Validation of Self-Reported Health Literacy Questions Among Diverse English and Spanish-Speaking Populations

Urmimala Sarkar, MD MPH[1,2], Dean Schillinger, MD[1,2], Andrea López, BS[1,2], and Rebecca Sudore, MD[3,4]

[1]San Francisco Center for Vulnerable Populations, San Francisco General Hospital, University of California, San Francisco, CA, USA; [2]Department of Medicine, Division of General Internal Medicine, University of California, San Francisco, CA, USA; [3]San Francisco VA Medical Center, San Francisco, CA, USA; [4]Division of Geriatrics, San Francisco Veterans Affairs Medical Center, University of California, San Francisco, CA, USA.

BACKGROUND: Limited health literacy (HL) contributes to poor health outcomes and disparities, and direct measurement is often time-intensive. Self-reported HL questions have not been validated among Spanish-speaking and diverse English-speaking populations.

OBJECTIVE: To evaluate three self-reported questions:[1] "How confident are you filling out medical forms?";[2] "How often do you have problems learning about your medical condition because of difficulty understanding written information?"; and[3] "How often do you have someone help you read hospital materials?" Answers were based on a 5-point Likert scale.

DESIGN: This was a validation study nested within a trial of diabetes self-management support in the San Francisco Department of Public Health.

PARTICIPANTS: English and Spanish-speaking adults with type 2 diabetes receiving primary care.

METHODS: Using the Test of Functional Health Literacy in Adults (s-TOFHLA) in English and Spanish as the reference, we classified HL as inadequate, marginal, or adequate. We calculated the C-index and test characteristics of the three questions and summative scale compared to the s-TOFHLA and assessed variations in performance by language, race/ethnicity, age, and education.

KEY RESULTS: Of 296 participants, 48% were Spanish-speaking; 9% were White, non-Hispanic; 47% had inadequate HL and 12% had marginal HL. Overall, 57% reported being confident with forms "somewhat" or less. The "confident with forms" question performed best for detecting inadequate (C-index = 0.82, (0.77-0.87)) and inadequate plus marginal HL (C index = 0.81, (0.76-0.86); p<0.01 for differences from other questions), and performed comparably to the summative scale. The "confident with forms" question and scale also performed best across language, race/ethnicity, educational attainment, and age.

CONCLUSIONS: A single self-reported HL question about confidence with forms and a summative scale of three questions discriminated between Spanish and English speakers with adequate HL and those with inadequate and/or inadequate plus marginal HL. The "confident with forms" question or the summative scale may be useful for estimating HL in clinical research involving Spanish-speaking and English-speaking, chronically-ill, diverse populations.

INTRODUCTION

Health literacy has been defined as "the ability to obtain, process, or understand basic health information needed to make appropriate health care decisions"[1]. Both inadequate (i.e. very low) and marginal health literacy (HL) appear to be important factors in the causal pathway to health disparities, especially in low income patients with chronic diseases[2–5]. Given the high prevalence (46% of the US population) of inadequate (i.e. very low) plus marginal HL, often described as 'limited HL'[6] and limited literacy's association with poor health outcomes[3,7–12], there has been great interest in including HL assessments in epidemiologic and clinical research[13]. However, because standard HL measurements require face-to-face interviews[14–16], take from 3 to 20 minutes, and cannot be administered by phone, they are often not feasible in large epidemiologic and public health research.

Chew and colleagues developed three self-reported HL "screening" questions and found that a single item about "confidence with completing forms" with a response cut-point of "somewhat," may be sufficient to detect patients with inadequate HL (C-index 0.74 (0.69-0.79)), sensitivity, 0.60; specificity, 0.82), but the items did not perform as well in patients with inadequate plus marginal HL (C-index 0.72 (0.69-0.76)[17,18]. Chew also found that a scale combining the three questions offered no additional benefit to the one question about confidence with forms. A recent review article

endorsed the use of the 'confidence with forms' item to assess HL in clinical settings[19]. However, these self-reported items have only been validated among largely homogeneous English-speaking populations[17,18,20]. The performance of the self-reported HL questions within Spanish-speaking and ethnically diverse patient subgroups has not been assessed[19].

It is important to validate these three self-reported HL items both individually and as a scale among Spanish speakers, patients with low-income, and minorities because the prevalence of limited HL is highest among these groups[6,21]. HL and limited English proficiency have a complex relationship, adding to the importance of measuring HL in languages other than English[22]. However, Spanish HL assessment currently requires face-to-face, multi-item, interviewer-administered assessments[23]. Therefore, we examined the performance of three self-reported HL questions individually and as a summative scale among English and Spanish-speaking, diverse, low-income, populations with type 2 diabetes. We further explored whether the self-reported questions performed equally well across language, race/ethnicity, educational attainment, age, gender, and health status subgroups.

METHODS

This validation study was nested within a trial of diabetes self-management support interventions in the San Francisco Department of Public Health (SFDPH). The methods have been previously reported[24,25]. Briefly, patients were included if they were over age 17 years, had ICD-9 codes consistent with type 2 diabetes, self-reported fluency in English and Spanish, made ≥1 primary care visit at one of four (SFDPH) clinics in the prior year, and had a hemoglobin A1c value (HbA1c) ≥8.0% at the time of recruitment. All participants provided informed consent, and the Committee on Human Research at the University of California, San Francisco approved the study protocol.

Self-Reported HL Measure

Bilingual research assistants administered the following three self-reported HL questions in person in English or Spanish:[1] How confident are you filling out medical forms by yourself? (¿Qué tan seguro(a) se siente al llenar formas usted solo(a)?) "confident with forms";[2] How often do you have problems learning about your medical condition because of difficulty understanding written information? (¿Qué tan seguido tiene problemas aprendiendo sobre su condición médica porque es difícil entender información escrita?) "problems learning"; and[3] How often do you have someone like a family member, friend, hospital or clinic worker or caregiver, help you read hospital materials? (¿Qué tan seguido tiene usted, un familiar, un amigo(a), un empleado(a) del hospital o la clínica u otra persona que le ayude a leer materiales del hospital?) "help reading"[17,18]. The self-reported HL questions were translated into Spanish, back-translated, and extensively pilot-tested. For "confident with forms" the categories were "not at all, a little, somewhat, quite a bit, and extremely"[17,20]. For "problems learning" and "help reading," response categories were "always, often, sometimes, rarely, or never". To create the summative scale, responses were assigned a number from 1 to 5. For "confident with forms" 1 was assigned for a Likert response of "extremely," and 5 for "not at all". For "problems learning" and "help reading" number assignments were reversed. Scores ranged from 3-15 with higher scores reflecting worse self-reported HL.

Standard Health Literacy Measure

As the reference measure, we administered the validated short Test of Functional Health Literacy in Adults (sTOFHLA) in English and Spanish[14]. Higher scores (range 0-36) indicate better reading comprehension. We used standard cut-offs in which scores from 0-16 represent inadequate HL, 17-22 marginal HL, and 23-36 adequate HL[14]. S-TOFHLA scores of 0-22 are collectively referred to as inadequate plus marginal HL. We assessed the performance of the self-reported questions and the summative scale compared to the s-TOFHLA categories of inadequate (scores 0-16) and inadequate plus marginal literacy, (scores 0-22).

Patient Characteristics

We assessed self-reported: language, defined as the language in which participants chose to be interviewed (i.e. English and Spanish); race/ethnicity (Hispanic White, non-Hispanic White, non-Hispanic Black, Asian/Pacific Islander); educational attainment (< high school versus ≥ high school/GED); age (mean and <65 years versus ≥65 years); gender; and health status (fair-to-poor versus good-to-excellent) – patient characteristics which have been associated with HL level[26]. We considered race/ethnicity jointly because of our relatively modest sample size.

Analysis

We used percentages and means to describe our study population. We calculated C-Indices (the area under the receiver operator curve (ROC)), for each question and for multiple cut off points of the summative scale for the HL categories of inadequate (comparing TOFHLA scores of 0-16 versus 17-36) and inadequate plus marginal (comparing TOFHLA scores of 0-22 versus 23-36). A C-index of 1.0 reflects perfect prediction, with both sensitivity and specificity being equal to 1. A C-index of 0.5 reflects discrimination no better than chance[27]. We also calculated multilevel likelihood ratios with 95% confidence intervals (CI), sensitivity, and specificity for each question and the summative scale. In cases of zero responses, a standard continuity correction was applied by adding 0.5 to all of the cells in the two-by-two table prior to computing the LR and the confidence interval[28]. We then assessed whether these questions and the summative scale were equally valid in analyses stratified by language. We used asymptotic methods to determine whether observed differences in the C-indices between the individual questions and the summative scale and between stratified language subgroups were statistically significant. Using the same methods, we also stratified by age, gender, educational attainment, health status, and race/ethnicity to ensure the questions were

equally valid in diverse patient subgroups. This is particularly important among race/ethnic subgroups, because prior studies suggest health literacy may partly explain racial/ethnic health disparities[29]. In these subgroup analyses, comparisons were made between 48 pairs of subgroups; we therefore regarded a difference as statistically significant at a Bonferroni-corrected level of p<0.001[27,30].

RESULTS

Of 296 participants, 48% were Spanish-speaking, and only 9% were white, non-Hispanic (Table 1). Limited HL was prevalent: 47% had inadequate HL as measured by the sTOFHLA and 12% had marginal literacy. For the self-reported HL questions, 57% reported being confident with forms "somewhat" or less, 45% of participants reported problems learning "sometimes" or more frequently, and 42% reported needing help reading "sometimes" or more frequently.

Overall, participants who reported less confidence with forms (C-index 0.82, CI 0.77-0.87), more problems learning (C-index 0.72, CI 0.67-0.78), needing more help reading (C-index 0.68, CI 0.62-0.74), and higher summative scale measures (worse HL) (C-index 0.82, CI 0.77-0.86) were consistently more likely to have inadequate HL (sTOFHLA 0-16), as demonstrated by C-indices >0.5 (range for the questions and scale 0.68-0.84). Overall, these questions also successfully differentiated those with inadequate plus marginal HL (sTOFHLA 0-22) compared to those with adequate HL (sTOFHLA 23-36) (C-indices ranging from 0.69-0.81). ("confident with forms," C-index 0.81, 0.76-0.86; "problems learning," C-index 0.74, 0.68-0.79; "help reading," C-index 0.69, 0.64-0.75; scale, C-index 0.82, 0.77-0.87) The performance of the summative scale was not statistically significantly different from the "confident with forms" question (p for inadequate HL=0.85; p for inadequate plus marginal =0.77). Both the "confident with forms" item and the summative scale performed better than the other 2

questions for both inadequate and inadequate plus marginal HL (p<0.01 for all comparisons).

In our stratified analyses by language, for inadequate (Table 2) and inadequate plus marginal HL (Table 3) the C-indices did not significantly differ between English and Spanish speakers. However, the three questions demonstrated higher sensitivity and lower specificity at any given cut point among Spanish speakers compared to English speakers. Sensitivity, specificity, and likelihood ratios were highest for the "confident with forms" question, among English and Spanish speakers, for identifying both inadequate HL (English, C-index 0.76; Spanish, C-index 0.74) (Table 2) and inadequate plus marginal HL (English, C-index 0.70; Spanish, C-index 0.80) (Table 3). For both inadequate (Table 2) and inadequate plus marginal HL (Table 3), a cut point of "somewhat" or less confident with forms, a cut point used in prior studies[18], appeared to maximize both sensitivity and specificity for English speakers. However, for both literacy levels a cut point of "a little" or less confident with forms functioned best, among Spanish speakers (Table 2 & 3). The test characteristics for the summative scale (See Online Appendix) demonstrate that a cut point of 9, corresponding to answers of "sometimes/somewhat" on all three questions, appeared to maximize both sensitivity and specificity for English and Spanish speakers. (See Online Appendix)

In stratified analyses, after adjustment for multiple comparisons, we found that the self-reported questions performed well and consistently across age, gender, educational attainment, health status, and race/ethnicity participant subgroups for identifying inadequate HL. For inadequate plus marginal HL there was slightly more variation between groups, but none of these differences were statistically significant (See Online Appendix, all P>0.01).

DISCUSSION

Because of its well-established role in health outcomes and health disparities, HL is an important factor to study in public health and epidemiological research[13]. To our knowledge, this is the first study to test the performance of self-reported HL questions among an ethnically diverse, English and Spanish-speaking population, and to compare the performance of the questions between language and other patient characteristic subgroups. We found that three self-reported HL questions could identify those with inadequate, and inadequate plus marginal HL within this ethnically diverse, English and Spanish-speaking population with a moderate degree of discrimination. The "confident with forms" question performed best among the individual items and within both language and all other patient characteristic subgroups. The summative scale performed similarly to the individual "confident with forms" question.

Our findings build on previous studies of the three self-reported HL measures. As in prior studies[17,18,20], the "confident with forms" question performed the best out of the three questions. In contrast to prior work, we found that both the "confident with forms" question and the summative scale could discriminate moderately well between those with inadequate plus marginal vs. adequate HL, in addition to inadequate HL, for both English and Spanish speakers. For

Table 1. Patient Characteristics (N=296)

	N(%)
Age, mean (SD)	54.9 (12.1)
Gender	
Male	126 (42.6)
Female	170 (57.4)
Race	
White, non-Hispanic	25 (8.5)
White, Hispanic	156 (52.7)
Black	70 (23.7)
Asian	40 (13.5)
Multiracial / Other	5 (1.7)
Language	
English	154 (52%)
Spanish	142 (48%)
Income <$20,000	215 (72.6)
Education	
Less than high school	149 (50.3)
High school graduate/GED	52 (17.6)
More than high school	95 (32.1)
Fair or poor health status	223 (75.3%)
Health literacy level by s-TOFHLA	
Inadequate (score, 0-16)	140 (47.3)
Marginal (score, 17-22)	34 (11.5)
Adequate (score, 23-36)	122 (41.2)

Table 2. Test Characteristics for Health Literacy Questions Compared to sTOFHLA Scores for Inadequate Health Literacy

	AUROC	sTOFHLA <17 number of subjects	sTOFHLA≥17 number of subjects	Sensitivity (95% CI)	Specificity (95% CI)	Multilevel LR (95% CI)
ENGLISH						
[a]Confident with Forms	0.76 (0.67 – 0.85)					
[e]Extremely		6	67	1.00 (0.91 – 1.00)	0.00 (0.00 – 0.03)	0.28 (0.13 – 0.60)
Quite a bit		10	21	0.84 (0.69 – 0.92)	0.57 (0.48 – 0.66)	1.51 (0.78 – 2.90)
Somewhat		7	24	0.57 (0.41 – 0.71)	0.75 (0.67 – 0.82)	0.92 (0.43 – 1.97)
A little		8	3	0.38 (0.24 – 0.54)	0.96 (0.90 – 0.98)	8.43 (2.36 – 30.16)
Not at all		6	2	0.16 (0.08 – 0.31)	0.98 (0.94 – 1.00)	9.49 (2.00 – 45.01)
[b]Problems Learning	0.72 (0.63 – 0.82)					
[e]Never		11	78	1.00 (0.91 – 1.00)	0.00 (0.00 – 0.03)	0.44 (0.27 – 0.74)
Rarely		6	15	0.70 (0.54 – 0.83)	0.67 (0.58 – 0.75)	1.26 (0.53 – 3.02)
Sometimes		9	19	0.54 (0.38 – 0.69)	0.79 (0.71 – 0.86)	1.50 (0.74 – 3.02)
Often		3	3	0.30 (0.17 – 0.46)	0.96 (0.90 – 0.98)	3.16 (0.67 – 15.00)
Always		8	2	0.22 (0.11 – 0.37)	0.98 (0.94 – 1.00)	12.65 (2.81 – 56.96)
[c]Help Reading	0.65 (0.55 – 0.75)					
[e]Never		14	69	1.00 (0.91 – 1.00)	0.00 (0.00 – 0.03)	0.64 (0.41 – 1.00)
Rarely		3	14	0.62 (0.46 – 0.76)	0.59 (0.50 – 0.67)	0.68 (0.21 – 2.23)
Sometimes		10	25	0.54 (0.38 – 0.69)	0.71 (0.62 – 0.78)	1.26 (0.67 – 2.38)
Often		1	7	0.27 (0.15 – 0.43)	0.92 (0.86 – 0.96)	0.45 (0.06 – 3.55)
Always		9	2	0.24 (0.13 – 0.40)	0.98 (0.94 – 1.00)	14.23 (3.22 – 62.94)
[d]Summative Scale	0.76 (0.67 – 0.85)					
SPANISH						
[a]Confident with Forms	0.74 (0.66 – 0.83)					
[e]Extremely		4	4	1.00 (0.97 – 1.00)	0.00 (0.00 – 0.09)	0.38 (0.10 – 1.44)
Quite a bit		8	8	0.96 (0.90 – 0.98)	0.10 (0.04 – 0.24)	0.38 (0.15 – 0.94)
Somewhat		15	14	0.88 (0.81 – 0.93)	0.31 (0.19 – 0.46)	0.41 (0.22 – 0.76)
A little		34	10	0.74 (0.65 – 0.81)	0.67 (0.51 – 0.79)	1.29 (0.71 – 2.35)
Not at all		42	3	0.41 (0.32 – 0.50)	0.92 (0.80 – 0.97)	5.30 (1.74 – 16.11)
[b]Problems Learning	0.63 (0.54 – 0.73)					
[e]Never		26	12	1.00 (0.96 – 1.00)	0.00 (0.00 – 0.09)	0.80 (0.45 – 1.42)
Rarely		6	9	0.75 (0.66 – 0.82)	0.32 (0.19 – 0.47)	0.25 (0.09 – 0.64)
Sometimes		34	11	0.69 (0.59 – 0.77)	0.55 (0.40 – 0.70)	1.14 (0.65 – 2.01)
Often		8	3	0.36 (0.27 – 0.46)	0.84 (0.70 – 0.93)	0.98 (0.28 – 3.52)
Always		29	3	0.28 (0.20 – 0.38)	0.92 (0.79 – 0.97)	0.37 (1.15 – 11.02)
[c]Help Reading	0.68 (0.60 – 0.77)					
[e]Never		30	19	1.00 (0.96 – 1.00)	0.00 (0.00 – 0.09)	0.58 (0.38 – 0.90)
Rarely		14	7	0.71 (0.61 – 0.79)	0.50 (0.33 – 0.64)	0.74 (0.32 – 1.69)
Sometimes		24	11	0.57 (0.48 – 0.66)	0.68 (0.51 – 0.80)	0.80 (0.44 – 1.48)
Often		6	0	0.34 (0.26 – 0.44)	0.97 (0.86 – 1.00)	4.88 (0.28 – 84.51)
Always		29	1	0.28 (0.20 – 0.38)	0.97 (0.86 – 1.00)	10.70 (1.51 – 75.84)
[d]Summative Scale	0.74 (0.66 – 0.83)					

[a]*How confident are you filling out medical forms?*
[b]*How often do you have problems learning about your medical condition because of difficulty understanding written information?*
[c]*How often do you have someone help you read hospital materials?*
[d]*Summed responses to all three questions. Complete information about the test characteristics of the summative scale can be found in Online Appendix 1.*
[e]*Referent Category*

the "confident with forms question" Chew et al found a C-index of 0.72 for inadequate plus marginal HL while we found a C-index of 0.81 for the overall sample. This is important because marginal HL, in addition to inadequate HL, has been associated with poor health outcomes including mortality and health disparities[4,12,26]. Because dose response associations have been found between HL level and poor patient outcomes,[31] some investigators may want to identify both literacy level subgroups. Our results also mirror those of prior studies in finding similar performance between the "confidence with forms" item and the summative scale[17].

In stratified analysis by language, the C-indices for the "confidence with forms" question were similar for Spanish and English speakers. However, the item seemed to have higher sensitivity but lower specificity among Spanish speakers at every cut point. The optimum cut point for the "confident with forms" question for English speakers that maximized both sensitivity and specificity was "somewhat" or less, while for Spanish speakers the optimum cutpoint was "a little" or less. These findings may be the result of cultural variation and /or Spanish-speaking participants responding to the 'confident with forms" question for forms not only written in Spanish, but also in English. As such, researchers may want to consider different cut points for English and Spanish-speaking subgroups.

The utility of the "confident with forms" question and summative scale among the Spanish speakers in our population may also be affected by the relatively high prevalence of language concordant patient-physician dyads in this clinical setting and the ubiquitous access to Spanish transcription and translation services[22]. Patient–physician language concordance has been shown to be a powerful determinant of patient satisfaction with communication and may have leveled the playing field with their English-speaking counterparts in terms of patients feeling confident with forms[22]. As

Table 3. Test Characteristics for Health Literacy Questions Compared to sTOFHLA Scores for Inadequate + Marginal Health Literacy

	AUROC	sTOFHLA <23 number of subjects	sTOFHLA ≥ 23 number of subjects	Sensitivity (95% CI)	Specificity (95% CI)	Multilevel LR (95% CI)
ENGLISH						
[a]Confident with Forms	0.70 (0.62 – 0.78)					
[e]Extremely		17	56	1.00 (0.94 – 1.00)	0.00 (0.00 – 0.04)	0.48 (0.31 – 0.74)
Quite a bit		14	17	0.72 (0.59 – 0.81)	0.60 (0.49 – 0.69)	1.29 (0.69 – 2.42)
Somewhat		12	19	0.48 (0.36 – 0.61)	0.78 (0.68 – 0.85)	0.99 (0.52 – 1.89)
A little		10	1	0.28 (0.19 – 0.41)	0.98 (0.93 – 0.99)	15.67 (2.06 – 119.3)
Not at all		7	1	0.12 (0.06 – 0.22)	0.99 (0.94 – 1.00)	10.97 (1.38 – 86.92)
[b]Problems Learning	0.69 (0.61 – 0.77)					
[e]Never		24	65	1.00 (0.94 – 1.00)	0.00 (0.00 – 0.04)	0.58 (0.41 – 0.81)
Rarely		6	15	0.60 (0.47 – 0.71)	0.69 (0.59 – 0.78)	0.63 (0.26 – 1.53)
Sometimes		17	11	0.50 (0.38 – 0.62)	0.85 (0.77 – 0.91)	2.42 (1.22 – 4.81)
Often		4	2	0.22 (0.13 – 0.34)	0.97 (0.91 – 0.99)	3.13 (0.59 – 16.58)
Always		9	1	0.15 (0.08 – 0.26)	0.99 (0.94 – 1.00)	14.1 (1.83 –108.49)
[c]Help Reading	0.66 (0.57 – 0.74)					
[e]Never		23	60	1.00 (0.94 – 1.00)	0.00 (0.00 – 0.04)	0.60 (0.42 – 0.86)
Rarely		5	12	0.62 (0.49 – 0.73)	0.64 (0.54 – 0.73)	0.65 (0.24 – 1.76)
Sometimes		21	14	0.53 (0.41 – 0.65)	0.77 (0.67 – 0.84)	2.35 (1.30 – 4.25)
Often		2	6	0.18 (0.11 – 0.30)	0.91 (0.84 – 0.96)	0.52 (0.11 – 2.50)
Always		9	2	0.15 (0.08 – 0.26)	0.98 (0.93 – 0.99)	7.05 (1.58 – 31.52)
[d]Summative Scale	0.73 (0.64 – 0.81)					
SPANISH						
[a]Confident with Forms	0.80 (0.71 – 0.88)					
[e]Extremely		4	4	1.00 (0.97 – 1.00)	0.00 (0.00 – 0.12)	0.25 (0.07 – 0.92)
Quite a bit		9	7	0.96 (0.91 – 0.99)	0.14 (0.06 – 0.31)	0.32 (0.13 – 0.77)
Somewhat		18	11	0.89 (0.81 – 0.93)	0.39 (0.24 – 0.58)	0.40 (0.21 – 0.75)
A little		39	5	0.73 (0.64 – 0.80)	0.79 (0.60 – 0.90)	1.92 (0.83 – 4.41)
Not at all		44	1	0.39 (0.30 – 0.48)	0.96 (0.82 – 0.99)	10.81 (1.56 – 75.09)
[b]Problems Learning	0.70 (0.61 – 0.80)					
[e]Never		27	11	1.00 (0.97 – 1.00)	0.00 (0.00 – 0.12)	0.61 (0.35 – 1.07)
Rarely		7	8	0.76 (0.67 – 0.83)	0.39 (0.24 – 0.58)	0.21 (0.09 – 0.54)
Sometimes		38	7	0.70 (0.61 – 0.78)	0.68 (0.49 – 0.82)	1.35 (0.67 – 2.69)
Often		10	1	0.36 (0.28 – 0.45)	0.93 (0.77 – 0.98)	2.46 (0.33 – 18.40)
Always		31	1	0.27 (0.20 – 0.36)	0.96 (0.82 – 0.99)	7.68 (1.10 – 53.88)
[c]Help Reading	0.71 (0.63 – 0.80)					
[e]Never		33	16	1.00 (0.97 – 1.00)	0.00 (0.00 – 0.12)	0.51 (0.33 – 0.79)
Rarely		16	5	0.71 (0.62 – 0.78)	0.57 (0.39 – 0.73)	0.79 (0.32 – 1.98)
Sometimes		28	7	0.57 (0.47 – 0.65)	0.75 (0.57 – 0.87)	0.99 (0.48 – 2.03)
Often		6	0	0.32 (0.24 – 0.41)	1.00 (0.88 – 1.00)	3.31 (0.19 – 57.02)
Always		30	0	0.27 (0.19 – 0.35)	1.00 (0.88 – 1.00)	15.52 (0.98 – 246.3)
[d]Summative Scale	0.82 (0.75 – 0.89)					

[a]*How confident are you filling out medical forms?*
[b]*How often do you have problems learning about your medical condition because of difficulty understanding written information?*
[c]*How often do you have someone help you read hospital materials?*
[d]*Summed responses to all three questions. Complete information about the test characteristics of the summative scale can be found in Online Appendix 1.*
[e]*Referent Category*

such, the self-reported measures in this population may have been detecting true HL deficits rather than those related to language discordance or limited English proficiency.

Because of a prior lack of brief, validated measures of HL for diverse populations, some have suggested using demographic characteristics to estimate HL[32]. This approach does not permit the ability to assess the independent effects of HL beyond demographic characteristics. This is important because HL levels have been shown to vary widely within patient demographic subgroups[6]. Therefore, we contend that independent measurement of HL, for example with the "confident with forms" question or summative scale, would contribute substantially to epidemiologic and clinical research. In the clinical setting, screening for limited health literacy is controversial, with the current expert recommendations against routine screening[32–34]. However, in selected clinical situations, such as the prescribing of high-risk medications, screening for limited health literacy has been advocated, and the use of a single-item screener would be more feasible in busy clinical settings than standard literacy assessments[19].

While imperfect in their precision, the summative scale, and specifically the single "confident with forms" question, have some clear advantages over direct, longer HL measurements. They are brief and can be administered via telephone. Our group has recently field-tested these questions both individually and as a scale within a large sample of diverse diabetes patients and have demonstrated robust, independent associations with a range of outcomes, including perceived need for self-management support[35], higher rates of hypoglycemia[36], and lower patient use of electronic health records[37]. While these studies did not assess performance of these items across demographic sub-groups, these associations lend support to the items' predictive validity.

Our study has some limitations. First, we included only patients with poorly controlled diabetes, which may limit generalizability to healthier populations. Second, this study

was conducted at four sites within one county health care system and may not reflect regional differences. Third, in our practice environment there is excellent access to translation services and many physicians and staff speak Spanish. Results may differ for Spanish-speaking patients in different linguistic environments. Finally, our results reflect the criterion validity of the self-reported HL questions, i.e., their relationship with a gold-standard HL measurement. Further work is needed to establish predictive validity of these questions in relation to health outcomes of interest.

In summary, although limited HL is associated with a range of health outcomes, it is often not feasible to measure directly in clinical, epidemiologic, or public health studies because standard measurement tools are lengthy and cannot be administered by telephone. Our study suggests that the single self-reported "confident with forms" question or the summative scale of the three self-reported HL questions discriminate diverse English speakers and Spanish speakers with adequate HL from those with inadequate and inadequate plus marginal HL to a degree that warrants application and further assessment in epidemiologic and clinical research involving diverse populations.

Corresponding Author: *Urmimala Sarkar, MD MPH; Department of Medicine, Division of General Internal Medicine, University of California, Box 1364, 1001 Potrero, Bldg 10, 3rd floor, San Francisco, CA 94143-1364, USA (e-mail: usarkar@medsfgh.ucsf. edu).*

REFERENCES

1. U.S. Department of Health and Human Services. (Government Printing Office. Originally developed for Ratzan SC, Parker RM. 2000. Introduction. In National Library of Medicine Current Bibliographies in Medicine: Health Literacy. Selden CR, Zorn M, Ratzan SC, Parker RM, Editors. NLM Pub. No. CBM 2000-1. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services). Healthy People 2010. 2000.
2. **Sudore RL, Mehta KM, Simonsick EM, et al.** Limited literacy in older people and disparities in health and healthcare access. J Am Geriatr Soc. 2006;54(5):770–6.
3. **Schillinger D, Grumbach K, Piette J, et al.** Association of health literacy with diabetes outcomes. JAMA. 2002;288(4):475–82.
4. **Baker DW, Wolf MS, Feinglass J, Thompson JA, Gazmararian JA, Huang J.** Health literacy and mortality among elderly persons. Arch Intern Med. 2007;167(14):1503–9.
5. **Rothman R, Malone R, Bryant B, Dewalt D, Pignone M.** Health literacy and diabetic control. JAMA. 2002;288(21):2687–8.
6. **Paasche-Orlow MK, Parker RM, Gazmararian JA, Nielsen-Bohlman LT, Rudd RR.** The prevalence of limited health literacy. J Gen Intern Med. 2005;20(2):175–84.
7. **Baker DW, Gazmararian JA, Williams MV, et al.** Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. Am J Public Health. 2002;92(8):1278–83.
8. **Baker DW, Parker RM, Williams MV, Clark WS.** Health literacy and the risk of hospital admission. J Gen Intern Med. 1998;13(12):791–8.
9. **Williams MV, Baker DW, Honig EG, Lee TM, Nowlan A.** Inadequate literacy is a barrier to asthma knowledge and self-care. Chest. 1998;114 (4):1008–15.
10. **Davis TC, Arnold C, Berkel HJ, Nandy I, Jackson RH, Glass J.** Knowledge and attitude on screening mammography among low-literate, low-income women. Cancer. 1996;78(9):1912–20.
11. **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(2):166–72.
12. **Sudore RL, Yaffe K, Satterfield S, et al.** Limited literacy and mortality in the elderly: the health, aging, and body composition study. J Gen Intern Med. 2006;21(8):806–12.
13. Institute of Medicine. (National Academy Press, Institute of Medicine Committee on Quality of Health Care in America). Crossing the quality chasm: A new health system for the 21st century. 2001
14. **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(1):33–42.
15. **Weiss BD, Mays MZ, Martz W, et al.** Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med. 2005;3(6):514–22.
16. **Davis TC, Long SW, Jackson RH, et al.** Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med. 1993;25(6):391–5.
17. **Chew LD, Bradley KA, Boyko EJ.** Brief questions to identify patients with inadequate health literacy. Fam Med. 2004;36(8):588–94.
18. **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(5):561–6.
19. **Powers BJ, Trinh JV, Bosworth HB.** Can this patient read and understand written health information? *JAMA*;304(1):76-84.
20. **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(8):874–7.
21. **Sudore RL, Landefeld CS, Williams BA, Barnes DE, Lindquist K, Schillinger D.** Use of a modified informed consent process among vulnerable patients: a descriptive study. J Gen Intern Med. 2006;21(8):867–73.
22. **Sudore RL, Landefeld CS, Perez-Stable EJ, Bibbins-Domingo K, Williams BA, Schillinger D.** Unraveling the relationship between literacy, language proficiency, and patient-physician communication. Patient Educ Couns. 2009;75(3):398–402.
23. **Lee SY, Stucky BD, Lee JY, Rozier RG, Bender DE.** Short Assessment of Health Literacy-Spanish and English: a comparable test of health literacy for Spanish and English speakers. *Health Serv Res*;45(4):1105-20.
24. **Schillinger D, Hammer H, Wang F, et al.** Seeing in 3-D: Examining the Reach of Diabetes Self-Management Support Strategies in a Public Health Care System. *Health Educ Behav.* 2007.
25. **Schillinger D, Handley M, Wang F, Hammer H.** Effects of self-management support on structure, process, and outcomes among vulnerable patients with diabetes: a three-arm practical clinical trial. Diabetes Care. 2009;32(4):559–66.
26. Institute of Medicine. (National Academy Press, Institute of Medicine Committee on Health Literacy). Health Literacy: A Prescription to End Confusion. 2004.
27. **Hanley JA, McNeil BJ.** The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29–36.
28. **Weinberger M, Kirkman MS, Samsa GP, et al.** A nurse-coordinated intervention for primary care patients with non-insulin-dependent diabetes mellitus: impact on glycemic control and health-related quality of life. J Gen Intern Med. 1995;10(2):59–66.
29. **Osborn CY, Cavanaugh K, Wallston KA, White RO, Rothman RL.** Diabetes numeracy: an overlooked factor in understanding racial disparities in glycemic control. Diabetes Care. 2009;32(9):1614–9.

30. **Hanley JA, McNeil BJ.** A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology. 1983;148(3):839–43.

31. **Wolf MS, Davis TC, Shrank W, et al.** To err is human: patient misinterpretations of prescription drug label instructions. Patient Educ Couns. 2007;67(3):293–300.

32. **Paasche-Orlow MK, Wolf MS.** Evidence does not support clinical screening of literacy. J Gen Intern Med. 2008;23(1):100–2.

33. **Parikh NS, Parker RM, Nurss JR, Baker DW, Williams MV.** Shame and health literacy: the unspoken connection. Patient Educ Couns. 1996;27 (1):33–9.

34. **Wolf MS, Williams MV, Parker RM, Parikh NS, Nowlan AW, Baker DW.** Patients' shame and attitudes toward discussing the results of literacy screening. J Health Commun. 2007;12(8):721–32.

35. **Sarkar U, Piette JD, Gonzales R, et al.** Preferences for self-management support: findings from a survey of diabetes patients in safety-net health systems. Patient Educ Couns. 2008;70(1): 102–10.

36. **Sarkar U, Karter A, Adler N, Liu J, Moffet H, Schillinger D.** Limited health literacy is associated with increased risk of hypoglycemia among insured, type 2 diabetes patients: the Diabetes Study of Northern California (DISTANCE) Society for General Internal Medicine. Pittsburgh, PA. J Gen Intern Med. 2008;23 (s2):338.

37. **Sarkar U, Karter AJ, Liu JY, et al.** The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system—results from the Diabetes Study of Northern California (DISTANCE). J Health Commun. 2010;15(1):183–196.