

BRIEF REPORT: Brief Instrument to Assess Geriatrics Knowledge of Surgical and Medical Subspecialty House Officers

Brent C. Williams, MD, MPH,¹ James T. Fitzgerald, PhD^{2,3}

¹Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA; ²Department of Medical Education, University of Michigan, Ann Arbor, MI, USA; ³Department of Veterans Affairs, University of Michigan, Ann Arbor, MI, USA.

PROBLEM: Initiatives are underway to increase geriatrics training in nonprimary care disciplines. However, no validated instrument exists to measure geriatrics knowledge of house officers in surgical specialties and medical subspecialties.

METHODS: A 23-item multiple-choice test emphasizing inpatient care and common geriatric syndromes was developed through expert panels and pilot testing, and administered to 305 residents and fellows at 4 institutions in surgical disciplines (25% of respondents), emergency medicine (29%), medicine subspecialties (19%), internal medicine (12%), and other disciplines (15%).

RESULTS: Three items decreased internal reliability. The remaining 20 items covered 17 topic areas. Residents averaged 62% correct on the test. Internal consistency was appropriate (Cronbach's α coefficient = 0.60). Validity was supported by the use of expert panels to develop content, and by overall differences in scores by level of training ($P < .0001$) and graded improvement in test performance, with 58%, 63%, 62%, and 69% correct responses among HO1, HO2, HO3, and HO4s, respectively.

CONCLUSIONS: This reliable, valid measure of clinical geriatrics knowledge can be used by a wide variety of surgical and medical graduate medical education programs to guide curriculum reform or evaluate program performance to meet certification requirements. The instrument is now available on the web.

KEY WORDS: measurement; internship and residency; education; surgery; specialists.

DOI: 10.1111/j.1525-1497.2006.00433.x
J GEN INTERN MED 2006; 21:490-493.

America is aging. Approximately 6,000 Americans turn 65 each day, a number that will increase to 10,000 per day by the year 2012.¹ Correspondingly, older adults account for a growing proportion of patients seen by specialty and subspecialty physicians.² Currently, few physicians who care for older adults have formal training in geriatrics.¹ Geriatricians, leading physicians in surgical specialties and medical subspecialties, and several specialty societies have called for increased training in geriatrics among surgical specialists and medical subspecialists.²⁻⁶ Initiatives by the John A. Hartford Foundation, the Donald W. Reynolds Foundation, and the American Geriatrics Society are underway to expand geriatrics training among house officers in surgical specialties and medical subspecialties.^{2,3,7,8}

To assess the effectiveness of geriatrics training initiatives, methods are needed to evaluate trainees' clinical knowledge, skills, and behaviors in the care of older patients. Although there are reliable and valid instruments to measure geriatrics attitudes among medical students and house officers⁹⁻¹¹ and instruments to assess geriatrics knowledge among medical students¹² and primary care house officers,¹³ there are no instruments that measure geriatrics knowledge of house officers in surgical specialties and medical subspecialties. Current knowledge assessment instruments are not appropriate for surgical specialty and medical subspecialty house officers; they are either too simple for this level of training or lack relevance for these practice areas. To address this limitation, as part of a larger initiative supported by the Donald W. Reynolds Foundation to develop curricula in the care of older adults in nonprimary care disciplines, we developed a brief written instrument to measure clinical geriatrics knowledge among house officers in surgical specialties and medical subspecialties. This paper describes the development and evaluation of the reliability and validity of the instrument, termed the University of Michigan (UM) Geriatrics Clinical Decision Making Assessment instrument.

METHODS

Instrument Development

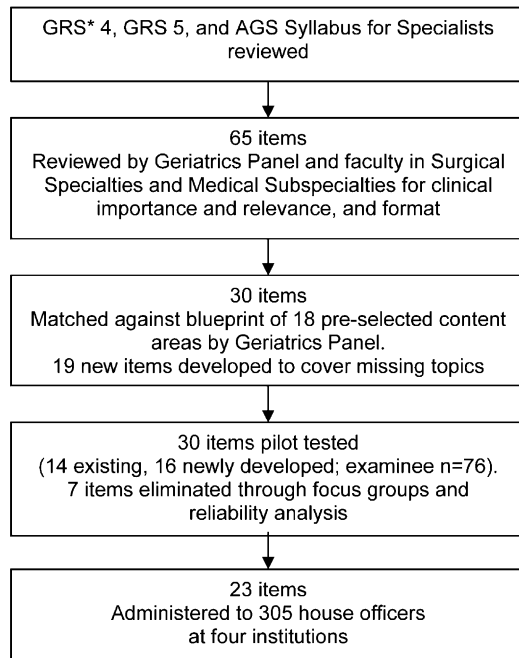
Test Criteria. The length and format of the instrument were modeled after the UCLA geriatrics knowledge test,¹³ which was designed to assess geriatrics knowledge among primary care residents. Specifically, our instrument was designed to include: (a) a focus on clinical management rather than nonclinical issues (e.g., social support, long-term care, health systems organization, and finance; to maintain coherence, relevance, and brevity); (b) a case-based multiple choice format (to enhance clinical relevance and reliability of scoring¹⁴); and (c) a length of 20 to 25 items (to allow a completion time of 30 minutes or less).

Item Selection and Review. A 23-item version of the instrument was developed at the UM through a multistage process (Fig. 1). In overview, candidate items were identified from existing instruments, supplemented with new items covering key topics, and refined through pilot testing. Specifically, a total of 65 candidate items were identified from Versions 4 and 5 of the American Geriatrics Society (AGS) Geriatrics Review Syllabus^{14,15} and the AGS Geriatrics Syllabus for Specialists¹⁶ that covered topics other than outpatient primary care encounters (e.g., hospital care, functional assessment, ethics). Thirty of 65 items were then selected through an iterative review process by an 8-member panel of geriatrics-trained specialists (1 cardiologist, gastroenterologist, and nephrologist; 2 general geriatricians; a psychiatrist; a neurologist, and a social worker) and a 6-member panel of surgeons and medical subspecialists who

The authors have no conflicts of interest to declare for this article.

Pilot data for the work described in this manuscript were presented at the American Geriatrics Society in May 2003, Baltimore, MD. The final results described in this manuscript were presented at a poster session at the national SGIM meeting in May, 2005, New Orleans, LA.

Address correspondence and requests for reprints to Dr. Williams: University of Michigan, 300 N. Ingalls Building, 7E18, Ann Arbor, MI 48109-0429 (e-mail: bwilliam@umich.edu).



* GRS: Geriatrics Review Syllabus; AGS: American Geriatrics Society

FIGURE 1. Development of the University of Michigan Geriatrics Clinical Decision Making Assessment Instrument.

were developing geriatrics curricula for their programs (emergency medicine, gynecology, oncology, rheumatology, thoracic surgery, and urology). These 2 panels selected items based on clinical importance, relevance to the clinical practice of physicians in surgical specialties and medical subspecialties, and wording, using published criteria for item-writing.¹⁷

Following the item selection process, the geriatrics panel compared the 30 items with a test blueprint developed by the panel that included 18 topic areas deemed essential for physicians in surgical specialties and medical subspecialties based on published recommendations.⁴ The 30 items covered the following 10 topic areas: abuse/neglect, advance directives, adverse drug reactions, alcohol abuse, delirium, depression, functional decline in the hospital, adjusting care when prognosis is limited, care level options (e.g., subacute care, nursing home care), and palliative care. Eight topics were not covered: pain management, dementia, constipation, pressure ulcers, falls in the hospital, iatrogenic urinary incontinence, deconditioning/immobility, and infections. To enable inclusion of at least 1 item from each topic area, 19 new items were developed by the panel. To keep the instrument brief, 30 items (14 existing, 16 developed items) that represented all 18 content areas were selected from the 49-item pool for pilot testing.

Pilot Study. The 30-item pilot instrument was administered to 56 house officers in emergency medicine, gynecology, medical oncology, rheumatology, and nephrology; 9 medical students, and 12 general medicine faculty members at the UM. Two focus groups of house officers in gynecology and oncology reviewed and provided feedback on each item for clarity and content. Seven items were reported by the focus groups to have poor response categories, demonstrated limited reliability, or had a greater than 95% correct response rate. The remaining 23 items had an appropriate Cronbach's α reliability coefficient

of 0.55. Validity was supported as individuals with more training tended to score better. For example, faculty had the highest scores and medical students had the lowest scores.

Participants. Participation in the multiinstitutional study was solicited at a national meeting (May, 2003) following a presentation of the pilot study results. From 15 institutions that expressed interest in administering the instrument locally, 6 that had at least 50 house officers in a surgical specialty and/or a medical subspecialty were selected. A standard protocol for administering the instrument under supervision in large group settings was used at each institution after obtaining Institutional Review Board (IRB) approval. Three institutions returned usable responses by June, 2004, the end of the study period. A total of 305 assessments were returned by the 3 participating institutions and the UM. Of these, 149 (49%) were from the UM, 110 (36%) from a second institution, and the remaining 45 (15%) from the other 2 institutions.

Statistical Methods. Cronbach's coefficient α was used to measure test reliability.¹⁸ To examine test validity, test scores were examined by level of training using one-way analysis of variance. Differences between training levels were determined by the Tukey-Kramer (honestly significant difference) HSD test. Validity is supported if test scores improve with increasing levels of training. To examine institutional effects, stratified analyses were carried out among UM and nonUM institutions separately.

RESULTS

Respondents

Among the 305 respondents, 105 (35%) were women. Ninety-three percent of respondents described themselves as either white (69%) or Asian (24%). A total of 25% of respondents were from surgical disciplines, 29% from emergency medicine, 19% from subspecialties of Internal Medicine, 12% from Internal Medicine residency, and 15% from other disciplines.

Reliability

Three items decreased internal reliability and were dropped. For the remaining 20 items covering 17 content areas, Cronbach's coefficient α among all respondents was 0.60 (Table 1). The reliability coefficient was lower for the UM participants (0.46) and higher for the participants from other institutions (0.65).

Validity

Test scores improved as training levels increased (Table 1). HO1s had the lowest mean score and HO4 to HO7 had the highest. This pattern was significant overall; similar patterns were observed among UM and nonUM institutions in stratified analyses, though UM HO1s scored relatively high (Table 1).

DISCUSSION

As the proportion of older adults cared for by surgical specialists and medical subspecialists grows, it will be important to demonstrate the effectiveness of training aimed at enhancing their ability to meet the needs of older patients. The UM Geriatrics Clinical Decision Making Assessment is intended to

Table 1. Reliability and Graded Response Measurement of 20-item UM Geriatrics Clinical Decision Making Assessment

House Officers	Number of HOs (% of Group)	Mean (SD) Percent Correct Responses	Statistics*
All	305 (100)	62 (16)	$\alpha=0.60$
HO1	109 (36)	58 (15)	A
HO2	73 (25)	63 (16)	A B
HO3	40 (13)	62 (15)	A B
HO4 to HO7 [†]	83 (27)	69 (14)	B
University of Michigan	146 (100)	66 (13)	$\alpha=0.46$
HO1	21 (14)	67 (9)	A B
HO2	38 (26)	62 (15)	A
HO3	22 (15)	63 (15)	A B
HO4 to HO7	65 (44)	70 (11)	B
Other institutions	159 (100)	59 (17)	$\alpha=0.65$
HO1	88 (55)	55 (17)	A
HO2	35 (22)	64 (16)	A
HO3	18 (11)	61 (17)	A
HO4 to HO7	18 (11)	66 (20)	A

* α , Cronbach's coefficient α . Letters (A, B, C) indicate subgroups that are not statistically significantly different by Tukey-Kramer HSD test. Overall $P < .02$ from 1-way analyses of variance for all 3 analyses.

[†]Includes all HOs at or above HO4 level. Number of HOs: HO4=38, HO5=24, HO6=17, HO7=4.

HO, house officer (resident or fellow); SD, standard deviation; HSD, honesty significant difference; UM, University of Michigan.

provide program directors a general indicator of geriatrics clinical knowledge among groups of house officers, rather than a comprehensive assessment of individual physicians' knowledge related to the care of older patients. The instrument was designed to be brief to enhance administration to large groups of house officers with busy schedules. The instrument therefore included only a small number of items for each clinical topic. Differences within individuals in mastery of different content areas within geriatrics will not be well detected by this instrument because of the limited number of items related to each topic.

The instrument demonstrated adequate internal consistency. Although the Cronbach's α was modest at 0.60, we consider this appropriate for an instrument intended to measure knowledge in a variety of knowledge domains related to clinical geriatrics, and adequate to achieve our primary objective – to provide a meaningful measure of overall knowledge of geriatrics among groups of learners.

The results of this study also support the validity of our instrument. Item content was chosen through an iterative review process by experts using a predetermined "blueprint." Test performance improved with increasing levels of training, although no detectable difference was observed in the performance of HO2s and HO3s. Also, no ceiling or floor effects were found among individual or all items, and mean scores were appropriate to allow for responsiveness to change over time.

There are important limitations of this study. Some differences in instrument performance were observed across institutions. However, patterns were similar, and differences were smaller where larger numbers of respondents were available, as would be expected. Finally, the predictive validity, potential learning effects on scores of subsequent administrations, and responsiveness to changes in knowledge over time have not been demonstrated.

Surgical specialists and medical subspecialists should be able to recognize and manage issues common among older patients in order to maximize functional and health outcomes. The UM Geriatrics Clinical Decision Making Assessment should prove useful to training programs in surgical specialties and medical subspecialties to gather baseline data, implement curricula, and measure the effects of curricular change on the knowledge of trainees in caring for the complex and interacting needs of older adults. The instrument is now available free on the Web at <http://www.med.umich.edu/geriatrics/educationalprograms/gme.htm>.

The authors wish to acknowledge the contribution of the faculty that developed the assessment instrument: Pam Davies, Gary Faerber, Norman Foster, Scott Gitlin, Karen Hall, Robert Hogikyan, Helen Kales, Irene Kazmers, Eve Losman, Sue Maxixner, Kathie Supiano, Peter Vaitkevicius, Jocelyn Wiggins, and John Yee; the facilitators at participating institutions including Lowell Gerson (Northeastern Ohio University College of Medicine), Manish Shah (University of Rochester School of Medicine and Dentistry), and Belinda Vicioso (University of Texas Southwestern Medical Center); and John B. Murphy, editor of the American Geriatrics Society Geriatric Review Syllabus (GRS) for permission to incorporate GRS items in the instrument. This work was supported by a grant from the Donald W. Reynolds Foundation.

REFERENCES

1. **Alliance for Aging Research.** Medical never-never land: 10 reasons why America is not ready for the coming age boom. February, 2002. Available at: <http://www.agingresearch.org/bookshelf.cfm>. Accessed March 10, 2003.
2. **Solomon DH, Burton JR, Lundebjerg NE, Eisner J.** The new frontier: increasing geriatrics expertise in surgical and medical specialties. *J Am Geriatr Soc.* 2000;48:702-4.
3. **Hazzard WR.** Integrating geriatrics into the subspecialties of internal medicine: Project overview. September 2000. Available at: http://www.americangeriatrics.org/ags_hart_main.shtml. Accessed March 13, 2003.
4. A statement of principles: toward improved care of older patients in surgical and medical specialties. *J Am Geriatr Soc.* 2000;48:699-701.
5. **Council on Resident Education in Obstetrics and Gynecology.** Basic Geriatric Care Objectives for Residency Training in Obstetrics and Gynecology. Washington, DC: American College of Obstetricians and Gynecologists; 1999.
6. Sanders AB (ed). Emergency Care of the Elder Person. Geriatric Medicine Task Force. Society for Academic Emergency Medicine. St. Louis, MO: Beverly Cracom Publications; 1996.
7. American Geriatrics Web site. Available at: http://www.americangeriatrics.org/ags_hart_main.shtml. Accessed March 13, 2003.
8. D.W. Reynolds Foundation Website. Available at: <http://www.dwreynolds.org/Financials/Grants.htm>. Accessed March 13, 2003.
9. **Maxwell AJ, Sullivan N.** Attitudes toward the geriatric patient among family practice residents. *J Am Geriatr Soc.* 1980;28:341-5.
10. **Reuben DB, Fullerton JT, Tschann JM, Croughan-Minihane M.** Attitudes of beginning medical students toward older persons: a five-campus study. The University of California Academic Geriatric Resource Program Student Survey Research Group. *J Am Geriatr Soc.* 1995; 43:1430-6.
11. **Intrieri RC, von Eye A, Kelly JA.** The aging semantic differential: a confirmatory factor analysis. *Gerontologist.* 1995;35:616-21.
12. **Palmore E.** Facts on aging: a short quiz. *Gerontologist.* 1977;17: 315-20.
13. **Reuben DB, Lee M, Davis JW, Eslami MS, Osterweil DG, Melchiorre S, Weintraub NT.** Development and validation of a geriatrics knowledge test for primary care residents. *J Gen Intern Med.* 1997;12:450-2.
14. Cobbs EL, Duthie EH, Murphy JB, eds. Geriatrics Review Syllabus: A Core Curriculum in Geriatric Medicine. 5th edn. Malden, MA: Blackwell Publishing for the American Geriatrics Society; 2002.

-
15. Cobbs EL, Duthie EH, Murphy JB, eds. *Geriatrics Review Syllabus: A Core Curriculum in Geriatric Medicine*. 4th edn. Dubuque, IA: Kendall/Hunt Publishing Company for the American Geriatrics Society; 1999.
 16. **Katz PR, Grossberg GT, Potter JF, et al.** *Geriatrics Syllabus for Specialists*. New York, NY: American Geriatrics Society; 2002.
 17. **Case SM, Swanson DB.** *Constructing Written Test Questions for the Basic and Clinical Sciences*. 2nd edn. Philadelphia, PA: National Board of Medical Examiners (3rd edition available at: <http://www.nbme.org/PDF/2001iwgindex.pdf>, 1998).
 18. **Cronbach LJ.** Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297-334.