Validity of Single-Item Screening for Limited Health Literacy in English and Spanish Speakers

Wendy Pechero Bishop, MS, Simon J. Craddock Lee, PhD, MPH, Celette Sugg Skinner, PhD, Tiffany M. Jones, DrPH, Katharine McCallister, BA, and Jasmin A. Tiro, PhD

Objectives. To evaluate 3 single-item screening measures for limited health literacy in a community-based population of English and Spanish speakers.

Methods. We recruited 324 English and 314 Spanish speakers from a community research registry in Dallas, Texas, enrolled between 2009 and 2012. We used 3 screening measures: (1) How would you rate your ability to read?; (2) How confident are you filling out medical forms by yourself?; and (3) How often do you have someone help you read hospital materials? In analyses stratified by language, we used area under the receiver operating characteristic (AUROC) curves to compare each item with the validated 40-item Short Test of Functional Health Literacy in Adults.

Results. For English speakers, no difference was seen among the items. For Spanish speakers, "ability to read" identified inadequate literacy better than "help reading hospital materials" (AUROC curve = 0.76 vs 0.65; $P = .019$).

Conclusions. The "ability to read" item performed the best, supporting use as a screening tool in safety-net systems caring for diverse populations. Future studies should investigate how to implement brief measures in safety-net settings and whether highlighting health literacy level influences providers' communication practices and patient outcomes. (*Am J Public Health.* 2016;106:889–892. doi:10.2105/AJPH.2016.303092)

Health literacy—ability to obtain, process, and understand basic health information and services needed to make health decisions—is a key health determinant, particularly for Hispanic immigrants.[1,2] Recent reforms following the Affordable Care Act ask health care systems to identify low health literacy patients, provide special assistance, and incorporate health literacy into quality metrics.[3] Administration of validated measures such as the Short Test of Functional Health Literacy in Adults (STOFHLA) takes 3 to 8 minutes.[4] Brief single-item measures that indirectly assess literacy and are documented in electronic health records are needed. However, these measures have been studied mostly among English speakers[5–7] who are chronically ill.[8–10] Safety-net systems are willing to conduct literacy screening[1] but need valid measures for their diverse population, particularly Spanish-speaking patients (40% of US Hispanic individuals are foreign-born, and fewer than 25% of them report speaking English very well[11]).

We evaluated 3 single items against the STOFHLA, a well-accepted, commonly used measure, in a community-based population of English and Spanish speakers.

METHODS

We randomly selected study participants (n = 638) in 2011 to 2012 from our community research registry of individuals who joined between 2009 and 2012.[12,13] Our registry enrolls Dallas County, Texas, community members by (1) inviting adults attending local health events or waiting in ambulatory clinics of the Dallas County's safety-net system[14] or (2) enabling current registry members to refer their friends or family. Registry members aged 18 to 70 years received an invitation letter (with a toll-free number to opt out) requesting help to identify strategies that improve communication between patients and providers. Bilingual research assistants called potential participants 1 week later to ascertain interest, assess eligibility, and schedule an in-person appointment. To be eligible, individuals had to report ability to read English or Spanish; if bilingual, they were asked to complete study procedures in the language they preferred to use with their provider. At the appointment, research assistants used a script to obtain consent, administer a paper version of the STOFHLA, and verbally ask the single-item measures; 49.2% of the participants completed study procedures in Spanish and 50.8% in English.

STOFHLA, a 40-item scale with validity established in Spanish and English,[15] assesses reading comprehension and numerical ability. We used recommended cutpoints to dichotomize scores (inadequate vs marginal or adequate).[16] The 3 single items described in Table 1 used 5-point responses: (1) How would you rate your ability to read?; (2) How confident are you filling out medical forms by yourself?; and (3) How often do you have someone help you read hospital materials?

ABOUT THE AUTHORS

Wendy Pechero Bishop, Simon J. Craddock Lee, Celette Sugg Skinner, and Jasmin A. Tiro are with Department of Clinical Sciences, UT Southwestern Medical Center, Dallas, TX, and Harold C. Simmons Comprehensive Cancer Center, Dallas. Tiffany M. Jones is with the School of Public Health, University of Texas Health Science Center at Houston. Katharine McCallister is with Department of Clinical Sciences, UT Southwestern Medical Center.

Correspondence should be sent to Wendy Pechero Bishop, MS, Department of Clinical Sciences, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390-9066 (e-mail: wendy.bishop@utsouthwestern.edu). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.

This brief was accepted January 16, 2016.

doi: 10.2105/AJPH.2016.303092

TABLE 1—Sociodemographic Characteristics and Health Literacy Items Among English- and Spanish-Speaking Participants Enrolled in a Community Registry: Dallas County, Texas, 2011–2012

	English (n = 324), No. (%)	Spanish (n = 314), No. (%)
Sociodemographic Characteristics[a]		
Age, y[a]		
18–34	110 (34.0)	78 (24.8)
35–49	110 (34.0)	170 (54.1)
50–70	104 (32.1)	66 (21.0)
Sex[a]		
Male	114 (35.2)	91 (29.0)
Female	210 (64.8)	223 (71.0)
Race[a]		
Black/American Indian/Alaska Native/Asian	171 (52.8)	4 (1.3)
White[b]	153 (47.2)	310 (98.7)
Marital status[a]		
Married or living with partner	128 (39.5)	214 (68.2)
Single/divorced/widowed/separated/other	196 (60.5)	100 (31.9)
Education[a,c]		
Grade school	4 (1.2)	72 (23.0)
Some high school	31 (9.6)	95 (30.4)
High school diploma/GED/technical school	102 (31.5)	101 (32.3)
Some college/graduated college	187 (57.7)	45 (14.4)
Health Literacy Items		
STOFHLA[d]		
Inadequate (score = 0–16)	9 (2.8)	58 (18.5)
Marginal or adequate (score = 17–36)	315 (97.2)	256 (81.5)
How would you rate your ability to read?		
Very poor (1)	1 (0.3)	2 (0.6)
Poor (2)	8 (2.5)	3 (1.0)
OK (3)	42 (13.0)	76 (24.2)
Good (4)	81 (25.0)	151 (48.1)
Very good (5)	192 (59.3)	82 (26.1)
How often do you have someone help you read hospital materials?		
Always (1)	8 (2.5)	6 (1.9)
Often (2)	16 (4.9)	2 (0.6)
Sometimes (3)	39 (12.0)	64 (20.4)
Occasionally (4)	66 (20.4)	63 (20.1)
Never (5)	195 (60.2)	179 (57.0)
How confident are you filling out medical forms by yourself?		
Not at all (1)	0 (0.0)	1 (0.3)
A little bit (2)	12 (3.7)	13 (4.1)
Somewhat (3)	44 (13.6)	31 (9.9)
Quite a bit (4)	57 (17.6)	164 (52.2)
Very (5)	211 (65.1)	105 (33.4)

Note. GED = general equivalency diploma; STOFHLA = Short Test of Functional Health Literacy in America. The sample size was n = 638.

[a]Chi-square analysis for each sociodemographic characteristic revealed significant differences between English and Spanish speakers ($P < .001$).

[b]Includes "don't know" or "did not want to reply."

[c]One Spanish-speaking participant was excluded because of missing data.

[d]STOFHLA: inadequate (score = 0–16) vs marginal or adequate (score = 17–36).

In analyses stratified by language, we calculated area under the receiver operating characteristic (AUROC) curves comparing each item with dichotomized STOFHLA scores. This comparison used a nonparametric approach[17,18] based on generalized U-statistics theory following a χ^2 distribution. Past literature examining single-item measures indicated that AUROC curve values greater than 0.7 are justified for use in health care settings[5,9,19,20]; therefore, we used this value to compare performance separately for English- and Spanish-speaking samples. We calculated sensitivity and specificity with cutpoints suggested by the AUROC curves. Analyses were conducted in SAS version 9.3 (SAS Institute, Cary, NC).

RESULTS

Compared with English speakers, Spanish speakers were more likely to be White, middle-aged, married or living with a partner, and less educated (Table 1). More Spanish than English speakers received an inadequate STOFHLA score (18.5% vs 2.8%). More Spanish than English speakers rated "ability to read" as no better than "OK" (25.8% and 15.8%). "Help reading hospital materials" followed a similar pattern—20.4% of Spanish and 12.0% of English speakers reported "sometimes" having help. For "confidence filling out medical forms," 85.6% and 82.7% of Spanish and English speakers, respectively, were "quite a bit or very confident."

AUROC curve estimates for English and Spanish speakers had similar ranges (Table 2). For English speakers, there were no significant differences among the 3 questions ($P = .39$). For Spanish speakers, "ability to read" identified inadequate literacy better than did "help reading hospital materials" (AUROC curve = 0.76 vs 0.65; $P = .019$). Overall, 95% confidence intervals were wider (indicating less precision) for English speakers. Table 2 reports sensitivity and specificity based on optimal cutpoints suggested by the AUROC curves.

DISCUSSION

Single-item health literacy screening tools have advantages over longer, time-consuming

TABLE 2—Area Under the Receiver Operating Characteristic (AUROC) Curves, Sensitivity, Specificity, and 95% Confidence Intervals (CIs) Comparing Each Health Literacy Screening Item With the STOFHLA, Stratified by Language: Dallas County, TX, 2011–2012

	English (n = 324)			Spanish (n = 314)		
	AUROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
How would you rate your ability to read? (very poor/poor/OK/good vs very good)	0.73 (0.56, 0.91)	0.78 (0.40, 0.97)	0.60 (0.55, 0.66)	0.76 (0.70, 0.83)	0.93 (0.83, 0.98)	0.30 (0.25, 0.37)
How confident are you filling out medical forms by yourself? (not at all/a little bit/somewhat/quite a bit vs very)	0.77 (0.60, 0.94)	0.78 (0.40, 0.97)	0.66 (0.61, 0.72)	0.69 (0.62, 0.76)	0.84 (0.73, 0.93)	0.38 (0.32, 0.44)
How often do you have someone help you read hospital materials? (always/often/sometimes/occasionally vs never)	0.66 (0.47, 0.84)	0.67 (0.30, 0.93)	0.61 (0.55, 0.66)	0.65 (0.57, 0.73)	0.59 (0.45, 0.71)	0.61 (0.54, 0.67)

Note. The sample size was n = 638. AUROC curves compare each item with STOFHLA scores dichotomized as inadequate vs marginal or adequate. STOFHLA (Short Test of Functional Health Literacy in America): inadequate (score = 0–16) vs marginal or adequate (score = 17–36).

standard measures. Administration is quick and can be performed by any health care team member. Results are easy to interpret and document in the electronic health records and can guide providers' communication practices.

Our findings from a community sample indicated support for the "ability to read" item. It performed the best at distinguishing inadequate from marginal or adequate health literacy (AUROC curve = 0.73 and 0.76 for English and Spanish speakers, respectively). Although our findings for English speakers were consistent with those of past studies,[5,9,10,19] the same was not true for our Spanish-speaking sample. Specifically, our AUROC curve estimates for "confidence filling out medical forms" and "help reading hospital materials" were lower than those previously reported[9,20] and slightly lower than the 0.7 threshold recommended by the literature.[5,8,9,19,20] Differences in AUROC curve estimates between our study and previous studies may stem from sample characteristics and STOFHLA score distributions, highlighting the importance of replication studies that build evidence for validity in a variety of populations. Previous studies examined item performance with chronically ill populations who have more frequent contact with health care systems and more opportunity to practice their health literacy skills. In addition, the Sarkar et al.[9] sample had a wider distribution of STOFHLA scores.

Our study had some limitations. One limitation was that few participants gave responses at the lower threshold of the 3 questions. Eligibility criteria requiring ability to read in English or Spanish (to complete STOFHLA) may have excluded individuals with the lowest health literacy levels.[4] Second, the STOFHLA is well accepted but is not a gold standard; thus, the degree to which it is an imperfect reference standard potentially introduced inaccuracy in the AUROC analysis.[21] We did not measure social desirability, so we cannot determine the extent to which that type of bias affected participants' survey responses.[22] Finally, researchers acknowledge that health literacy is a multidimensional construct, but there is little agreement about what dimensions must be assessed. The single items we examined focus on reading comprehension.[4] These measures do not assess verbal communication, ability to navigate the health care system, or health care decision-making. It is unclear whether awareness of these skills would improve the delivery of health care.

This study highlights the importance of replication studies in building evidence for validity in various populations. Our study contributes data on the health literacy of Dallas County community members, of whom 39% are Hispanic, 35% speak Spanish at home, and 23% are uninsured.[23,24] Given our low sensitivity and specificity estimates, alternative screening items should be evaluated. Future studies also should investigate how to implement brief measures in safety-net settings[22] and whether highlighting health literacy level influences providers' communication practices and, in turn, patient outcomes. AJPH

CONTRIBUTORS

W. Pechero Bishop, C. Sugg Skinner, and J. A. Tiro conceptualized and designed the study. W. Pechero Bishop, K. McCallister, and J. A. Tiro collected the data. W. Pechero Bishop, T. M. Jones, and J. A. Tiro analyzed and interpreted the data. All authors drafted the brief and provided critical intellectual feedback on writing.

ACKNOWLEDGMENTS

This work was conducted with support from the National Institutes of Health (NCATS UL1TR001105) to the UT Southwestern Center for Translational Medicine, the National Cancer Institute (5 P30 CA142543-02) to the Harold C. Simmons Comprehensive Cancer Center, and the Agency for Healthcare Research and Quality (HS022418) to the UT Southwestern Center for Patient-Centered Outcomes Research.

The results of this study were presented at the 140th American Public Health Association Conference; October 29, 2012; San Francisco, CA.

The authors thank staff members Maria Funes, Adam Loewen, Saddyna Belmashkan, and Clare Stevens for their assistance with data collection.

HUMAN PARTICIPANT PROTECTION

This study was approved by the institutional review board at UT Southwestern Medical Center (STU 042011-042 Tiro).

REFERENCES

1. Nielsen-Bohlman L, Panzer AM, Kindig DA, eds. *Health Literacy: A Prescription to End Confusion*. Washington, DC: National Academies Press; 2004.

2. Kutner M, Greenburg E, Jin Y, Paulsen C. *The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy (NCES 2006-483)*. Washington, DC: US Department of Education, National Center for Education Statistics; 2006.

3. Koh HK, Berwick DM, Clancy CM, et al. New federal policy initiatives to boost health literacy can help the nation move beyond the cycle of costly 'crisis care.' *Health Aff (Millwood)*. 2012;31:434–443.

4. Al Sayah F, Williams B, Johnson JA. Measuring health literacy in individuals with diabetes: a systematic review and evaluation of available measures. *Health Educ Behav*. 2013;40(1):42–55.

5. Wallace LS, Rogers ES, Roskos SE, Holiday DB, Weiss BD. Brief report: screening items to identify patients with limited health literacy skills. *J Gen Intern Med*. 2006;21: 874–877.

6. Stagliano V, Wallace LS. Brief health literacy screening items predict newest vital sign scores. *J Am Board Fam Med*. 2013;26:558–565.

7. Kiechle ES, Hnat AT, Norman KE, Viera AJ, DeWalt DA, Brice JH. Comparison of brief health literacy screens in the emergency department. *J Health Commun*. 2015;20: 539–545.

8. Morris NS, MacLean CD, Chew LD, Littenberg B. The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability. *BMC Fam Pract*. 2006;7:21.

9. Sarkar U, Schillinger D, López A, Sudore R. Validation of self-reported health literacy questions among diverse English and Spanish-speaking populations. *J Gen Intern Med*. 2011;26(3):265–271.

10. Jeppesen KM, Coyle JD, Miser WF. Screening questions to predict limited health literacy: a cross-sectional study of patients with diabetes mellitus. *Ann Fam Med*. 2009;7:24–31.

11. Escarce JJ, Kapur K. Access to and quality of health care. In: Tienda M, Mitchell F, eds. *Hispanics and the Future of America*. Washington, DC: National Academies Press; 2006:410–446.

12. Bishop WP, Tiro JA, Lee SJ, Bruce CM, Skinner CS. Community events as viable sites for recruiting minority volunteers who agree to be contacted for future research. *Contemp Clin Trials*. 2011;32:369–371.

13. Bishop WP, Tiro JA, Sanders JM, Craddock Lee SJ, Skinner CS. Effectiveness of a community research registry to recruit minority and underserved adults for health research. *Clin Transl Sci*. 2015;8(1):82–84.

14. Pickens S, Boumbulian P, Anderson RJ, Ross S, Phillips S. Community-oriented primary care in action: a Dallas story. *Am J Public Health*. 2002;92: 1728–1732.

15. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. *J Gen Intern Med*. 1995;10:537–541.

16. Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J. Development of a brief test to measure functional health literacy. *Patient Educ Couns*. 1999;38: 33–42.

17. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29–36.

18. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–845.

19. Chew LD, Griffin JM, Partin MR, et al. Validation of screening questions for limited health literacy in a large VA outpatient population. *J Gen Intern Med*. 2008;23:561–566.

20. Garcia CH, Hanley J, Souffrant G. A single question may be useful for detecting patients with inadequate health literacy. *J Gen Intern Med*. 2008;23(9):1545.

21. Eng J. Receiver operating characteristic analysis: a primer. *Acad Radiol*. 2005;12:909–916.

22. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. *Qual Quant*. 2013; 47(4):2025–2047.

23. US Census Bureau. State and County QuickFacts, Dallas County, Texas. 2015. Available at: http://quickfacts.census.gov/qfd/states/48/48113.html. Accessed May 15, 2015.

24. US Census Bureau. Selected Characteristics of the Native and Foreign-Born Populations: 2009-2013 American Community Survey 5-Year Estimates. Available at: http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_14_5YR_S0501&prodType=table. 2015. Accessed May 15, 2015.