Screening for Common Problems in Ambulatory Elderly: Clinical Confirmation of a Screening Instrument

Alison A. Moore, MD, MPH, Albert L. Siu, MD, MSPH, *Los Angeles, California*

PURPOSE: To develop a test that will be administered by nonphysician office staff to screen elderly persons seen in ambulatory settings for problems commonly contributing to functional disability.

METHODS: We reviewed the literature to identify problems that reduce function and screening measures appropriate for use in office settings. Using this information, we developed an instrument including screening items for malnutrition/weight loss, visual impairment, hearing loss, cognitive impairment, urinary incontinence, depression, physical disability, and reduced leg mobility. The instrument was tested on 109 new patients at a university-based ambulatory geriatrics medicine clinic and validated using two standards: blinded and unblinded geriatricians' assessments. For each of the individual items on the screening instrument, we calculated sensitivity and specificity using both the blinded and unblinded geriatricians' evaluations as the reference standards, prevalence of the disorders, positive and negative predictive values, inter-rater reliability, and the direct annual costs of administering the test for an individual physician's office.

RESULTS: The screen was administered in 8 to 12 minutes. Inter-rater agreement varied by item from 77% to 100%. The sensitivities of the items varied between 0.65 to 0.93 (blinded) and 0.70 to 0.95 (unblinded). Specificities ranged between 0.50 to 0.95 (blinded) and 0.64 to 0.95 (unblinded). Problem prevalences varied from 21% to 72%. Positive and negative predictive values were 0.60 to 0.91 and 0.77 to 0.96, respectively. Direct annual costs for a clinical practice include a one-time $530 fee for equipment and, depending on the screening administrator's salary, between $1 to $7 per patient screened.

CONCLUSIONS: The screening instrument is relatively inexpensive and brief and easy to use in the ambulatory setting. It has good validity and reliability when compared to the assessment of a geriatrician. We are currently conducting a randomized trial to assess the effectiveness of the screen among older persons seen in community physicians' offices.

Functional disability, or difficulty performing one's daily activities, is common among the elderly.[1,2] Although function is not typically evaluated in standard medical assessments, deficits in function predict mortality[3,4] and better predict patients' outcomes following hospitalization than do admitting diagnoses.[5,6] Recognizing the importance of the relationship between functional status and health, major medical organizations advocate physicians' evaluation of patients' physical, mental, and psychosocial function.[7,8] Unfortunately, previous studies using instruments that give physicians information about their patients' function have failed to modify physician's behavior or improve patients' function.[9,10]

One such intervention employed the Functional Status Questionnaire (FSQ), a 34-item self-administered questionnaire that is computer-scored and assesses physical, psychological, social, and role function in ambulatory patients.[11] The FSQ was used in a randomized controlled trial to see whether giving general internists information on their patients' function would enable them to address any disabilities identified.[9] Physicians participated in a 2-hour seminar on how to evaluate and treat functional disabilities and were also given information on their enrolled patients' function every 4 months for 1 year. The FSQ report was fashioned to look like a lab report and the individual's score was given along with a range of normal scores. The educational seminar and feedback on function had no apparent effect on physicians' management of patients.

Dartmouth COOP Charts have also been used in a randomized trial to test their effect on the process of care and satisfaction of patients seeing primary care physicians.[10] COOP Charts are simple cartoon figures that describe various levels of functioning. Six COOP Charts, including physical condition, emotional condition, daily work, social activities, health change, and overall condition, were filled out by patients in their

From the Multicampus Program in Geriatrics and Gerontology (AAM, ALS), the Robert Wood Johnson Clinical Scholars Program, University of California at Los Angeles School of Medicine, Los Angeles, California; and the Department of Health Policy (ALS), Mount Sinai School of Medicine, New York, New York.

This project was funded by the Robert Wood Johnson Clinical Scholars Program.

Requests for reprints should be addressed to Alison A. Moore, MD, MPH, Multicampus Program in Geriatrics and Gerontology, UCLA School of Medicine, Box 951687, Los Angeles, California 90095-1687.

Manuscript submitted April 18, 1995 and accepted in revised form October 31, 1995.

physicians' offices prior to being seen.[12] Showing the COOP Charts to physicians affected only two variables: the ordering of tests and procedures for women was increased, and men reported greater help with functional problems related to pain.

Various explanations have been advanced for why these interventions have achieved disappointing results and why other attempts to measure health status in clinical practice may fail.[13] We hypothesize that one reason these interventions have had little effect is that the functional status measures used in these studies provided physicians with information on broad concepts of function rather than on specific clinical problems that may affect function. Physicians may have difficulty translating information on general aspects of function (eg, overall emotional condition and quality of social interaction) into specific interventions to improve functioning. Clinicians might more effectively intervene to improve their patients' ability to do their daily routines if they could identify specific impairments (eg, hearing loss and incontinence) that adversely affect function. Indeed, the American College of Physicians supports screening for such problems in the ambulatory setting,[14] but no single, comprehensive screening measure or set of screening tools to identify such clinical problems has been tested.

Using and adapting existing screening measures, we developed and tested an instrument to efficiently screen (using nonphysician office staff) elderly patients seen in physician's offices for eight common problems that may adversely affect function and for which intervention may improve patient outcomes.

METHODS

Screen Development

We reviewed the literature to identify problems that commonly contribute to functional disability and brief, easy-to-administer screening measures for these problems. After reviewing this literature, we focused on the following eight problems that are often missed by traditional examinations: malnutrition/weight loss, visual impairment, hearing loss, cognitive impairment, urinary incontinence, depression, physical disability, and reduced leg mobility. For all of these items we then identified and, in some cases, modified screening measures appropriate for use in office settings and for administration by nonphysician office personnel. The selection of items was also based on available information on the feasibility, reliability, and validity of these measures.

The final screening package includes: (1) questions about visual impairment, depression, weight loss, urinary incontinence, and physical functioning; and (2) performance-based tests of malnutrition/weight loss, vision, hearing, cognition, and leg mobility (**Table I**). A Snellen eye chart (Welch-Allyn, Skaneateles Falls, NY) and a Welch-Allyn audioscope are used as part of the screening package.

Clinical Confirmation of the Screen

This study was performed in the University of California at Los Angeles Geriatrics Clinic among a consecutive sample of 109 new patients being seen for geriatric assessment on designated days. A trained research assistant first administered the screen to patients before they saw a geriatrician. Any one of seven geriatricians then comprehensively evaluated the patient, without knowing the screen's results. When evaluating a new patient, these geriatricians typically use the Snellen eye chart or Jaeger card and administer the whispered voice test,[25] Mini-Mental State Examination,[26] Geriatric Depression Scale,[27,28] Katz Activities of Daily Living Scale,[29] Lawton Instrumental Activities of Daily Living Scale,[30] and Performance-Based Mobility Assessment.[31] The geriatrician recorded the presence or absence of any relevant impairments before being given the screening findings.

First, we compared the screen's results to the blinded or initial geriatrician's evaluation as our "gold" standard. In order to further evaluate the clinical utility of the screening package, the geriatrician was then given the screen's results and he/she was given the opportunity to revise his/her original assessment. We also compared the screen's results to this unblinded geriatrician's assessment. We estimated each item's sensitivity, specificity, and predictive value using the blinded or the unblinded geriatrician evaluation as the "gold" standard for the presence or absence of the given impairment.

We chose to present data from both of the geriatrician's evaluations because the blinded assessment is more methodologically rigorous, but in the absence of an ideal gold standard, the combination of information from the screening package and the geriatrician's clinical assessment (unblinded evaluation) provided the most accurate information. This is because for some of the items the screen is better than the physician's assessment (eg, hearing); for some items the geriatrician's assessment is better than the screen (eg, cognition); and for some items there is no gold standard (eg, nutrition).

Using the kappa statistic, we assessed inter-rater reliability by having two nonphysicians independently administer the screening test to 22 patients. Finally, the direct annual costs of using the screening package in an individual medical practice were estimated using available data on the costs of equipment and personnel.

TABLE I

Screening Package Characteristics

Problem	Screening Measure	Positive Screen	Supporting Data
Vision	2 Parts: Ask: "Do you have difficulty driving, or watching television, or reading, or doing any of your daily activities because of your eyesight? If yes, then: Test each eye with Snellen chart while patient wears corrective lenses (if applicable).	Yes to question and inability to read greater than 20/40 on Snellen chart.	Question: derived from some of the most reliable items on the Boston Activities of Daily Vision Scale; test-retest reliability is 0.8[15]; Snellen chart: "gold" standard.
Hearing	Use audioscope set at 40 dB. Test hearing using 1,000 and 2,000 Hz.	Inability to hear 1,000 or 2,000 Hz in both ears or either of these frequencies in one ear.	In physicians' offices: sensitivity = 0.94; specificity = 0.72.[16]
Leg mobility	Time the patient after asking: "Rise from the chair. Walk 20 feet briskly, turn, walk back to the chair and sit down."	Unable to complete task in 15 seconds.	Modified version of the "Up & Go"; inter-rater and test-retest reliability = 0.99; good correlations with other measures of gait and balance (-0.6 to -0.8).[17]
Urinary incontinence	2 Parts: Ask: "In the last year, have you ever lost your urine and gotten wet?" If yes, then ask: "Have you lost urine on at least 6 separate days?"	Yes to both questions.	83% agreement between patient response and urologic assessment.[18,19]
Nutrition/weight loss	2 Parts: Ask: "Have you lost 10 lbs. over the past 6 months without trying to do so?"* Weigh the patient.	Yes to the question or weight <100 lb.	Question: relative risk of death = 2.0* (NHEFS); weight: PPV of malnutrition = 0.99.[20,21]
Memory	Three-item recall.	Unable to remember all three items after 1 minute.	Likelihood ratios: recalls all 3 = 0.06; recalls 2 = 0.5; recalls <2 = 3.1.[22]
Depression	Ask: "Do you often feel sad or depressed?"	Yes to the question.	Sensitivity = 0.78; specificity = 0.87.[23]
Physical disability	Six questions: "Are you able to...: "Do strenuous activites like fast walking or bicycling?" "Do heavy work around the house like washing windows, walls, or floors?" "Go shopping for groceries or clothes?" "Get to places out of walking distance?" "Bathe, either a sponge bath, tub bath, or shower?" "Dress, like putting on a shirt, buttoning and zipping, or putting on shoes?"	Yes to any of the questions.	Coefficient of scalability 0.86; coefficient of reproducibility 0.96; test-retest reliability 0.88; good correlation with other measures of physical function 0.63–0.89.[24]

*Personal communication from Tamara B. Harris, MD.
NHEFS = National Health Epidemiologic Follow-up Study. PPV = postive predictive value.

RESULTS

The mean age of the study population was 79 years (range 62 to 96); 66% were female; 80% were white, 12% were black, 6% were Hispanic, and 1% were Asian. The screening package consists of 9 to 16 items, and it took 8 to 12 minutes for a nonphysician to administer.

Inter-rater agreement varied by item from 77% to 100% ($\kappa > 0.8$, excepting cognition) (**Table II**). Using the first evaluation, or blinded geriatrician's evaluation, as the gold standard, the sensitivities of each of the items ranged from 0.65 for nutrition to 0.93 for hearing (**Table III**). Using the final, or unblinded geriatrician's, assessment as the gold standard, the sensitivity of each of the screen's items was at least 0.70. The specificity of each of the items using the blinded assessment as the gold standard varied between 0.50 for physical disability to 0.95 for incontinence (**Table IV**). Similar values were obtained using the unblinded geriatrician's evaluation as the gold standard.

Using the geriatricians' final evaluation to determine the prevalences of the various problems in which we were interested, the prevalence rates in this population of older persons ranged from 21% for malnutrition to 72% for physical disabilities (**Table V**). Using the final geriatrician's assessment as the gold standard, the positive predictive values for each item varied between 0.60 for memory to 0.91 for mobility. Negative predictive values in this population ranged between 0.77 for physical disability to 0.96 for urinary incontinence.

Finally, we estimated the direct annual costs of the screening package for an individual physician practice. Fixed costs are approximately $530 and include the Welch-Allyn audioscope and an eye chart. To estimate variable costs, we assumed that 100 elderly patients being seen for first visits and yearly comprehensive physical exams would be screened annually. We also assumed that it would take 10 minutes on average to screen each patient. We determined a range of personnel costs depending on the qualifications of the person administering the screen. For example, a "front-office person" without any certification who earns between $5 to $6/h might cost a physician's practice between $83 to $100 per year to screen 100 patients, while a registered nurse who earns between $35 to $40/h might cost a practice $585 to $667 per year to screen these same patients. Since personnel costs account for essentially all of the costs after equipment is paid for, annual variable costs, excluding fringe benefits and overhead, might range from $83 to $667 per practice or $1 to $7 per patient. Physician time may increase (to evaluate screen findings) or decrease (from having certain aspects of the exam performed by other staff). We did not have a basis for estimating this cost nor the cost of follow-up tests and procedures.

TABLE II

Inter-Rater Reliability (n = 22)

Problems	Overall Agreement (%)	Kappa
Nutrition	100	1.00
Vision	100	1.00
Hearing	91	0.81
Memory	77	0.55
Incontinence	95	0.90
Depression	100	1.00
Physical disability	100	1.00
Leg mobility	91	0.82

TABLE III

Screening Package: Sensitivity

Problems	Blinded Assessment % (± 95% CI)	Unblinded Assessment % (± 9.5% CI)
Nutrition	0.65 (0.56, 0.74)	0.70 (0.61, 0.79)
Vision	0.67 (0.58, 0.76)	0.75 (0.67, 0.83)
Hearing	0.93 (0.88, 0.98)	0.95 (0.91, 0.99)
Memory	0.90 (0.84, 0.96)	0.90 (0.84, 0.96)
Incontinence	0.89 (0.83, 0.95)	0.89 (0.83, 0.95)
Depression	0.83 (0.76, 0.90)	0.85 (0.78, 0.92)
Physical disability	0.91 (0.86, 0.96)	0.92 (0.87, 0.97)
Leg mobility	0.88 (0.82, 0.94)	0.88 (0.82, 0.94)

CI = confidence interval.

TABLE IV

Screening Package: Specificity

Problems	Blinded Assessment % (± 95% CI)	Unblinded Assessment % (± 9.5% CI)
Nutrition	0.87 (0.81, 0.93)	0.88 (0.83, 0.95)
Vision	0.86 (0.79, 0.93)	0.89 (0.83, 0.95)
Hearing	0.60 (0.51, 0.69)	0.64 (0.55, 0.73)
Memory	0.64 (0.55, 0.73)	0.64 (0.55, 0.73)
Incontinence	0.95 (0.91, 0.99)	0.95 (0.91, 0.99)
Depression	0.79 (0.71, 0.87)	0.80 (0.72, 0.88)
Physical disability	0.50 (0.41, 0.59)	0.67 (0.58, 0.76)
Leg mobility	0.94 (0.89, 0.99)	0.94 (0.89, 0.99)

CI = confidence interval.

TABLE V

Prevalence of Problems in Study Population and Other Diagnostic Characteristics

Problems	Prevalence (%)	Positive Predictive Value	Negative Predictive Value
Nutrition	21	0.62	0.92
Vision	30	0.75	0.89
Hearing	54	0.75	0.91
Memory	37	0.60	0.92
Incontinence	26	0.86	0.96
Depression	37	0.71	0.90
Physical disability	72	0.88	0.77
Leg mobility	40	0.91	0.92

CONCLUSIONS

This simple screening package to assess common problems in the elderly is feasible for nonphysicians to administer in the context of an office visit and it is brief and easy to use. The screen exhibits good interrater reliability when administered by two nonphysicians. In general, its results agree with a geriatrician's clinical assessment and improves it in some areas. In this population of older persons, the screen performed well in that the sensitivity and specificity for detecting each of the screen-included problems was relatively high. Additionally, given the prevalences of the screened disorders in this population, the screening measure also exhibited high positive and negative predictive power. Finally, the screening package may be a relatively inexpensive way to efficiently and effectively assess problems that commonly adversely affect function in the elderly.

Our goal was to select screening tests that could be implemented by office staff to reasonably rule out the presence of some impairments and thus permit the busy clinician to focus on a more extensive evaluation of the screen's positive findings. For that purpose, each of our screening items would ideally have very high sensitivity along with specificity that was sufficiently high so that the clinician's time would not be wasted evaluating many false positives.

Although the sensitivity of most of our screening items was high, nutrition and vision had notably lower sensitivity. This should not be surprising, given that simple office-based gold standards for these impairments have not previously been identified. It suggests that clinicians who have other reasons to suspect the presence of these conditions need to be aware about possible false negative responses for these tests.

On the other hand, sensitivity was high with low specificity for our hearing, memory, and physical disability screening items. Depending on the prevalence of these conditions, clinicians using the screen should not be surprised by the need to evaluate screen-positive patients who turn out to be false positives on these items. Also, depending on the prevalence of the conditions in the patient population, the number of false positive screens for nutrition and vision may also be considerable (Table V).

The greatest limitations in this study are those of generalizability and our gold standard. Our study was performed in an academic geriatrics practice, and we cannot draw conclusions about the screen's utility in other settings. The prevalence estimates for the screened conditions in the ambulatory elderly population are generally between one third to two thirds of the prevalence rates observed in the study population. The prevalences of the screened disorders in the population of older persons who see physicians in their offices is likely to be in between that of the study population and the general ambulatory elderly population. Since the prevalence rates for most of the studied clinical problems is likely to be lower among elderly persons seeking care from family physicians and internists, the positive predictive values for each of the screen's items will also be lower and the negative predictive values will be higher. We are currently conducting a randomized trial to assess the impact of the screen on the health of elderly patients seen in community physicians' offices.

Perhaps a better standard with which to test the screen would have been a multidisciplinary team of health professionals to evaluate the various problems included in the studied screening package, that is, an audiologist to assess hearing, a psychiatrist to evaluate depression, etc. However, despite the absence of a better or ideal gold standard to test the screening package, we argue that using a combination of the geriatrician's clinical judgment and data from the screener provides the most clinically useful standard to evaluate the utility of the screening package.

The screening package has many features that should make it useful in busy clinical settings for identifying patients who are at risk of functional decline. This measure can be administered by any trained member of a physician's staff. In addition, patients may be screened during longer visits such as an initial visit or annual physical examination, when the physician may be more interested in an overall evaluation of the patients' clinical problems rather than focusing on a specific complaint. Although the screening is brief, physicians may wish to administer different components of the screening package over the course of a number of visits rather than administering it in its entirety in one visit. Further, clinicians may want to use this instrument or modify it so that they may identify problems prevalent among their particular patient populations. We believe that this package meets the needs of primary care physicians who are being asked to care for an increasing number of elderly persons in an increasingly limited amount of time by enabling them to focus on particular problems that may commonly compromise their patients' health and functioning.

ACKNOWLEDGMENT

The authors would like to thank Jennifer M. Partridge for her invaluable assistance with data collection and management.

REFERENCES

1. Guralnik JM, Simonsick EM. Physical disability in older Americans. *J Gerontol.* 1993;48:3–10.

2. Manton KG, Corder LS, Stallard E. Estimates of change in chronic disability and institutional incidence and prevalence rates in the U.S. elderly population from the 1982, 1984, and 1989 National Long Term Care Survey. *J Gerontol.* 1993;48:S153–S166.

3. Manton KG. A longitudinal study of functional change and mortality in the United States. *J Gerontol.* 1988;48:S153–S161.

4. Reuben DB, Rubenstein LV, Hirsch SH, Hays RD. Value of functional status as a predictor of mortality: results of a prospective study. *Am J Med.* 1992;93: 663–669.
5. Narain P, Rubenstein LZ, Wieland GD, et al. Predictors of immediate and 6-month outcomes in hospitalized elderly patients. *J Am Geriatr Soc.* 1988;36: 775–783.
6. Winograd CH, Gerety MB, Chung M, et al. Screening for frailty: criteria and predictors of outcomes. *J Am Geriatr Soc.* 1991;39:778–784.
7. Almy TP. Comprehensive functional assessment for elderly patients. *Ann Intern Med.* 1988;109:70–72.
8. Rubenstein LV, Calkins DR, Greenfield S, et al. Health status assessment for elderly patients: report of the Society of General Internal Medicine Task Force on Health Assessment. *J Am Geriatr Soc.* 1989;37:562–569.
9. Rubenstein LV, Calkins DR, Young RT, et al. Improving patient function: a randomized trial of functional disability screening. *Ann Intern Med.* 1989;10: 836–842.
10. Wasson J, Hays R, Rubenstein L, et al. The short-term effect of patient health status assessment in a health maintenance organization. *Qual Life Res.* 1992;1:99–106.
11. 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;1:143–149.
12. Nelson EC, Landgraf JM, Hays RD, et al. The functional status of patients: how can it be measured in physicians' offices? *Med Care.* 1990;28:1111–1126.
13. Nelson EC, Berwick DM. The measurement of health status in clinical practice. *Med Care.* 1989;27:S77–S90.
14. 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.
15. Mangione CM, Phillips RS, Seddon JM, et al. Development of the "Activities of Daily Vision Scale": a measure of visual functional status. *Med Care.* 1992;30:1111–1126.
16. Lichtenstein MJ, Bess FH, Logan SA. Validation of screening tools for identifying hearing-impaired elderly in primary care. *JAMA.* 1988;259:2875–2878.
17. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc.* 1991;39:142–148.
18. Diokno AC, Brock BM, Brown MR, Herzog AR. Prevalence of urinary incontinence and other urological symptoms in the noninstitutionalized elderly. *J Urol.* 1986;136:1022–1025.
19. Diokno AC, Brown MR, Brock BM, et al. Clinical and cystometric characteristics of continent and incontinent noninstitutionalized elderly. *J Urol.* 1988; 140:567–571.
20. National Center for Health Statistics. Plan and Operation of the NHANES I-Epidemiologic Follow-up Study 1982–1984. Vital and health statistics. DHHS Pub. No. (PHS) 87-1324. Hyattsville, Md: Public Health Service; 1987: Ser. 1, No. 22.
21. Manson A, Shea S. Malnutrition in the elderly ambulatory medical patients. *Am J Public Health.* 1991;81:1195–1197.
22. Siu AL. Screening for dementia and investigating its causes. *Ann Intern Med.* 1991;115:122–132.
23. Abler R, Drinka T, Mahoney J, et al. Depression in patients of a geriatric medicine clinic: comparison of two screening instruments. *Gerontologist.* 1991;31 (special issue II:)325. Abstract.
24. Siu AL, Reuben DB, Hays RD. Hierarchical measures of physical function in ambulatory geriatrics. *J Am Geriatr Soc.* 1990;38:1113–1119.
25. Mulrow CD, Lichtenstein MJ. Screening for hearing impairment in the elderly: rationale and strategy. *J Gen Intern Med.* 1991;6:249–258.
26. 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;12:189–198.
27. Brink TL, Yesavage JA, Lum O, et al. Screening tests for geriatric depression. *Clin Gerontol.* 1982;1:37–43.
28. 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.
29. Katz S, Ford AB, Moskowitz RW, et al. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. *JAMA.* 1963;185:914–919.
30. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist.* 1969;9:179–186.
31. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. *J Am Geriatr Soc.* 1986;34:119–126.