Health Outcomes Assessment in Vulnerable Populations: Measurement Challenges and Recommendations

Elizabeth A. Hahn, MA, David Cella, PhD

ABSTRACT. Hahn EA, Cella D. Health outcomes assessment in vulnerable populations: measurement challenges and recommendations. Arch Phys Med Rehabil 2003;84 Suppl 2:S35-42.

With growing recognition that some population subgroups are particularly vulnerable to receiving suboptimal health care and achieving poor health outcomes, innovative techniques are required for collecting and evaluating health outcomes data. Research is also needed to better understand the causal pathways linking vulnerability with health outcomes. This article focuses on patients with a chronic illness (cancer) who also have low literacy and/or poor English language skills. We summarize the association among literacy, language, ethnicity, and health outcomes; describe innovative technologies to enhance communication; and discuss the advantages of using psychometric measurement models in health outcomes assessment. Results from our ongoing research projects are presented, including the development of an audiovisual computer-based testing platform for self-administration of questionnaires. Such innovative multimedia technologies allow patients with limited or even no reading ability to participate in outcomes assessment and have the potential to be incorporated into a clinical setting with minimal burden on staff and patients. Appropriate methods are also needed to evaluate measurement equivalence across diverse patient groups, that is, the extent to which items in a questionnaire perform similarly across groups. Item response theory measurement models provide a strategy for differentiating between measurement bias and real differences that may exist between groups. Recommendations for clinical practice and research are offered specifically to address medically underserved and vulnerable populations.

Key Words: Cancer; Computers; Medically underserved areas; Outcome assessment (health care); Rehabilitation.

© 2003 by the American Congress of Rehabilitation Medicine

THERE IS GROWING RECOGNITION that some population subgroups are particularly vulnerable to receiving suboptimal health care and achieving poorer health outcomes compared with the general population. In 1998, the President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry[1] recognized that vulnerability is mul-

tifaceted and may be because of financial circumstances or place of residence; health, functional, or developmental status; ability to communicate effectively; or age, race, ethnicity, or gender. This definition encompasses populations who are vulnerable because of a chronic or terminal illness or disability and those with literacy or language difficulties. This article focuses on the health care challenges confronting cancer patients who may also have low literacy and/or poor English language skills. Cancer is an excellent model for understanding the impact of chronic, life-threatening illness on health outcomes and preferences for treatment over time. We briefly summarize the literature available on the association among literacy, language, ethnicity, and health outcomes; describe innovative technologies to enhance communication with diverse patient populations; and discuss some advantages of using psychometric measurement models in health outcomes assessment. We share some results from our ongoing research projects and conclude with some recommendations for clinical practice and research.

LITERACY, LANGUAGE, ETHNICITY, AND HEALTH OUTCOMES

Low literacy is a widespread but neglected problem in the United States. The 1992 National Adult Literacy Survey[2] (NALS) measured 3 kinds of literacy tasks that adults encounter in daily life (prose literacy, document literacy, quantitative literacy). Results from the NALS suggest that approximately 44 million Americans are functionally illiterate in the English language, and another 50 million have marginal literacy skills.[3] This means that almost half of the adult population experiences difficulty in using reading, speaking, writing, and computational skills in everyday life situations. "Health literacy," the constellation of skills required to function in the health care environment, may be significantly worse than functional literacy because of the unfamiliar context and vocabulary of the health care system.[3] The National Assessment of Adult Literacy, planned for the year 2003, will include a special health literacy component that will provide national statistics on health literacy.

Contributing to poor understanding of the importance of literacy skills is the fact that low literacy is often underreported. The NALS reported that 66% to 75% of adults in the lowest reading level and 93% to 97% in the second-lowest reading level described themselves as being able to read or write English "well" or "very well."[2] In addition, many low-literate individuals are ashamed of their reading difficulties and try to hide the problem, even from their family. As many as 67% never tell a family member about their literacy problem.[4] Lack of recognition and denial of reading problems creates a barrier to health care. Because they are ashamed of their reading difficulties, low-literacy patients have acknowledged avoidance of medical care.[4,5] And because there are generally only moderate reading demands in everyday life, individuals may not be aware of their reading problems until a literacy-challenging event (ie, reviewing treatment options, reading a consent document, completing health assessment forms) occurs.[6]

In 1992, the National Cancer Institute established the National Work Group on Literacy and Health[7] to focus attention on communication difficulties associated with limited literacy skills. The findings of this group indicated that limited literacy has implications not just for cancer but also for all areas of health. Patients with limited literacy skills often have difficulty understanding diagnoses, discharge instructions, and treatment recommendations, partially because health education pamphlets and other written materials often require at least ninth-grade reading comprehension levels.[8-14] This puts them at risk for noncompliance with treatment regimens and for adverse reactions (eg, medication errors resulting from inability to read labels).[6,11] Low literacy has also been associated with poorer physical health[15,16] and with advanced stage prostate cancer at presentation[17]; these associations persist even after adjustment for demographic and socioeconomic indicators, including years of education. The NALS found that 75% of American adults with a physical or mental health condition scored in the 2 lowest literacy levels.[2] Low literacy may also result in higher health care costs because of ineffective health care, poorer health status, or higher hospitalization rates.[7,15]

Although the majority of adults with poor literacy are white, native-born Americans, ethnic minorities are disproportionately represented.[15,18] By using a newly developed multilingual test of functional health literacy, Williams et al[13] reported that over one third of English-speaking patients and nearly two thirds of Spanish-speaking patients at 2 public hospitals had inadequate or marginal health literacy. For example, some were unable to understand information regarding medication directions, appointments, and informed consent.

There is also increasing evidence of racial and ethnic disparities in health care quality and outcomes, although the association may be mediated by socioeconomic disparities.[19] Barriers to health care and poorer health outcomes have been consistently reported for Latino patients with poor English language skills.[20-25] Both Latinos and African Americans have been more likely to present with advanced stage breast or prostate cancer compared with whites.[26-28] Latino women have the second highest incidence of invasive cervical cancer, and they have lower rates of breast cancer, but higher mortality.[29-31] Latino and African-American respondents in the Medical Outcomes Study both reported slightly lower health status scores compared with other ethnic groups.[32]

One of the key outcome measures being evaluated in current research is self-reported health status, or health-related quality of life (HRQOL), a recognized and highly valued endpoint of care.[33-40] Administration of HRQOL questionnaires is usually performed with paper-and-pencil instruments, and multilingual versions of questionnaires are often not available. Interviewer administration is labor intensive and cost prohibitive in most health care settings. Therefore, patients with low literacy, those with certain functional limitations, or those who do not speak English are typically excluded, either explicitly or implicitly, from any outcome evaluation in a clinical trial or practice setting in which patient-reported data are collected on forms. An Institute of Medicine (IOM) panel concluded that inadequate attention has been devoted to the specific needs of cancer survivors in ethnic minority and medically underserved communities.[41]

As patient-reported outcomes continue to play a greater role in medical decision making and evaluation of the quality of health care, sensitive and efficient methods of measuring those outcomes among underserved populations must be developed and validated. Minority status, language preference, and literacy level may be critical variables in differentiating those who receive and respond well to treatment from those who do not.

These patients may experience different health outcomes because of disparities in care or barriers to care. Outcome measurement in these patients may provide new insight into disease or treatment problems that may have gone undetected simply because many studies have not been able to accommodate the special needs of such patients.

MULTIMEDIA TECHNOLOGY AND ITS APPLICATION TO HEALTH CARE

Interactive computerized outcomes assessment provides an innovative method for gathering and using self-report data.[42-44] Computer-based testing (CBT) has been successfully implemented in oncology practice,[42,45-54] in other clinical settings,[55-61] and on the Internet. Debriefing interviews have revealed a high level of acceptability, even with inexperienced and very ill patients. Findings regarding the comparability of CBT and paper-and-pencil assessment in health care settings are also encouraging.[51,53,56,58,60,61] All of these CBT applications, however, have been accessible only to respondents who can read the text on the computer screen or who have someone to read it aloud to them.

The Center on Outcomes, Research and Education Approach to Computerized Assessment in Vulnerable Populations

Recently, audio components have been added to computer-based assessment and programmed into administration software, creating an audiovisual computer-based testing (AVCBT) platform. The technology was originally developed to administer complex survey questionnaires in an audio format and record responses without an interviewer.[62,63] At the Center on Outcomes, Research and Education (CORE), we have developed an AVCBT program in English and Spanish to assess health status and preference-based HRQOL outcomes in low literacy cancer patients. Specifically, the AVCBT program allows respondents to listen with headphones to spoken questions that have been digitally recorded and stored on a computer. To answer, respondents press an answer button on the touchscreen monitor. Questions are also displayed on the computer screen, and patients may respond to the visual presentation of the question rather than waiting until the audio reading has been completed. The text for each questionnaire element (instructions, items, response categories) appears on the screen in different colors and boxes, and a small picture of an ear appears near each text element. Patients may touch the ear icons as many times as they wish to replay the sound for a particular element. Figure 1 shows a sample screen for 1 item in English and in Spanish. The simultaneous audio and visual presentation of questionnaires in AVCBT extends the eligibility of respondents, permitting those with limited or no reading abilities to participate. There is also anecdotal evidence that adding an audio component enhances the concentration of respondents with good reading skills.[64]

Concern has been expressed about the feasibility of obtaining information via computer from older and less-educated participants. And yet, the over 65 age group is the fastest-growing group of users of the Internet. In addition, Kohlmeier et al[43] have shown that elderly people are willing and able to become computer literate, as long as ergonomic adaptations are made. Elderly participants in their focus group were flattered that they were asked to use a high-tech method and expressed support for the program and concept. Low-income participants were also able to use a computer-based demonstration program on their own.[43]

Our multicenter AVCBT studies include an evaluative component to receive feedback from patients on the acceptability of
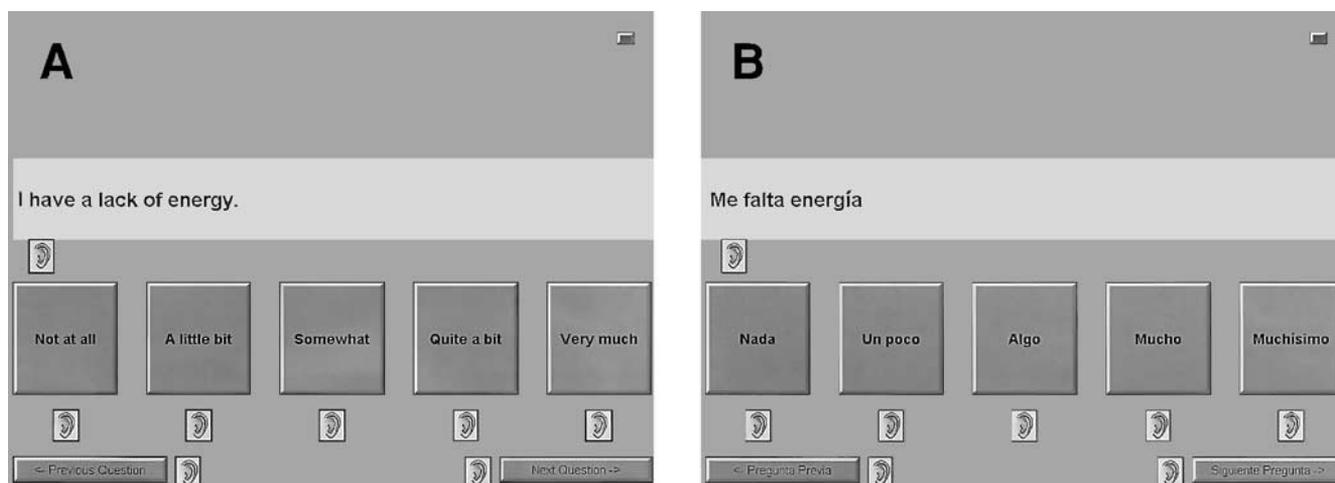
Fig 1. Examples of the AVCBT screen in (A) English and (B) Spanish.

the technology, the case of questionnaire completion, and preferences for future assessments. Responses to the evaluation to date provide convincing evidence of the feasibility of using multimedia technology with a diverse group of cancer patients. Sixty percent of the first 324 patients enrolled in the English language study are African American, two thirds are women, and the average age is 54 years (range, 20–85y); 57% had never used a computer. Table 1 shows some of the evaluation results by literacy level. Patients were classified as low literate (below seventh-grade reading comprehension) or high literate, based on the Woodcock Language Proficiency Battery passage

Table 1: Evaluation of the AVCBT

	Low Literacy* (n=170) (%)	High Literacy (n=154) (%)	All Patients (n=324) (%)
How hard was it for you to use the touchscreen?			
Very easy	40 (25.0)	85 (57.4)	125 (40.6)
Easy	116 (72.5)	60 (40.5)	176 (57.1)
Hard	4 (2.5)	1 (0.7)	5 (1.6)
Very Hard	0	2 (1.4)	2 (0.6)
Did you feel anxious or nervous while completing the touchscreen surveys?			
Yes	26 (16.1)	13 (8.8)	39 (12.6)
No	135 (83.9)	135 (91.2)	270 (87.4)
Preference			
Prefer touchscreen	101 (62.0)	103 (68.2)	204 (65.0)
Prefer interviewer	34 (20.9)	9 (6.0)	43 (13.7)
No preference	28 (17.2)	39 (25.8)	67 (21.3)
Would you be willing to do the surveys each time you visit the doctor?			
Yes	64 (39.0)	62 (41.1)	126 (40.0)
Sometimes	79 (48.2)	67 (44.4)	146 (46.3)
No	21 (12.8)	22 (14.6)	43 (13.7)

NOTE: Values are number of patients (%). Missing responses were excluded.
* Low literacy was defined as less than seventh-grade reading comprehension (see text).

comprehension subtest.[65] Nearly all patients (98%) felt the touchscreen was easy or very easy to use. Only a small proportion (13%) felt anxious or nervous while completing the touchscreen surveys; reasons included nervousness using a computer for the first time as well as anxiety related to the content of some of the questions. When asked about preferences, the majority preferred the touchscreen to having an interviewer ask the questions. It is likely that some of the high-literacy patients might have preferred paper-and-pencil forms, but we did not include this option in the evaluation question because it would not have been applicable to low-literacy patients. Most patients said they would be willing to complete an HRQOL assessment at future visits.

MEASUREMENT EQUIVALENCE

The extent to which items in a questionnaire perform similarly across different reference groups is critical when determining whether a given questionnaire can be used as an unbiased basis for comparing groups. Possible cultural differences in interpreting questions[66,67] and in response styles[68] may limit comparisons between members of different cultural groups. Without assurances that a self-report instrument is "culture-fair," detected differences caused by biased items will be incorrectly interpreted to reflect real differences between groups. Similar bias could be present when performing outcomes assessment with low-literacy patients.

Commonly used approaches to measuring equivalence in the tradition of classical test theory[69] include cross-cultural comparisons of clinical validity (eg, known-groups comparisons), multitrait scaling, exploratory and confirmatory factor analysis, regression modeling with an external criterion of bias, item-by-group interaction terms in analysis of variance, and comparison of the proportion of subjects favorably endorsing each item. Two major limitations to the classical test theory approach are that the statistics are sample dependent, and the obtained scores are test dependent.[70,71] This means that items on a questionnaire may function differently depending on the sample, and the estimate of an individual's trait level may vary depending on the particular items on the questionnaire. For example, to obtain useful and meaningful estimates of HRQOL, low-literacy patients with high (good) HRQOL should have the same probability of giving a favorable response to a particular item as high-literacy patients with high HRQOL. In addition, if multiple representative items are used

to measure a latent trait such as HRQOL, the same measure (score) should be obtained for a particular patient regardless of the items used, as long as the items really do represent the underlying construct.

Additional limitations and caveats for using traditional (classical test theory) analysis methods to detect bias across different groups of subjects have been discussed by Wright et al[72] and Liang.[73] Using an external criterion of bias in regression models (eg, comparing the extent to which HRQOL scores predict a selected outcome across cultural groups) requires the criterion itself to be unbiased. This suggests that techniques using only an internal criterion (ie, the item responses) may be more useful. Comparisons of factor configurations, which rely on correlation matrices, are confounded by differences in sample variances. Significant item-by-group interaction in analysis of variance, purported to be an indicator of bias, is affected by heterogeneity of item difficulty variances. Evaluating differences in the proportion of each group that endorses each item depends on the unlikely assumption that the groups have comparable distributions of the latent trait.

A solution to these measurement challenges is offered by an approach based on item response theory (IRT) measurement models.[70,71,74-78] These measurement models have been used extensively in educational testing and are now being implemented in health status assessment,[79,80] international cross-cultural outcomes measurement projects,[81,82] longitudinal observational studies of patients with chronic medical and psychiatric conditions,[83] and development and validation studies in the field of rehabilitation medicine.[84-86]

Several IRT-related models exist for the analysis of ordinal rating scale data, although they vary in the number of parameters that are modeled.[70,71,74-78,87] All IRT models describe the association between a respondent's level on a latent trait and the probability of a particular item response. Hays et al[79] and Cella and Chang[80] provide useful summaries of the various IRT models and some advantages of using IRT in health outcomes assessment.

IRT models also offer strategies to evaluate measurement equivalence, that is, whether there are similar probabilities of a particular item response for individuals who are equal on the latent trait. Items that appear to display differential item functioning (DIF), or measurement bias, can be identified and investigated as to their content to determine the likely source of DIF. The most important indicator of DIF is not whether items systematically differentiate relevant subgroups, but whether they do so in an unmodeled (ie, unpredicted) way.

The CORE Approach to Evaluating Measurement Equivalence

In a recent multicenter study, we used IRT models and DIF detecting procedures[70,78,88,89] to evaluate the measurement properties of the Spanish language version of the Functional Assessment of Cancer Therapy–General[90-92] (FACT-G). The FACT-G is the cornerstone of the Functional Assessment of Chronic Illness Therapy (FACIT) measurement system, which is a multi-item compilation of a generic chronic illness questionnaire (now 27 items) and many specific subscales, which reflect issues or problems associated with different diseases, including cancer, human immunodeficiency virus/acquired immune deficiency syndrome, and multiple sclerosis. There are 270 additional items in the FACIT system, and many are available in over 30 different languages.[90] The FACT-G measures physical well-being (7 items), social and family well-being (7 items), emotional well-being (6 items), and functional well-being (7 items).

Latino patients were classified as low literate (below the seventh-grade reading level) or high literate based on the Spanish version of the Woodcock Language Proficiency Battery passage comprehension subtest.[65] High-literacy patients were randomly assigned to self- or interviewer-administration of the paper-and-pencil version of the FACT-G, and low-literacy patients were assigned to interviewer administration. By using an IRT measurement model for each of the 4 FACT-G subscales, we obtained separate item location calibrations for the 3 patient groups and plotted them against each other. The item location calibration is an estimate of the item's location on the continuum of measurement, representing the probability of answering in a manner reflecting a high level of the particular trait being measured. The item location calibration plots for the 2 interviewer-administration groups are shown in figure 2. An identity line was drawn through the origin of each plot to enhance interpretability. The expectation is that the item calibrations will be close to the identity line. A t statistic was computed for each item, by using the item location calibrations (d_i) and standard errors (s_i),

$$t_i = (d_{i,\text{Low literacy}} - d_{i,\text{High literacy}})/(s_{i,\text{Low literacy}}^2 + s_{i,\text{High literacy}}^2)^{1/2}.$$

An absolute t value greater than 1.96 was considered statistically significant and was represented by a square in the figures. We found that item locations were highly similar across literacy levels. There were a few items that appeared to function differently between the 2 groups. Work is ongoing to distinguish between statistically significant and clinically meaningful differences. We will apply these same techniques to evaluate equivalence of health status measures across low- and high-literacy patients enrolled in our AVCBT projects.

DISCUSSION

Research on literacy, language, ethnicity, and health outcomes has been limited because of (1) difficulties in determining causal pathways of direct and indirect effects linking outcomes with patient factors; (2) difficulties in separating literacy from socioeconomic, demographic, and cultural factors; and (3) exclusion of low-literacy and non-English-speaking patients when self-report data are collected on forms. Computerized assessment of HRQOL outcomes, using multimedia technology, has the potential to provide a simple and efficient administration method that can be easily incorporated into a busy clinical setting with minimal burden on staff and patients. Presentation of a question on a computer screen, accompanied by a recorded reading of the same question (AVCBT) can allow people with limited or even no reading ability to participate in the assessment. By offering the same assessment method to both low- and high-literacy patients, those with lower literacy skills may avoid being labeled as requiring extra staff effort to obtain accurate and complete data. In light of research findings that many low-literacy patients are ashamed of their reading deficiencies and would prefer not to disclose them,[4,6] a standardized approach across levels of literacy has the added advantage of reducing the low-literacy stigma. Additional ergonomic adaptations will need to be developed for patients with vision or hearing problems. Interactive health communications have the potential to improve outcomes, decrease costs, and enhance satisfaction but have rarely been implemented in chronic illness treatment.[93-95]

As reforms in health care are implemented, more patients who previously received care only in the public sector will receive care in the private sector, which means that more providers will become involved in treating low-literacy patients,[9] creating unprecedented educational challenges to many providers. And yet, we know very little about the quality of life
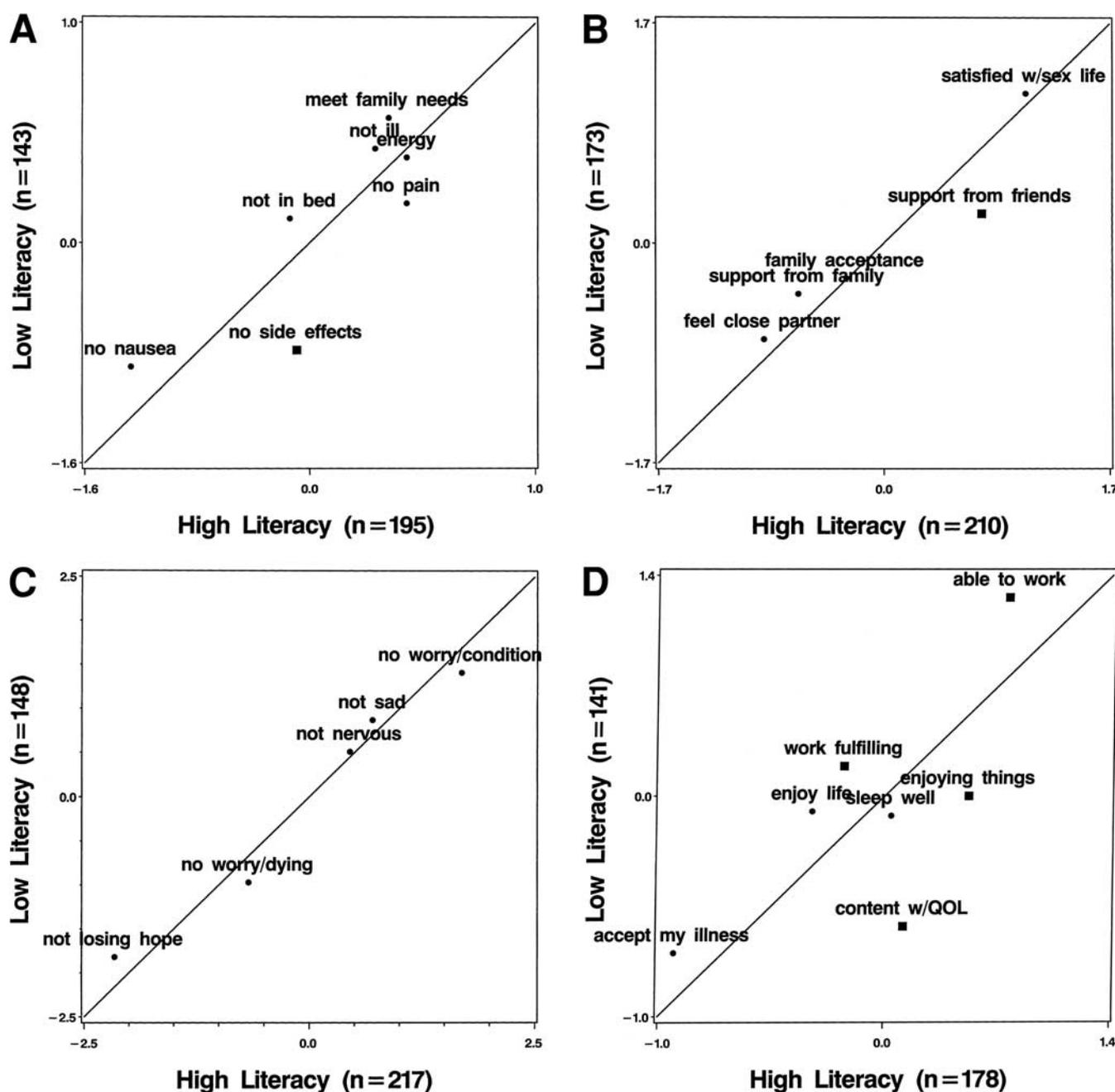
Fig 2. FACT-G item locations, low-literacy versus high-literacy patients, in logits (log-odds units). The diagonal line is the identity line. A square represents item locations that were significantly different from one another ($t > 1.96$). (A) FACT-G physical well-being, (B) social/family well-being, (C) emotional well-being, and (D) functional well-being.

of these patients because they often avoid, refuse, or are excluded from quality of life assessment. The American Medical Association's Ad Hoc Committee on Health Literacy[3] recently outlined 4 key research needs: (1) health literacy screening, (2) improving communication with low-literacy patients, (3) medical outcomes and economic costs of poor literacy, and (4) understanding the causal pathway of how health literacy influences health. In addition, an IOM panel[41] concluded that inadequate attention has been devoted to the specific needs of cancer survivors in ethnic minority and medically underserved communities.

RECOMMENDATIONS

We offer the following recommendations for future clinical practice and research.

Recommendations for Clinical Practice

1. Implement routine literacy screening when complicated medical regimens will be prescribed, provide appropriate follow-up to ensure comprehension and compliance, and include staff sensitivity training.[3,4,6,93]
2. Simplify written material.

3. Provide multilingual health materials and assessment tools.
4. Use culturally appropriate communication interventions.[93,96,97]
5. Use multimedia approaches such as pictures, graphics, and computers.

Research Needs

1. Determine which strategies are effective for communicating health information to people with chronic illnesses, regardless of their reading ability, language, or culture.
2. Evaluate whether innovative communication strategies or simplified materials improve compliance or outcomes.[3,93]
3. Evaluate the association between literacy and health care costs, use of health services, and disease outcomes.
4. Better understand the causal pathways by which low literacy is associated with poorer health, higher health care use, and higher costs.[3]

CONCLUSION

Implementation of these practice recommendations will be required to meet the identified research needs; similarly, research findings will inform the development of additional practice improvements. The strategies are intertwined, and both will be needed to meet the needs of medically underserved and vulnerable populations.

References

1. Quality First: better health care for all Americans. Final Report of the President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry. Washington (DC): US Government Printing Office; 1998.
2. Kirsch IS, Jungblut A, Jenkins L, Kolstad A. Adult literacy in America: a first look at the result of the National Adult Literacy Survey. Washington (DC): National Center for Education Statistics, US Department of Education; 1993.
3. Health literacy: report of the Council on Scientific Affairs. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. JAMA 1999;281:552-7.
4. Parikh NS, Parker RM, Nurss JR, Baker DW, Williams MV. Shame and health literacy: the unspoken connection. Patient Educ Couns 1996;27:3-39.
5. Baker DW, Parker RM, Williams MV, et al. The health care experience of patients with low literacy [abstract]. Clin Res 1993; 41:584A.
6. Baker DW, Parker RM, Williams MV, et al. The health care experience of patients with low literacy. Arch Fam Med 1996;5: 329-34.
7. Communicating with patients who have limited literacy skills. Report of the National Work Group on literacy and health. J Fam Pract 1998;46:168-76.
8. Dollahite J, Thompson C, McNew R. Readability of printed sources of diet and health information. Patient Educ Couns 1996; 27:123-34.
9. Glazer HR, Kirk LM, Bosler FE. Patient education pamphlets about prevention, detection, and treatment of breast cancer for low literacy women. Patient Educ Couns 1996;27:185-9.
10. Jolly BT, Scott JL, Feied CF, Sanford SM. Functional illiteracy among emergency department patients: a preliminary study. Ann Emerg Med 1993;22:573-8.
11. Kalichman SC, Ramachandran B, Catz S. Adherence to combination antiretroviral therapies in HIV patients of low health literacy. J Gen Intern Med 1999;5:267-73.
12. Spandorfer JM, Karras DJ, Hughes LA, Caputo C. Comprehension of discharge instructions by patients in an urban emergency department. Ann Emerg Med 1995;25:71-4.
13. Williams MV, Parker RM, Baker DW, et al. Inadequate functional health literacy among patients at two public hospitals. JAMA 1995;274:1677-82.
14. Williams MV, Baker DW, Parker RM, Nurss JR. Relationship of functional health literacy to patients' knowledge of their chronic disease. A study of patients with hypertension and diabetes. Arch Intern Med 1998;158:166-72.
15. Baker DW, Parker RM, Williams MV, Clark WS, Nurss J. The relationship of patient reading ability to self-reported health and use of health services. Am J Public Health 1997;86:1027-30.
16. Weiss BD, Hart G, McGee DL, D'Estelle S. Health status of illiterate adults: relation between literacy and health status among persons with low literacy skills. J Am Board Fam Pract 1992;5: 257-64.
17. Bennett CL, Ferreira MR, Davis TC, et al. Relation between literacy, race, and stage of presentation among low-income patients with prostate cancer. J Clin Oncol 1998;16:3101-4.
18. Weiss BD, Reed RL, Kligman EW. Literacy skills and communication methods of low-income older persons. Patient Educ Couns 1995;25:109-19.
19. Fiscella K, Franks P, Gold MR, Clancy CM. Inequality in quality: addressing socioeconomic, racial, and ethnic disparities in health care. JAMA 2000;283:2579-84. Comments in: JAMA 2000;284: 2053, 2001;285:883.
20. Derose KP, Baker DW. Limited English proficiency and Latinos' use of physician services. Med Care Res Rev 2000;57:76-91.
21. Harlan LC, Bernstein AB, Kessler LG. Cervical cancer screening: who is not screened and why? Am J Public Health 1991;81:885-90.
22. Kirkman-Liff B, Mondragon D. Language of interview: relevance for research of southwest Hispanics. Am J Public Health 1991; 81:1399-404.
23. Polednak AP. Reported pap test used by Hispanic women in Connecticut and Long Island. Conn Med 1996;60:3-8.
24. Schur CL, Albers LA. Language, sociodemographics, and health care use of Hispanic adults. J Health Care Poor Underserved 1996;7:140-58.
25. Solis JM, Marks G, Garcia M, Shelton D. Acculturation, access to care, and use of preventive services by Hispanics: findings from HHANES 1982-84. Am J Public Health 1990;80 Suppl:11-9.
26. Hoffman RM, Gilliland FD, Eley JW, et al. Racial and ethnic differences in advanced-stage prostate cancer: the Prostate Cancer Outcomes Study. J Natl Cancer Inst 2001;93:388-95.
27. Lannin DR, Mathews HF, Mitchell J, Swanson MS, Swanson FH, Edwards MS. Influence of socioeconomic and cultural factors on racial differences in late-stage presentation of breast cancer. JAMA 1998;279:1801-7.
28. Thompson I, Tangen C, Tolcher A, Crawford E, Eisenberger M, Moinpour C. Association of African-American ethnic background with survival in men with metastatic prostate cancer. J Natl Cancer Inst 2001;93:219-25.
29. American Cancer Society. Facts and figures, 1997. New York: American Cancer Society; 1997.
30. Eidson M, Becker TM, Wiggins CL, Key CR, Samet JM. Breast cancer among Hispanics, American Indians and non-Hispanic whites in New Mexico. Int J Epidemiol 1994;23:231-7.
31. Ries LA, Kosary CL, Hankey BF, Miller BA, Edwards BK, editors. SEER cancer statistics review, 1973-1995. Bethesda (MD): National Cancer Institute; 2000.
32. Cunningham WE, Hays RD, Burton TM, Kington RS. Health status measurement performance and health status differences by age, ethnicity, and gender: assessment in the medical outcomes study. J Health Care Poor Underserved 2000;11:58-76.
33. Cella D. Quality of life: the concept. J Palliat Care 1992;8:8-13.
34. Cella D. Quality of life. In: Holland JC, editor. Psychooncology. New York: Oxford Univ Pr; 1998. p 1135-46.
35. Cella D, Tulsky DS. Quality of life in cancer. Definition, purpose, and method of measurement. Cancer Invest 1993;11:327-36.
36. Ellwood PM. Outcomes management: a technology of patient experience. N Engl J Med 1988;18:1549-56.
37. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. Ann Intern Med 1993;118:622-9.
38. Schroeder SA. Outcome assessment 70 years later: are we ready? New Engl J Med 1987;316:160-2.
39. Tarlov AR. The increasing supply of physicians, the changing structure of the health-services system, and the future practice of medicine. New Engl J Med 1983;398:1235-44.

40. Ware JE Jr. The status of health assessment 1994. Ann Rev Public Health 1995;16:327-54.
41. Haynes MA, Smedley BD, editors. The unequal burden of cancer: an assessment of NIH Research and Programs for ethnic minorities and the medically underserved (summary). Washington, (DC): Natl Acad Pr; 1999.
42. Cella DF, Lloyd SR. Data collection strategies for patient-reported information. Qual Manag Health Care 1994;2(4):28-35.
43. Kohlmeier L, Mendez M, McDuffie J, Miller M. Computer-assisted self-interviewing: a multimedia approach to dietary assessment. Am J Clin Nutr 1997;65(4 Suppl):1275S-1281S.
44. Hahn EA, Cella D, Odom L, AVCBT Investigators. The talking touchscreen: a new method to assess QOL in people with very limited reading ability. Paper presented at: 9th Annual Meeting of the International Society for Quality of Life Research; 2002 Nov 1; Orlando (FL).
45. Buxton J, White M, Osoba D. Patients' experiences using a computerized program with a touch-sensitive video monitor for the assessment of health-related quality of life. Qual Life Res 1998;6:513-9.
46. Cella D, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy (FACT) scale: development and validation of the general measure. J Clin Oncol 1993;11:570-9.
47. Harpole LH, Nolte KC, D'Amico TA, Crawford JN, Harpole DH. Assessing lung cancer patient outcomes at point of service. In: Proceedings of the American Society of Clinical Oncology's 37th Annual Meeting; 2001 May 12-15; San Francisco (CA). No. 2817.
48. McLachlan S, Allenby A, Matthews J, et al. A randomised trial of the collection of quality of life (QL) and cancer needs information via a computer touch screen by ambulatory patients with cancer. In: Proceedings of the American Society of Clinical Oncology's 37th Annual Meeting; 2001 May 12-15; San Francisco (CA). No. 987.
49. Newell S, Girgis A, Sanson-Fisher RW, et al. Are touch-screen computer surveys acceptable to medical oncology patients? J Psychosocial Oncol 1997;15:37-46.
50. Sigle J, Porzolt F. Practical aspects of quality of life measurement, design and feasibility: study of the quality of life recorder and the standardized measurement of quality of life in an outpatient clinic. Cancer Treat Rev 1996;22 Suppl A:75-89.
51. Taenzer PA, Speca M, Atkinson MJ, et al. Computerized quality-of-life screening in an oncology clinic. Cancer Pract 1997;5:68-175.
52. Taenzer P, Bultz BD, Carlson LE, et al. Impact of computerized quality of life screening on physician behaviour and patient satisfaction in lung cancer outpatients. Psychooncology 2000;9:203-13.
53. Velikova G, Wright EP, Smith AB, et al. Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. J Clin Oncol 1999;17:998-1007.
54. Wright EP, Smith AB, Velikova G, et al. Automated screening and recording of quality of life data: a comparison of data capture methods [abstract]. Psychooncology 1998;7:67.
55. Burke JD, Burke KC, Baker JH, et al. Test-retest reliability in psychiatric patients of the SF-36 Health Survey. Int J Methods Psychiatr Res 1995;5:189-94.
56. Crawley JA, Kleinman L, Dominitz J. User preferences for computer administration of quality of life instruments. Drug Information J 2000;34:137-44.
57. Lenert LA, Hornberger JC. Computer-assisted quality of life assessment for clinical trials. Proc AMIA Annu Fall Symp 1996;992-6.
58. Pouwer F, Snoek FJ, van der Ploeg HM, Heine RJ, Brand AN. A comparison of the standard and the computerized versions of the Well-being Questionnaire (WBQ) and the Diabetes Treatment Satisfaction Questionnaire (DTSQ). Qual Life Res 1998;7:33-8.
59. Robinson R, West R. A comparison of computer and questionnaire methods of history-taking in a genito-urinary clinic. Psychol Health 1992;6:77-84.
60. Roizen MF, Coalson D, Hayward RS, et al. Can patients use an automated questionnaire to define their current health status? Med Care 1992;30(5 Suppl):74-84.
61. Swanston M, Abraham C, Macrae WA, et al. Pain assessment with interactive computer animation. Pain 1993;53:347-51.
62. O'Reilly JM, Hubbard M, Lessler J, et al. Audio and video computer-assisted self interviewing: preliminary tests of new technologies for data collection. J Off Stat 1994;10:197-214.
63. Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use and violence: increased reporting with computer survey technology. Science 1998;280(5365):867-73.
64. Lessler JT, O'Reilly JM. Mode of interview and reporting of sensitive issues: design and implementation of audio computer-assisted self-interviewing. In: Harrison L, Hughes A, editors. The validity of self-reported drug use: improving the accuracy of survey estimates. Rockville (MD): US Dept of Health and Human Services, National Institutes of Health, National Institute of Drug Abuse; 1997. p 366-82. NIDA Res Monogr 167.
65. Woodcock RW. Examiner's manual: Woodcock Language Proficiency Battery-Revised. Allen (TX): DLM Publisher; 1991.
66. Angel R, Thoits P. The impact of culture on the cognitive structure of illness. Culture Med Psychiatry 1987;11:23-52.
67. Johnson TP, O'Rourke D, Chavez N, et al. Cultural variations in the interpretation of health questions. In: Warnecke RB, editor. Health survey research methods conference proceedings. Hyattsville (MD): National Center for Health Statistics; 1996. p 57-62.
68. Hayes RP, Baker DW. Methodological problems in comparing English-speaking and Spanish-speaking patients' satisfaction with interpersonal aspects of care. Med Care 1998;36:230-6.
69. Lord FM, Novick NR. Statistical theories of mental test scores. Reading (MA): Addison-Wesley; 1968.
70. Wright BD, Masters GN. Rating scale analysis: Rasch measurement. Chicago: MESA Pr; 1982.
71. Hambleton RK, Swaminathan H, Rogers HJ. Fundamentals of item response theory. Newbury Park (CA): Sage; 1991.
72. Wright BD, Mead R, Draba R. Detecting and correcting test item bias with a logistic response model. Chicago: MESA Psychometric Lab, Univ Chicago; 1976.
73. Liang J. Assessing cross-cultural comparability in mental health among older adults. J Ment Health Aging 2001;7:21-30.
74. Birnbaum A. Some latent trait models and their use in inferring an examinee's ability. In: Lord FM, Novick MR, editors. Statistical theories of mental test scores. Reading (MA): Addison-Wesley; 1968. p 397-479.
75. Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Denmarks Paedogogiske Institut; 1960.
76. Samejima, F. Estimation of latent ability using a response pattern of graded scores. Iowa City (IA): Psychometric Society; 1969. Psychometric Monogr No. 17.
77. van der Linden WJ, Hambleton RK. Handbook of modern item response theory. New York: Springer-Verlag; 1997.
78. Wright BD, Stone MH. Best test design: Rasch measurement. Chicago: MESA Pr; 1979.
79. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. Med Care 2000;38(9 Suppl):II28-42.
80. Cella D, Chang CH. A discussion of item response theory and its applications in health status assessment. Med Care 2000;38(9 Suppl):II66-72.
81. Raczek AE, Ware JE, Bjorner JB, et al. Comparison of Rasch and summated rating scales constructed from SF-36 physical functioning items in seven countries; results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998:51;1203-14.
82. Tennant A. Rasch-based medical research at University of Leeds, England. Rasch Meas Trans 1998;11:601.
83. Haley SM, McHorney CA, Ware JE Jr. Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the RASCH item scale. J Clin Epidemiol 1994;47:671-84.
84. Heinemann AW, Hamilton BB, Granger CV, Wright BD, Linacre JM, Mamott BD. Rating scale analyses of functional assessment

measures. Final report to the National Institute on Disability and Rehabilitation Research. Chicago: Rehabilitation Inst Chicago; 1991.

85. Ludlow LH, Haley SM, Gans BM. A hierarchical model of functional performance in rehabilitation medicine: the Tufts Assessment of Motor Performance. Eval Health Professions 1992; 15:59-74.

86. Silverstein B, Kilgore KM, Fisher W. Implementing patient tracking systems and using functional assessment scales. Wheaton (IL): Marianjoy Rehabilitation Center; 1989.

87. Andrich D. A rating formulation for ordered response categories. Psychometrika 1978;43:561-73.

88. Smith RM. A comparison of the Rasch separate calibration and between fit methods of detecting item bias. Educ Psychol Meas 1996;56:403-18.

89. Cella DF, Lloyd SR, Wright BD. Cross-cultural instrument equating: current research and future directions. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd ed. Philadelphia: Lippincott-Raven; 1996. p 707-15.

90. Cella DF. Manual of the Functional Assessment of Chronic Illness Therapy (FACIT Scales)-version 4. Evanston (IL): Center on Outcomes, Research and Education (CORE), Evanston Northwestern Healthcare & Northwestern Univ; 1997.

91. Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy (FACT) scale: development and validation of the general measure. J Clin Oncol 1993;11:570-9.

92. Cella DF, Hernandez L, Bonomi AE, et al. Spanish language translation and initial validation of the functional assessment of cancer therapy quality-of-life instrument. Med Care 1998;36: 1407-18.

93. Baker DW. Reading between the lines: deciphering the connections between literacy and health. J Gen Intern Med 1999;14: 315-7.

94. Eng TR, Gustafson DH, Henderson J, et al. Introduction to evaluation of interactive health communication and applications. Science Panel on Interactive Communication and Health. Am J Prevent Med 1999;16:10-5.

95. Houts PS, Bachrach R, Witmer JT, Tringali CA, Bucher JA, Localio RA. Using pictographs to enhance recall of spoken medical instructions. Patient Educ Couns 1998;35:83-8.

96. Hodge FS, Toms FD, Guillermo T. Achieving cultural competency and responsive health care delivery. Cancer Suppl 1998;83: 1714-6.

97. Science Panel on Interactive Communication and Health. Wired for health and well being: the emergence of interactive health communication. Washington (DC): US Department of Health and Human Services; 1999.