physicians (WONCA) Classification Committee (eds). *Functional Status Measurement in Primary Care*. New York, NY: Springer; 1990:3-26.
- Verbrugge LM, Jette AM. The disablement process. *Soc Sci Med*. 1994;38:1-14.
- World Health Organization. *International Classification of Impairments, Disabilities, and Handicaps*. Geneva, Switzerland: WHO; 1980.

28. Nagi SZ. Disability concepts revisited: implications for prevention. In: Pope AM, Tarlov AR, eds. *Disability in America: Toward a National Agenda for Prevention*. Washington, DC: Division of Health Promotion and Disease Prevention, Institute of Medicine, National Academy Press; 1991:309-27.
29. Ware JE Jr, Kisiniski M, Keller SD. A 12-item short form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34:220-33.
30. Rubenstein LV, Calkins DR, Young RT, et al. Improving patient function: a randomized trial of functional disability screening. *Ann Intern Med*. 1989;111:836-42.
31. Rubenstein LV, McCoy JM, Cope DW, et al. Improving patient quality of life with feedback to physicians about functional status. *J Gen Intern Med*. 1995;10:607-14.
32. Moore AA, Siu AL, Partridge JM, Hays RD, Adams J. A randomized trial of office-based screening for common problems in older persons. *Am J Med*. 1997;102:371-8.
33. Myers AM, Holliday PJ, Harvey KA, Hutchinson KS. Functional performance measures: are they superior to self-assessments? *J Gerontol*. 1993;48(5):M196-206.
34. Elam JT, Graney MJ, Beaver T, El Derwi D, Applegate WB, Miller ST. Comparison of subjective ratings of function with observed functional ability of frail older persons. *Am J Public Health*. 1991;81(9):1127-30.
35. Sager MA, Dunham NC, Schwantes A, Mecum L, Halverson K, Harlowe D. Measurement of activities of daily living scales in hospitalized elderly: a comparison of self-report and performance-based methods. *J Am Geriatr Soc*. 1992;40:457-62.
36. Rozzini R, Frisoni GB, Bianchetti A, Zanetti O, Trabucchi M. Physical performance test and activities of daily living scales in the assessment of health status in elderly people. *J Am Geriatr Soc*. 1993;41:1109-13.
37. Judge JO, Schechtman K, Cress E, and the FICSIT Group. The relationship between physical performance measures and independence in instrumental activities of daily living. *J Am Geriatr Soc*. 1996;44:1332-41.
38. Reuben DB, Valle LA, Hays RD, Siu AL. Measuring physical function in community-dwelling older persons: a comparison of self-administered, interviewer-administered, and performance-based measures. *J Am Geriatr Soc*. 1995;43:17-23.
39. Kempen GJ, Steverink N, Ormel J, Deeg DJH. The assessment of ADL among frail elderly in an interview survey: self-report versus performance-based tests and determinants of discrepancies. *J Gerontol*. 1996;51B:P254-60.

ANNOUNCEMENT

American Board of Internal Medicine

1999 ABIM Certification Examination in Internal Medicine

Registration Period: September 1, 1998 – December 1, 1998
Examination Dates: August 24-25, 1999

1999 ABIM Certification Examination in Sports Medicine

Registration Period: July 1, 1998 – November 1, 1998
Examination Dates: April 16, 1999

Important Note: The 1999 Sports Medicine Examination is the last one for which Diplomates may qualify through a practice pathway.

For more information and application forms, please contact:

Registration Section
American Board of Internal Medicine
510 Walnut Street, Suite 1700
Philadelphia, PA 19106-3699
Telephone: (800) 441-2246 or (215) 446-3500 Fax: (215) 446-3590 E-mail: request@abim.org