BRIEF REPORT

Development and Validation of a Geriatrics Knowledge Test for Primary Care Residents

David B. Reuben, MD, Ming Lee, PhD, James W. Davis, MD, Michelle S. Eslami, MD, Daniel G. Osterweil, MD, Susan Melchiorre, MD, Nancy T. Weintraub, MD

This study reports the development and preliminary validation of an instrument to measure geriatrics knowledge of primary care residents. A 23-item test was developed using questions selected from the American Geriatrics Society's Geriatrics Review Syllabus. Ninety-six internal medicine and family practice residents, 14 geriatrics fellows, and 11 geriatrics faculty members participated in the study. Findings support the reliability (Cronbach's $\alpha = 0.66$) and validity (content and "known groups") of this short test. Predictive validity and sensitivity of the test to changes in knowledge will have to be further explored as residents progress through their training.

KEY WORDS: geriatrics knowledge assessment; evaluation of residency programs; geriatrics education.

J GEN INTERN MED 1997;12:450-452.

As primary care residency programs increasingly incorporate geriatrics into their training, it will be important to document their effectiveness. One of the difficulties in evaluating residency programs is lack of instruments that are geared to the level of geriatrics knowledge in residency training. Although Palmore's Facts on Aging Quiz series¹⁻⁵ is the most frequently used instrument for measuring knowledge about aging, the instrument is not designed to measure clinical knowledge of geriatric patient care, and its psychometric properties have been questioned.^{6,7} In this study, we developed an instrument that measures geriatric content knowledge, and conducted validity assessments to assess the appropriateness of using the instrument to evaluate residency programs and to help identify specific geriatrics educational needs of residents.

Received from the University of California, Los Angeles (DBR, ML, JWD, MSE, DGO, NTW), and Adults Primary Care, Coeur d'Alene, Idaho (SM).

Presented at the annual meeting of the American Geriatrics Society, Chicago, Ill., May 2-5, 1996.

Supported by the John A. Hartford Foundation and the Department of Veterans Affairs Medical Centers—West Los Angeles and Sepulveda, Calif.

Address correspondence and reprint requests to Dr. Reuben: Department of Medicine-Geriatrics, University of California, Los Angeles, 10945 Le Conte Ave., Suite 2339, Los Angeles, CA 90095-1687.

METHODS

Instrument Development

The test was designed to span content knowledge relevant to different practice settings and a variety of geriatric content areas, and to be brief enough for practical administration to physicians-in-training. We also recognized that different dimensions of geriatrics knowledge (e.g., knowledge of clinical syndromes vs knowledge of social support) may exist, but that such dimensions would probably not be adequately discriminated in a brief instrument; hence, we conceptualized the test as measuring general geriatrics knowledge.

Thirty-three questions that were appropriate to the level of residency training were selected by five geriatrics faculty from 355 items included in the American Geriatrics Society (AGS) Geriatrics Review Syllabus (GRS) series,^{8,9} with permission of the AGS. Most of the questions were case-oriented questions about medical assessment or management.

The 33-item test was then administered to UCLA internal medicine family practice residents, geriatrics faculty, and fellows of the VA-UCLA Geriatric Medicine Fellowship Program. On the basis of psychometric analyses, a final set of 23 items was selected. The criteria used for eliminating the 10 questions were (1) low item correlation with the total score of the remaining items, (2) poor item discriminatory power among groups as indicated by the percentage of correct responses made by each group, and (3) poor item construction or extremely difficult items as inferred from poor item performance of the faculty and GRS participants.

Analysis

To evaluate the psychometric properties of the test, reliability was estimated by Cronbach's coefficient α ,¹⁰ floor and ceiling effects were determined, known-groups validity¹¹ was assessed by analyses of variance, and two-tailed pairwise *t* tests were performed for individual group comparisons. Pearson correlation coefficients were calculated to compare the performance of individual items with the total score of the remaining items and to compare the percentages of correct item responses of study participants and of GRS subscribers using data provided by the

AGS. Finally, analysis of variance was performed to compare scores of residents with different degrees of interest in geriatrics as a career. We hypothesized that with progressive training and interest in geriatrics, knowledge scores would be higher.

RESULTS

The response rate in all groups was 100%. A total of 121 subjects participated in the study, including 72 internal medicine residents (postgraduate year-1 [PGY-1] = 30, PGY-2 = 25, PGY-3 = 17), 24 family practice residents (8 in each year of the residency), 14 geriatrics fellows, and 11 geriatrics faculty members. Seventy-three respondents (60%) were men and 48 (40%) were women. The majority (57%) of the respondents who gave their racial identity ($n = 106$) were whites, 37% were Asian Americans, and the remaining were Hispanic-Americans (5%) and African-Americans (2%).

The Cronbach's coefficient α of the 23-item test was 0.66. Only one faculty member obtained a perfect score (ceiling), while no one failed all items (floor). The correlation between the percentage of correct answers made by the study respondents and those made by the GRS participants was 0.64 ($p = .001$). The percentage of correct responses for each item ranged from 0.17 to 0.82 (median = 0.56), and the correlation of each item with the total score of the remaining items ranged from 0.12 to 0.50 (median = 0.22).

Mean scores (Table 1) increased with the level of training ($p = .0001$). Faculty scored higher than each trainee group, PGY-3 residents and geriatric fellows scored equally well, and the mean score for PGY-3 residents was higher than the score for PGY-1 residents.

No significant differences were found between resident groups identified by specialty (internal medicine vs family practice), gender, race, or age. Ninety-three residents responded to the question about their interest in selecting geriatrics as their profession. Among them, 20% ($n = 19$) expressed being "very interested" or "somewhat interested," 36% ($n = 33$) "uncertain," and 44% ($n = 41$) "not interested" or having "no interest at all." There was a small but significant ($p = .039$) difference in knowledge

scores among these three groups, with mean scores of 11.00, 10.45, and 12.02, respectively.

DISCUSSION

The ability of residency programs to document the effectiveness of efforts to increase geriatrics training will provide credibility and may facilitate dissemination of successful educational methods. Such documentation requires that reliable and valid evaluation instruments be available. The 23-item knowledge test described in this study meets many of the criteria for reliability and validity. It showed no floor or ceiling effects as a whole or on any individual item among physician trainees. The average score of residents and fellows was approximately 11, a desirable mean¹² indicating substantial room for improvement. The test also demonstrated acceptable reliability for research purposes or for group comparisons.¹² The validity of this instrument is supported by high correlation of scores between the study participants and AGS GRS participants and the higher scores with increasing levels of geriatrics training (known-groups validity).¹¹

We observed two unexpected findings. First, the mean score of the fellows was not substantially higher than that of the resident groups. This may be explained in part by the previous training and background of these trainees. The majority (64%) of the participating fellows were international medical graduates, whereas 96% of the residents received training in U.S. medical schools. Second, residents expressing no interest in choosing geriatrics as a career scored higher than those who were uncertain or interested. This finding may indicate that some of the brightest and most knowledgeable residents have learned basic principles of geriatrics but have little interest in pursuing the discipline as a career. Although this explanation of findings is plausible, further research will have to determine whether this finding is, in fact, evidence against the validity of the instrument.

Despite the potential usefulness of this knowledge test, some limitations should be recognized. First, this was a single-site validation study, and performance of the instrument may vary at other institutions. Second, the in-

Table 1. Differences in Knowledge Test Scores by Level of Training

Group	n	Mean	SD	Range*	F Test	p Value†			
						PGY-1	PGY-2	PGY-3	Fellows
Total	121	12.12	3.69	5-23	28.18‡				
PGY-1	38	10.39	2.77	5-17					
PGY-2	33	11.61	2.76	6-19		0.059			
PGY-3	25	12.00	2.31	8-17		0.021	0.579		
Fellows	14	12.00	3.06	5-17		0.057	0.645	1.000	
Faculty	11	20.00	2.28	15-23		<.0001	<.0001	<.0001	<.0001

*The range of possible scores is 0-23.

†Values were calculated using two-tailed *t* tests.

‡Statistically significant at the $p = .0001$ level.

strument may not cover all of the content that is deemed important in geriatrics education. Third, by the nature of our cross-sectional study design, the validation process is incomplete. Still remaining to be determined are the instrument's sensitivity to changes, a potential learning effect on post-test scores, and the predictive validity of the test.

In summary, this study supports the reliability and validity of a short test of geriatrics knowledge designed for medical residents. The sensitivity of the test to changes in knowledge after training will have to be determined later as the residents progress through their education.

REFERENCES

1. Palmore E. Facts on aging: a short quiz. *Gerontologist*. 1977;17:315-20.
2. Palmore E. The Facts on Aging Quiz: part two. *Gerontologist*. 1981;21:431-7.
3. Duerson MC, Thomas JW, Chang J, Stevens B. Medical students' knowledge and misconceptions about aging: responses to Palmore's Facts on Aging Quizzes. *Gerontologist*. 1992;32:171-4.
4. Holtzman JM, Beck JD. Palmore's Facts on Aging Quiz: a reappraisal. *Gerontologist*. 1979;19:116-20.
5. Shahidi S, Devlen J. Medical students' attitudes to and knowledge of the aged. *Med Educ*. 1993;27:286-8.
6. Klemmack DL. Comments: an examination of Palmore's Facts on Aging Quiz. *Gerontologist*. 1978;18:403-6.
7. Miller RB, Dodder RA. A revision of Palmore's Facts on Aging Quiz. *Gerontologist*. 1980;20:673-9.
8. Beck JC, ed. *Geriatrics Review Syllabus: A Core Curriculum in Geriatrics Medicine, Book I: Syllabus and Questions*. 2nd ed. New York, NY: The American Geriatrics Society; 1991.
9. Reuben DB, Yoshikawa TT, Besdine RW, eds. *Geriatrics Review Syllabus Supplement: A Core Curriculum in Geriatrics Medicine, Book I: Syllabus and Questions*. 2nd ed. New York, NY: The American Geriatrics Society; 1993.
10. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297-334.
11. Stewart A. Psychometric considerations in functional status instruments. In: *World Organization of National Colleges, Academies and Academic Associations of General Practice/Family Physicians (WONCA) Classification Committee (ed.), Functional Status Measurement in Primary care*. New York, NY: Springer; 1990.
12. Nunnally JC. *Psychometric Theory*. 2nd ed. New York, NY: McGraw-Hill; 1978.