


Development and Validation of the HL6: a Brief, Technology-Based Remote Measure of Health Literacy



Stacy Cooper Bailey, PhD MPH¹ , James W. Griffith, PhD², Chandana Vuyyuru, MPH¹, Stephanie Batio, MS¹, Evelyn Velazquez, BA¹, Delesha M. Carpenter, PhD MSPH³, Terry C. Davis, PhD⁴, Ruth M. Parker, MD MACP⁵, Michelle Taddeo, MPH², and Michael S. Wolf, PhD MPH¹

¹Center of Applied Health Research on Aging, Division of General Internal Medicine & Geriatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA; ²Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA; ³Division of Pharmaceutical Outcomes and Policy, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA; ⁴Department of Medicine, Louisiana State University Health Shreveport, Shreveport, LA, USA; ⁵Center for Ethics, Emory University, Atlanta, GA, USA.

BACKGROUND: Most health literacy measures require in-person administration or rely upon self-report.

OBJECTIVE: We sought to develop and test the feasibility of a brief, objective health literacy measure that could be deployed via text messaging or online survey.

DESIGN: Participants were recruited from ongoing NIH studies to complete a phone interview and online survey to test candidate items. Psychometric analyses included parallel analysis for dimensionality and item response theory. After 9 months, participants were randomized to receive the final instrument via text messaging or online survey.

PARTICIPANTS: Three hundred six English and Spanish-speaking adults with ≥ 1 chronic condition

MAIN MEASURES: Thirty-three candidate items for the new measure and patient-reported physical function, anxiety, depression, and medication adherence. All participants previously completed the Newest Vital Sign (NVS) in parent NIH studies.

KEY RESULTS: Participants were older (average 67 years), 69.6% were female, 44.3% were low income, and 22.0% had a high school level of education or less. Candidate items loaded onto a single factor (RMSEA: 0.04, CFI: 0.99, TLI: 0.98, all loadings $>.59$). Six items were chosen for the final measure, named the HL6. Items demonstrated acceptable internal consistency ($\alpha=0.73$) and did not display differential item functioning by language. Higher HL6 scores were significantly associated with greater educational attainment ($r=0.41$), higher NVS scores ($r=0.55$), greater physical functioning ($r=0.26$), fewer depressive symptoms ($r=-0.20$), fewer anxiety symptoms ($r=-0.15$), and fewer barriers to medication adherence ($r=-0.30$; all $p<.01$). In feasibility testing, 75.2% of participants in the text messaging arm completed the HL6 versus 66.2% in the online survey arm ($p=0.09$).

Prior presentations: findings related to psychometric testing were presented virtually in an oral abstract for the 2021 International Conference on Communication in Healthcare (ICCH) on October 20, 2021. Findings from feasibility testing were presented in an oral abstract for the 2022 Society of General Internal Medicine (SGIM) Annual Meeting in Orlando, FL on April 8, 2022.

Received March 16, 2022

Accepted July 1, 2022

Socioeconomic disparities in completion were more common in the online survey arm.

CONCLUSIONS: The HL6 demonstrates adequate reliability and validity in both English and Spanish. This performance-based assessment can be administered remotely using commonly available technologies with fewer logistical challenges than assessments requiring in-person administration.

KEY WORDS: health literacy; measurement; technology

J Gen Intern Med

DOI: 10.1007/s11606-022-07739-3

© The Author(s), under exclusive licence to Society of General Internal Medicine 2022

INTRODUCTION

Health literacy, or one's ability "to find, understand, and use information and services to inform health-related decisions and actions for themselves and others," has been linked to numerous health behaviors and health outcomes over the past few decades.¹⁻⁴ As the evidence detailing the impact of health literacy on health outcomes has proliferated, so has the need for accurate and reliable measurement of individuals' health literacy skills.^{5,6} As a result, researchers have developed a number of health literacy assessments that can be used across various disease states, languages, age groups, and research contexts.⁷

Despite the growth of this field, there is a paucity of tools available to measure health literacy via currently available consumer technologies; the most commonly used health literacy measures require in-person administration by a trained interviewer, which can be both costly and time-consuming. In fact, the Health Literacy Tool Shed, a National Library of Medicine-funded database of available health literacy assessments, shows that although more than 205 health literacy measures exist, there is no single measure that is a brief (<5 min), performance-based, general health literacy assessment that can be administered remotely in English or Spanish to adults.^{6,7} To our knowledge, there are also no measures that

have been validated for completion via text messaging, currently one of the most commonly used methods of communication.^{8,9}

As 93% of American adults report accessing and using the Internet and 97% own a cell phone (85% own a smartphone), opportunities now exist to measure health literacy skills remotely via commonly available consumer technologies.^{8,9} For this study, we sought to develop a brief, performance-based health literacy measure for adults—in both English and Spanish—that could be administered via online survey or text messaging. We also sought to explore the feasibility of deploying this measure using these modalities. The availability of such a measure could provide opportunities to study health literacy in new, previously unexplored contexts, utilizing fewer resources. Importantly, such a measure could also be used safely during COVID-19 or other pandemics without the need for in-person contact to measure health literacy.

METHODS

We used qualitative and quantitative methodology to develop and validate a new measure of health literacy. This multi-phase study included (1) item generation and refinement by health literacy and cognitive science experts and with input from community-dwelling adults, (2) psychometric testing via structured in-person interviews and online surveys, and (3) feasibility testing of the final measure among participants randomized to receive the instrument via either text messaging or online survey. All study procedures were approved by our Institutional Review Board.

Item Generation

We sought to create a brief measure of health literacy that could be easily completed remotely among individuals from diverse backgrounds and with varying levels of comfort with technology. As such, we acknowledged that the measure would be limited in scope and unable to comprehensively measure all skills and attributes reflected in the concept of health literacy.^{1,4} To ground the item generation process, we therefore reviewed health forms and tasks that patients are commonly required to complete in healthcare settings. To generate candidate items for the measure, we sought expert opinions from researchers, healthcare providers, educators, and psychometricians. We considered multiple types of assessments and test formats used to assess reading and cognitive skills, such as spot-the-word, Cloze procedure, multiple choice, and open-ended response.^{10,11} With this background, we generated a set of candidate items that used different formats for review and potential inclusion in the measure.

To obtain input from the target audience on candidate items, we recruited convenience samples of English- and Spanish-speaking adults. Participants were identified via Internet advertisements or from prior participation in research studies led by our team. A total of 34 adults (19 English speakers and 15

Spanish speakers) provided feedback on items via cognitive interviews and discussion groups. During these sessions, we asked participants to review candidate items, reflect on how they interpreted the items, and offer specific suggestions for improvement.

Candidate Item Selection and Testing

Through the item development process described above, we created 33 candidate items for further testing. Based on the input from patients and our scientific advisors, all items used a multiple-choice format with 4 response options. Due to the COVID-19 pandemic, our original methodology for psychometric testing via in-person interviews had to be modified. Instead, we tested items among a unique cohort of patients who had already completed in-person health literacy assessments prior to the pandemic.

Study Participants. Study participants were active enrollees in the COVID-19 and Chronic Conditions (C3) study. This study was launched in March 2020 during the first week of the outbreak. C3 leveraged five active NIH projects (e.g., “parent studies” R01AG030611; R01AG046352; R01DK110172; R01HL126508; R01NR015444) to rapidly recruit participants to form the C3 cohort. The objective of C3 was to capture the experiences of middle-aged and older adults during the pandemic. All of the parent studies had uniform data collection of sociodemographic, health literacy, and patient-reported outcomes, as well as access to participants’ electronic health record (EHR) data. In brief, parent studies included a longitudinal cohort study examining cognitive function and aging among older adults (called LitCog) and three randomized trials evaluating strategies to improve medication adherence and safety. Participants were originally recruited from academic or community health centers in Chicago, IL.¹²

Eligibility criteria for each parent study varied and have been described elsewhere in detail.¹³ In general, the target population for these studies was middle-aged or older, English-speaking patients, who had been diagnosed with one or more chronic conditions according to electronic health record data. One trial also recruited Spanish-speaking adults.¹⁴ Only participants who provided consent to be contacted for future research were eligible for the C3 study or this instrument development study. Additional eligibility criteria to enroll in this instrument development study included (1) access to a personal email, (2) prior completion of the Newest Vital Sign, an in-person, validated, performance-based health literacy measure,¹⁵ (3) ownership of a personal cell phone, and (4) willingness to send and receive text messages.

Data Collection Procedures. Trained research assistants (RAs) contacted patients and completed a phone-based survey with them as part of the C3 study. At the conclusion of the interview, participants were invited to participate in the instrument development study. Eligible and interested participants

were sent a link to an online consent form. After providing consent, participants were automatically directed to the online survey of 33 candidate items. If a participant did not complete the survey within 7 days, a reminder email was sent.

To maximize the participation of Spanish-speaking patients, we enrolled an additional 51 patients who had not participated in the C3 study but were active enrollees in one of the C3 parent studies (R01AG046352). These individuals completed the same relevant measures as the C3 study via telephone and were then directed to complete the online survey of 33 candidate items. Participants received \$40 for their participation in the psychometric testing phase of the study.

Measures. All participants completed standardized assessments of personal attributes as part of their participation in the parent studies. This included items assessing sociodemographic (e.g., age, sex, race, ethnicity, poverty level, education, employment status) and health characteristics (chronic conditions, overall health) as well as health literacy (The Newest Vital Sign).¹⁵ Participants in the LitCog cohort study (R01AG030611) also completed two other commonly used, validated measures of health literacy: the Rapid Estimate of Adult Literacy in Medicine (REALM) and Test of Functional Health Literacy Assessment (TOFHLA).^{16,17}

As part of the phone interviews for C3 and/or this study, participants also completed the Ask-12 Barriers to Medication Adherence survey and Patient-Reported Outcomes Measure Information Systems (PROMIS) assessments of anxiety (Short Form 4a and 7a), depression (Short Form 4a), and physical function (Short Form 10a).^{18,19}

Analyses. Psychometric analyses were performed to select items for the brief health literacy measure. To assess dimensionality, a scree plot and parallel analysis using Lubbe's method compared the empirically observed eigenvalues to eigenvalues drawn from randomly shuffled data.²⁰ Additionally, Differential Item Functioning (DIF) analyses were performed to choose items not impacted by language. Then, we performed item response theory (IRT) analyses, using a 2PL model, as well as a simpler 1PL model with a common slope for all items. Empirical reliability was estimated, along with Pearson's correlations for convergent and predictive validity. IRT reliability was implemented using the mirt package of R. Kappa statistic was used to test the agreement between the new tool and other health literacy measures.²¹

Descriptive statistics, including means with standard deviation and percentage frequencies, were calculated for all patient characteristics and item responses. Associations between the new tool and patient characteristics were measured using χ^2 test, *t*-test, or analysis of variance, as appropriate. Psychometric analyses were performed using R 4.1.2 and all other analyses using SAS, version 9 (Cary, NC).

Finalizing the HL6 Instrument

The final items selected for the tool were compiled in a draft measure, named the Health Literacy-6 or HL6. Using a videoconferencing platform, cognitive interviews were conducted with 9 participants (4 Spanish speakers, 5 English speakers) to finalize measure instructions and appearance. Participants in the cognitive interviews were shown a draft of the HL6 on a screen and asked to complete the tool in a "think aloud" interview. Based on participant feedback, the instructions and item order were finalized, as well as a suggested time limit of 8 min for completion. This time limit was considered long enough to give participants with low literacy skills and/or poor familiarity with technology enough time to read and answer items, but not so much time that participants would be tempted to search for answers online or from others. An additional 10 participants pilot-tested the tool in each modality ($n=5$ per modality, per language) to ensure that the measure was easy to complete and that the 8-min time limit was appropriate in a "real-world" context. All participants were recruited via convenience sampling and provided informed consent.

Feasibility Testing

In the final phase of the study, we investigated the feasibility of completing the new HL6 instrument via the two modalities. Specifically, we sought to determine whether participants' willingness and ability to complete the measure online or via text message varied by participant characteristics, such as age, race, ethnicity, health literacy, and education.

No additional participants were recruited for feasibility testing. Instead, we used the same participants from prior psychometric testing. Participants were randomized to receive the HL6 via either (1) text messaging or (2) online survey. Randomization stratified by language was performed using the Proc Rank statement in SAS version 9 (Cary, NC). Participants were then sent the final instrument via their assigned modality approximately 9 months after the conclusion of the psychometric testing phase. Participants did not receive any financial incentives for completing the measure. Differences between completion rates by modality and participant characteristics were assessed using χ^2 test, *t*-test, or analysis of variance, as appropriate.

RESULTS

Study Sample

A total of 488 participants were approached for this study. Of those, 42 could not be reached, 32 refused, 4 were lost to follow-up, 39 were ineligible, and 65 never completed the online survey. A total of 306 participants were enrolled (Table 1). On average, participants were older (average of 67 years), 69.6% were female, and 22.0% had a high

Table 1 Participant Characteristics and HL6 Score (N=306)

Variable	Overall*	HL6 total score (0–6)		P
		Mean (SD)		
Age				0.002
< 60	62 (20.3)	4.6 (1.6)		
60–69	112 (36.6)	4.9 (1.5)		
≥ 70	132 (43.1)	5.3 (1.2)		
Language spoken				<0.001
English	225 (73.5)	5.3 (1.2)		
Spanish	81 (26.5)	4.2 (1.7)		
Sex				<0.001
Female	213 (69.6)	4.8 (1.5)		
Male	93 (30.4)	5.5 (1.0)		
Race and ethnicity				<0.001
Hispanic	98 (32.0)	4.5 (1.7)		
Non-Hispanic White	133 (43.5)	5.7 (0.7)		
Non-Hispanic Black	65 (21.2)	4.6 (1.5)		
Other	10 (3.3)	5.1 (1.9)		
Education				<0.001
High school graduate or less	67 (22.0)	3.9 (1.8)		
Some college or technical school	76 (24.9)	5.1 (1.3)		
College graduate or higher	162 (53.1)	5.5 (1.0)		
Employed				0.29
Yes	95 (31.1)	4.9 (1.6)		
No	210 (68.9)	5.1 (1.3)		
Poverty level				<0.001
Above	170 (55.7)	5.7 (0.7)		
Below	135 (44.3)	4.2 (1.6)		
Health literacy, NVS				<0.001
Low	65 (21.4)	3.9 (1.7)		
Marginal	79 (26.1)	4.6 (1.5)		
Adequate	159 (52.5)	5.7 (0.6)		
Health literacy, REALM ^a				0.001
Limited	12 (9.0)	3.5 (1.6)		
Adequate	122 (91.0)	5.5 (1.0)		
Health literacy, TOFHLA ^b				0.002
Limited	15 (11.0)	3.9 (1.7)		
Adequate	121 (89.0)	5.5 (1.0)		
≥3 chronic conditions				0.11
Yes	135 (44.3)	4.9 (1.4)		
No	170 (55.7)	5.1 (1.4)		

*Values are frequencies and percentages, unless otherwise noted.

^an=134, ^bn=136

school level of education or less. The sample was racially and ethnically diverse, almost half (44.3%) lived below the 2020 US Federal poverty level, and 26.5% spoke Spanish as their primary language. Nearly a quarter (21.4%) had low health literacy according to the NVS.¹⁵

Item Selection

Of the 33 candidate items, 22 were removed from consideration as they displayed differential item functioning by language, using the lordif package for DIF analyses using a McFadden pseudo-R² of 0.18 as the cutoff. For the 11 remaining items, we used the mirt package to estimate model IRT parameters, model fit, and factor loadings. Factor loadings were all .59 or greater based on the 2PL model, and model fit was good; RMSEA = .04, CFI = .99, TLI = .98. We selected 6 items from the 11 based on factor loadings from the 2PL model, endorsement frequencies, and item content. This resulted in the final HL6 instrument, which demonstrated adequate internal consistency reliability ($\alpha = 0.73$; Table 2).

HL6 Performance

Individuals who spoke Spanish as their primary language, were female, were Hispanic or Black, were younger, had a high school level of education or less, and lived below the poverty line were significantly more likely to score lower on the HL6 (Table 1). The overall range of HL6 scores was 0 to 6, with an average score of 5 (standard deviation: 1.4).

The HL6 demonstrated moderate to high convergent validity with the NVS ($r=0.55$, $p<0.001$; Table 3). Among the cohort of LitCog participants who completed additional health literacy measures, the HL6 demonstrated moderate convergent validity with the REALM and TOFHLA ($r=0.51$, $p<0.001$; $r=0.45$, $p<0.001$, respectively, Table 3). Among the entire sample, higher HL6 scores were also predictive of greater

Table 2 Correctly Answered HL6 Items and IRT Analyses (n=306)**

Items	Correctly answered*	Factor loadings
The chance of having a side effect from a medication is 5%. If 1000 people take the medicine, about how many will have a side effect?	234 (76.5)	0.71
Anita weighs 250 pounds. She loses 10% of her body weight. How much does she weigh now?	270 (88.2)	0.60
A pulse is measured in beats per minute. If you count 40 beats in 30 s, how many beats is that per minute?	250 (81.7)	0.84
The flu vaccine reduces the risk of flu for 60% of a population. Among 1000 people, how many people have a reduced risk?	232 (75.8)	0.81
A blood pressure reading under 120/80 mmHg is considered normal. Which of these values is normal?	207 (67.6)	0.79
If you jog for 20 min 4 days a week, how much time in total do you spend jogging in a week?	285 (93.1)	0.66
You eat 2100 mg of sodium daily. Your doctor recommends reducing your daily intake by a third. How much sodium should you eat?	163 (53.2)	0.81
Half a cup of yogurt has 155 calories. If you eat 1 cup of yogurt, how many calories will you eat?	281 (91.8)	0.80
You want to consume 2500 calories a day. If you had a 300-calorie breakfast and a 1000-calorie lunch, how many calories should you consume for dinner?	288 (94.1)	0.83
Take 1–2 pills every 6 h for pain. Do not exceed 3000mg a day. Imagine each pill is 600mg. What is the maximum number of pills you can take daily?	256 (83.7)	0.78
This vaccine will protect against all three influenza strains that are expected to circulate this year.	266 (87.2)	0.76

*Values are frequencies and percentages, unless otherwise noted

†Bolded items are used in the final HL6 measure

Table 3 Psychometric Properties of HL6 (N=306)

	Overall*	Association with HL6 [†]	P
Convergent validity			
NVS	3.4 (2.0)	0.55	<0.001
REALM ^a	63.7 (5.0)	0.51	<0.001
TOFHLA ^b	82.6 (9.3)	0.45	<0.001
Predictive validity			
Physical function (PROMIS)	46.6 (8.9)	0.26	<0.001
Depression (PROMIS)	50.0 (8.5)	-0.20	<0.001
Anxiety (PROMIS)	51.9 (9.0)	-0.15	0.008
Medication adherence (Ask-12)	9.9 (7.8)	-0.30	<0.001

*Values are mean (SD), unless otherwise noted
[†]Pearson's correlations
^an=134, ^bn=136 (LitCog participants only)

physical function ($r=0.26, p<0.001$), fewer depressive symptoms ($r=-0.20, p<0.001$), fewer anxiety symptoms ($r=-0.15, p=0.008$), and fewer barriers to medication adherence ($r=-0.30, p<0.001$).

Feasibility Study

For the feasibility study, the final instrument was sent via text message to 154 participants and via an online survey to 152 participants. Of the 306 participants, 211 completed the HL6 in their assigned modality (online survey: 98, text message: 113). Of those who did not complete the HL6 in either modality ($n=95$), 88 did not initiate the instrument, 2 did not complete the entire measure, and 5 participants took longer than the 8 min allowed to finish the items. In the text message arm, 4 of the 38 patients who did not complete the HL6 never received it due to disconnected or wrong phone numbers. We do not know how many participants in the online survey arm did not receive the emailed survey link. We are also unable to determine which type of device (e.g., computer, tablet, smartphone) patients in the online survey arm used to complete the survey. Patients took an average of 3 min to complete the remote measure in either modality.

There were no significant differences in completion rates between modalities, with 75.2% of participants in the text message arm completing the measure versus 66.2% in the online survey arm ($p=0.09$; Table 4). However, socioeconomic differences in completion were more marked in the online

Table 4 Feasibility Testing of the HL6 (N=306)*

	Online HL6		P	Text HL6		P
	Did not complete (n=54)	Completed testing (n=98)		Did not complete (n=41)	Completed testing (n=113)	
Age			0.01			<0.001
< 60	15 (41.7)	21 (58.3)		14 (53.9)	12 (46.1)	
60-69	24 (48.0)	26 (52.0)		17 (27.4)	45 (72.6)	
≥ 70	15 (22.7)	51 (77.3)		10 (15.2)	56 (84.8)	
Language			<0.001			0.01
English	28 (25.0)	84 (75.0)		24 (21.2)	89 (78.8)	
Spanish	26 (65.0)	14 (35.0)		17 (41.50)	24 (58.5)	
Gender			0.92			0.90
Female	39 (35.8)	70 (64.2)		28 (26.9)	76 (73.1)	
Male	15 (34.9)	28 (65.1)		13 (26.0)	37 (74.0)	
Race			0.001			0.21
Hispanic	27 (54.0)	23 (46.0)		18 (37.5)	30 (62.5)	
Non-Hispanic Whites	14 (20.3)	55 (79.7)		13 (20.3)	51 (79.7)	
Non-Hispanic Blacks	11 (39.3)	17 (60.7)		9 (24.3)	28 (75.7)	
Others	2 (40.0)	3 (60.0)		1 (20.0)	4 (80.0)	
Education			0.005			0.32
HS graduate or less	21 (55.3)	17 (44.7)		10 (34.5)	19 (65.5)	
Some college	10 (38.5)	16 (61.5)		15 (30.0)	35 (70.0)	
College or more	22 (25.3)	65 (74.7)		16 (21.3)	59 (78.7)	
Employed			0.52			0.35
Yes	14 (31.8)	30 (68.2)		16 (31.4)	35 (68.6)	
No	40 (37.4)	67 (62.6)		25 (24.3)	78 (75.7)	
Poverty level ^a			<0.001			0.83
Above	35 (50.7)	34 (49.3)		24 (27.3)	64 (72.7)	
Below	19 (23.2)	63 (76.8)		17 (25.8)	49 (74.2)	
Health literacy, NVS			<0.001			0.86
Low	25 (78.1)	7 (21.9)		10 (30.3)	23 (69.7)	
Marginal	11 (30.6)	25 (69.4)		11 (25.0)	33 (75.0)	
Adequate	18 (21.7)	65 (78.3)		20 (26.0)	57 (74.0)	
≥3 chronic conditions			0.65			0.87
Yes	26 (37.7)	43 (62.3)		18 (27.3)	48 (72.7)	
No	28 (34.2)	54 (65.8)		23 (26.1)	65 (73.9)	
Physical function, mean (SD)	45.1 (8.2)	46.5 (9.3)	0.34	47.9 (9.6)	47.1 (8.6)	0.60
Depression, mean (SD)	52.2 (8.4)	49.9 (8.6)	0.10	50.9 (9.3)	48.7 (8.0)	0.14
Anxiety, mean (SD)	53.1 (9.4)	52.3 (8.9)	0.61	51.9 (9.0)	50.9 (8.8)	0.55
Medication adherence, mean (SD) ^b	14.4 (9.2)	7.8 (6.0)	<0.001	11.0 (8.2)	9.2 (7.5)	0.20

*Values are frequencies and percentages, unless otherwise noted. ^aAccording to 2020 US Federal Poverty Level Guidelines. ^bFive missing ASK12 adherence

survey arm. Patients receiving the measure via online survey were significantly less likely to complete it if they were Spanish-speaking, Hispanic, younger, less educated, living below the poverty level, and had lower health literacy according to the NVS.¹⁵ Among patients receiving the HL6 via text messaging, the only significant predictors of non-completion were younger age and speaking Spanish (Table 4).

DISCUSSION

Using mixed methods, we developed and tested a pool of items that could be used to broadly capture adults' health literacy skills. From this pool, we created a unidimensional, reliable, and brief measure: the HL6. The HL6 demonstrated convergent validity with the NVS, REALM, and TOFHLA, validated, commonly used measures of health literacy that are administered in-person.^{15,16,17} A summed score of the HL6 items was also significantly associated with key patient health behaviors and outcomes, including medication adherence, physical function, depression, and anxiety. Importantly, our study demonstrated that adults can complete the measure themselves, remotely, in an average of 3 min.

Interestingly, findings from our feasibility test indicate that text messaging may be an optimal modality for surveys, particularly among socioeconomically disadvantaged populations. This is consistent with evidence from other studies, which have shown a patient preference for completing surveys via text messaging, and have found socioeconomic differences in participation with online surveys.^{22–24} In our study, 3 out of 4 participants in the text messaging arm completed the HL6 when it was sent to them via text, even though they received no additional compensation for completion. The only significant predictors of non-completion were speaking Spanish and younger age. The latter, while statistically significant, may not be particularly meaningful as study participants were predominately older. In contrast, completion of the HL6 by online survey differed by key participant attributes, including education level, language spoken, ethnicity, income, age, and health literacy level according to the NVS. While not examined in this study, it is also plausible that deploying the HL6 via text messaging could help bridge the rural digital divide as cell phone ownership is more pervasive in rural communities than access to high-speed Internet.²⁵ Additional research is needed, however, to further investigate the feasibility of the HL6 in both modalities and across diverse populations and contexts.

There are limitations to this research that should be noted. First, because of the COVID-19 pandemic, we were unable to conduct this study as originally conceived as it would have required in-person interviews. Instead, we recruited and used data from existing and prior study participants, who had already completed an in-person assessment of health literacy, to develop and evaluate our measure. This affects the generalizability of our results, yet further underscores the need for a remote measure of health literacy, especially during the

pandemic. Second, because of the cross-sectional design of our study, we are unable to infer a causal relationship between HL6 scores and patient outcomes such as medication adherence, functional health, depression, and anxiety. However, findings did show a significant association between HL6 scores and these constructs. Finally, our study sample was predominately older and lower-income, all had at least one chronic condition and nearly half had 3 or more. While our sample represents arguably the most difficult patients to reach and those for whom health literacy is likely to be a challenge, it does affect the generalizability of our results to younger and healthier populations. Additional research is therefore needed to test the tool in other populations and to determine its ability to predict outcomes over time. Our team is now including the HL6 in multiple recently funded research studies, including two large, longitudinal NIH studies among younger and middle-aged adults (1R01DK127184; 1R01AG070212). Results will help determine the utility of this new tool and its performance in comparison to other commonly used, validated, in-person measures.

In summary, the HL6 appears to be a unidimensional, reliable, and brief assessment of adult health literacy skills that can be successfully completed by patients remotely. Given the ease of use of this tool, and its availability in both English and Spanish, we envision the HL6 to be a useful asset for future studies, particularly when in-person interviews are not feasible. The HL6 is available in the public domain, at no cost, and should be used for research purposes only. While additional validation studies are warranted, we hope the HL6 can help advance the field of health literacy measurement and research and help researchers overcome many of the challenges posed to this area of research during the pandemic.

Corresponding Author: Stacy Cooper Bailey, PhD MPH; Center of Applied Health Research on Aging, Division of General Internal Medicine & Geriatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA (e-mail: stacy-bailey@northwestern.edu).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11606-022-07739-3>.

Funding Supported in part by a research grant from the Investigator-Initiated Studies Program of Merck Sharp & Dohme Corp. The opinions expressed in this paper are those of the authors and do not necessarily represent those of Merck Sharp & Dohme Corp. Research reported in this publication was supported, in part, by the National Institutes of Health's National Center for Advancing Translational Sciences, Grant Number UL1TR001422, and National Institute on Aging, Grant Number P30AG059988. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Declarations:

Conflict of Interest: Dr. Bailey reports grants from the NIH, Merck, Pfizer, Gordon and Betty Moore Foundation, Retirement Research Foundation for Aging, Lundbeck, and Eli Lilly via her institution and personal fees from Sanofi, Pfizer, University of Westminster, Lundbeck, and

Luto UK outside the submitted work. Dr. Wolf reports grants from the NIH (NIA, NIDDK, NINR, NHLBI, NINDS), Gordon and Betty Moore Foundation, and Eli Lilly, and personal fees from Pfizer, Sanofi, Luto UK, University of Westminster, and Lundbeck outside the submitted work.

REFERENCES

- Institute of Medicine, Health literacy: A prescription to end confusion, ed. Nielsen-Bohman L, Panzer A, Kindig DA. 2004, Washington DC: National Academy Press.
- DeWalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. *J Gen Intern Med.* 2004; 19(12): 1228-1239.
- Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med.* 2011; 155(2): 97-107.
- Office of Disease Prevention and Health Promotion. Health Literacy in Healthy People 2030. Available at: <https://health.gov/our-work/national-health-initiatives/healthy-people/healthy-people-2030/health-literacy-healthy-people-2030>. Accessed 26 Jan 2022.
- Baker DW. The meaning and the measure of health literacy. *J Gen Intern Med.* 2006; 21(8): 878-883.
- Nguyen TH, Paasche-Orlow MK, McCormack LA. The state of the science of health literacy measurement. *Stud Health Technol Inform.* 2017; (240): 17-33.
- Boston University. Health Literacy Tool Shed: A Database of Health Literacy Measures. Available at: <https://healthliteracy.bu.edu/>. Accessed 21 Jan 2022.
- Pew Research Center. Internet/Broadband Fact Sheet. Available at: <https://www.pewresearch.org/internet/fact-sheet/internet-broadband/>. Accessed 17 June 2022.
- Pew Research Center. Mobile Fact Sheet. Available at: <http://www.pewinternet.org/fact-sheet/mobile/>. Accessed 17 June 2022.
- Baddeley A, Emslie H, Nimmo-Smith I. The Spot-the-Word test: a robust estimate of verbal intelligence based on lexical decision. *Br J Clin Psychol.* 1993; 32(1): 55-65.
- Taylor WL. Cloze procedure: a new tool for measuring readability. *Journalism Quarterly.* 1953; 30(4): 415-433.
- Bailey SC, Serper M, Opsasnick L, et al. Changes in COVID-19 Knowledge, Beliefs, Reported Behaviors, and Perceived Preparedness among High Risk Adults from the Onset to the Acceleration Phase of the U.S. Outbreak. *J Gen Intern Med.* 2020; 35(11): 3285-3292.
- Wolf MS, Serper M, Opsasnick L, et al. Awareness, Attitudes, and Actions Related to COVID-19 Among Adults With Chronic Conditions at the Onset of the U.S. Outbreak: A Cross-sectional Survey. *Ann Intern Med.* 2020; 173(2): 100-109.
- Bailey SC, Wismer GA, Parker RM, et al. Development and rationale for a multifactorial, randomized controlled trial to test strategies to promote adherence to complex drug regimens among older adults. *Contemp Clin Trials.* 2017; 62(21-26).
- 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.
- 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.
- 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.
- Matza LS, Park J, Coyne KS, Skinner EP, Malley KG, Wolever RG. Derivation and validation of the ASK-12 adherence barrier survey. *Ann Pharmacother.* 2009; 43(10): 1621-30.
- Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol.* 2010; 63(11): 1179-94.
- Lubbe D. Parallel analysis with categorical variables: Impact of category probability proportions on dimensionality assessment accuracy. *Psychol Methods.* 2019; 24(3): 339-351.
- Chalmers RP. mirt: a multidimensional item response theory package for the r environment. *J Stat Softw.* 2012; 48(1): 1-29.
- Dahl M, Morocco E, Rodriguez L, Dancz C. Acceptability of text messaging and disparities in mobile health access in Women's Health Care: a survey study. *Telemed J E Health.* 2021; 27(8): 866-873.
- Rai M, Moniz MH, Blaszcak J, Richardson CR, Chang T. Real-time data collection using text messaging in a primary care clinic. *Telemed J E Health.* 2017; 23(12): 955-963.
- Jang M, Vorderstrasse A. Socioeconomic status and racial or ethnic differences in participation: web-based survey. *JMIR Res Protoc.* 2019; 8(4): e11865.
- Vogels EA. Some Digital Divides Persist between Rural, Urban and Suburban America. Available at: <https://www.pewresearch.org/fact-tank/2021/08/19/some-digital-divides-persist-between-rural-urban-and-suburban-america/>. Accessed 17 June 2022.

Publisher's Note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.