CrossMark

Assessing Cancer Health Literacy among Spanish-Speaking Latinos

Margarita Echeverri[1] · David Anderson[2] · Anna María Nápoles[3]

Abstract Health literacy is a dynamic construct that changes with specific health conditions; thus, new disease-specific health literacy tools are needed. Since cancer is the leading cause of death among Latinos, the largest and fastest-growing minority population in the nation, there is a need to develop tools to assess cancer health literacy (CHL) among the Spanish-speaking population. The Cancer Health Literacy Test, Spanish version (CHLT-30-DKspa) was applied to identify Spanish-speaking individuals with low CHL and ascertain which items in the tool best discriminate between CHL level groups. Cross-sectional field test of the CHLT-30-DKspa among Spanish-speaking Latinos. Latent class analysis (LCA) identified participants with varying CHL levels. Probability of correct answers, odds ratios, and standardized errors were used to identify the items that allow the classification of individuals among the latent classes. LCA resulted in a three-latent-class model predicting 39.4% of participants to be in the HIGH class, 43.3% in the MEDIUM class, and 17.3% in the LOW class. Eleven items (the CHLT-11-DKspa) meet the criteria to clearly separate participants with HIGH and LOW classes of CHL. Although the best model fit was a three-class solution, results showed a clear separation of individuals from HIGH versus LOW levels of CHL, but separation of those in the MEDIUM level was not as clear. The CHLT-11-DKspa is a shorter measure that may be relatively easy to use in a clinical encounter to identify Spanish-speaking patients with the poorest levels of CHL who may require additional support to understand medical instructions and care plans.

✉ Margarita Echeverri
mechever@xula.edu

David Anderson
danders2@xula.edu

Anna María Nápoles
anapoles@ucsf.edu

1 Educational Coordinator Health Disparities, Diversity and Cultural Competence, Center for Minority Health and Health Disparities Research and Education, Xavier University of Louisiana, College of Pharmacy, 1 Drexel Drive, New Orleans, LA 70125, USA

2 Department of Mathematics, Xavier University of Louisiana, 1 Drexel Drive, New Orleans, LA 70125, USA

3 Department of Medicine, Center for Aging in Diverse Communities, University of California San Francisco, Box 0856, 3333 California Street, Suite 335, San Francisco, CA 94118, USA

Introduction

The creation of profiles of health literacy by the National Center of Education Statistics marked a new era in health literacy research and initiatives [1]. Health literacy (HL) refers to the wide range of skills and competencies that people develop over their lifetimes to seek out, comprehend, evaluate, and use health-related information and concepts to make informed decisions about their health behaviors, treatments, and use of services to reduce health risks and improve health outcomes and quality of life [2]. Consistent with this definition, the Institute of Medicine called for the development of new HL measures that go beyond reading skills and include other literacy skills such as comprehension and basic mathematical

🖄 Springer

calculations [3]. Considering that HL is positively associated with quality of healthcare services, and negatively associated with healthcare disparities and healthcare costs, there is a need to develop instruments to assess disease-specific domains of health literacy, specifically for chronic diseases, such as cancer [4–6]. Although Latinos are the largest and fastest-growing minority population in the US and 13.0% of US residents 5 and older speak Spanish at home [7], a systematic review of research on HL among US Latinos found that instruments in Spanish were rarely used [8]. Considering that cancer is the leading cause of death among Latinos in US [9], there is a need to develop tools to assess cancer literacy among this population.

Based on widely accepted definitions of HL [2, 3], Echeverri and colleagues defined Cancer Health Literacy (CHL) as the "individual's capacity to seek out, comprehend, evaluate, and use basic information and services needed to make appropriate decisions regarding cancer prevention, diagnosis, and treatment" [25, p70]. Only a few CHL measures were available in Spanish at the time we conducted this study: the *Cancer Literacy Measure* [10], the *Cultural Cancer Screening Scale* [11], and the *Cervical Cancer Literacy Assessment Tool* [12]. However, these tools were limited to breast or cervical cancer and focused on measurement of beliefs and behaviors rather than cancer health literacy constructs. To address these limitations, Echeverri and colleagues translated to Spanish and culturally adapted the 30-item Cancer Health Literacy Tool (CHLT-30) [13]. The new instrument, called the Cancer Health Literacy Tool, Spanish version (CHLT-30-DKspa), assesses knowledge and skills needed to locate, understand, and use information from texts and materials (prose and document literacy) and to apply simple arithmetic operations (quantitative literacy) [25]. The aim of the present study was to assess the ability of the CHLT-30-DKspa to (1) identify Spanish-speaking individuals with low CHL and (2) ascertain which items in the tool best discriminate between low and high CHL level groups. The ability to discriminate between low and high literacy level groups is important because this would facilitate the identification of Spanish speakers at increased risk of poor health outcomes who might derive particular benefit from targeted support to understand complex health-related written materials and instructions and interventions to increase their CHL [14].

Methods

This study is a secondary analysis of data collected among 500 Spanish-speaking Latinos during 2014–2016. Survey collection procedures are published elsewhere [25]. The CHLT-30-DKspa version (included as Supplement 1) consists of 30 multiple-choice questions with 3 or 4 response options where only one option is correct. Total score ranges from 0 to 30,

based on a sum of the number of correct answers. The tool differs primarily from the original CHLT-30 in that a "don't know" (DK) response option was included for each item because in pilot testing of the instrument, respondents did not feel comfortable guessing and preferred to indicate they did not know [25]. DK answers were scored as incorrect answers.

Because CHL is a construct that is not directly observable, we employed latent class analysis (LCA) methods to identify participants with varying CHL levels. LCA is a method used widely to discover unobservable (latent) differences in a population [15]. It allows us to "uncover unobserved heterogeneity in a population and to find substantively meaningful groups of people that are similar in their responses" [16, p536] The probability of correct answers, odds ratios, and standardized errors were used to identify the items that permit classification of individuals among the latent classes.

Although we considered using factor analytic (FA) methods, another popular technique used in the analysis of unobserved (latent) variables, we decided against this method for several reasons. First, factor analysis is concerned with the structure of variables, whereas LCA is more concerned with individuals. Specifically, while FA groups items (variables) according to the degree to which they are correlated with each other, LCA groups individuals into mutually exclusive and similar classes based on their response patterns to the collection of items. Second, in FA, the observable variables are assumed to be continuous and the resulting latent variables (factors) are treated as continuous normally distributed variables, while in LCA, the observable variables are discrete (dichotomous, ordinal, or nominal) and the resulting latent variables (classes) are assumed to be discrete and their conditional distributions are assumed to be binomial or multinomial [17]. Thus, in summary, considering that previous studies have confirmed the unidimensional structure and internal consistency of the CHLT [13, 25], and that the tool yields categorical data scored using binomial distributions (correct = 1; wrong or don't know = 0), LCA was selected as the preferred analytic method to identify participants with varying levels (classes) of CHL.

Based on the literature, a conservative and reliable group of fit indices and criteria was selected to evaluate the goodness of fit of the possible LCA solutions (models): Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and the adjusted BIC (ABIC). The best fitting models are those with the best balance between a smaller value on these three statistics, a higher log likelihood, and a fewer number of parameters in the model [16, 18, 19]. Thus, selecting among models is a trade-off between a higher log likelihood value and a lower number of parameters [20]. In LCA, it is desirable to have a high degree of class homogeneity (low within-class variability) along with a high degree of class separation (high between-class variability) for a clear and optimal separation of participants into latent classes

[20]. Class homogeneity allows the identification of individuals within a given class (within-class) who are similar to each other with respect to item responses, while class separation allows the identification of individuals across different classes (between-class) who are dissimilar with respect to item responses.

Based on the literature, three criteria were used to select the CHLT-30-DKspa items that best classify participants among different groups or classes [21, 22]. Probability of correct answers (Pr2) above 0.70 or below 0.30 was used to indicate high within-class homogeneity. Odds ratios (ORs) above 5 or below 0.2 were used to indicate between-class separation, with ORs expressing the likelihood of members in one class answering correctly the item when compared to members in the reference class. Additionally, standardized errors (SEs) were also used to assess between-class separation. Following the Chebbychev's Inequality result that in any probability distribution at least 8/9 of the estimate values are within three SE of the mean, items with differences of three times the SE (in either direction) between two groups were considered strong indicators of clear separation between the classes. Full-sample LCA procedures were conducted using R [23] and the poLCA statistical package [24] to predict class membership.

Introducing covariates in the LCA has been found to be generally beneficial because it has a statistical "effect" in the change in odds of membership [15] and increases proper class assignment [24]. Therefore, education and region of origin, both previously shown to have a significant impact on CHLT-30-DKspa total score [25], were used as covariates. Our analysis of variance showed a significant age × education interaction, so age and age × education were also added to the model as covariates. Throughout this paper, we refer to specific items using the item number and label from the original English-language validation study [13].

Results

As intended by the stratified sampling design, a total of 500 self-identified Spanish-speaking Latinos, half women, completed the CHLT-30-DKspa. The mean CHLT-30-DKspa total score was 17.17 (range 0 to 30; SD 6.58). While three participants had the maximum score of 30 points, four had the minimum score of 0 points; thus, the entire range was observed.

Participant's age ranged from 25 to 86 years old (M = 46.16) with 42.0% being between 25 and 40, 30.4% between 41 and 55, and 27.6% being 56 or older. The majority of participants were born in Central American (78.6%) and South American (12.8%) countries, and only 43 participants (8.6%) were born in the USA (second-generation immigrants). Most of the foreign-born participants were from Honduras (39.0%), Mexico (14%), Colombia (6.6%), Cuba (6%), El Salvador (5.6%), Nicaragua (5.2%), and Guatemala (5.0%).

The majority of participants (62.8%) had a high school diploma or higher but a substantial percentage (22.2%) had only primary school or lower educational level. Although all participants chose Spanish as their primary language at home, more than one third reported that they speak (37.8%) or read (36.2%) some English. In general, most of the participants living in USA less than 10 years (24.4%) spoke and read only Spanish while those being in USA between 10 and 20 years (29.4%), between 21 and 30 years (20.6%), and more than 30 years (25.6%) spoke and read also English. More than half of participants (53.6%) did not have health insurance even though most of them were employed full-time (41.0%), part-time (10.4%), or self-employed (13.0%). Income was distributed as follows: 32.0% of participants had an annual household income of US$10,000 or less, 41.8% made between 10,001 and 40,000, and 26.2% had incomes higher than 40,000.

Model fit

LCA resulted in a final model with three latent classes. We ran models using one class, two classes, three classes, and four classes that attempted to balance better model fit (higher log likelihood) with parsimony (fewer parameters in the model). We selected the three-class model because it has the minimum BIC value (16,946.9) of all models tested (better model fit) and is the simpler model [18, 19, 21]. The three classes identified by the final model are referred to as HIGH, MEDIUM, and LOW cancer health literacy, with a distribution of 39.4, 43.3, and 17.3% of participants, respectively. However, specific items distinguished between the classes in different ways (Fig. 1). For example Q2-*Next Pill* separated the LOW scoring class from the other two classes, while Q5-*Oral Cancer* separated the HIGH scoring class from the other two classes, and Q01-*High Calorie* did not separate any of them.

Homogeneity within Classes

Applying the high homogeneity within a class criterion defined earlier (Pr2s above 0.70 or below 0.30), only three items had high within-class homogeneity for all three classes (Q11-*Body temperature*, Q14-*Efficacy*, and Q26-*Complication rate*). Looking at each class separately,

- Eight items could not be used to adequately classify respondents in the HIGH scoring class because their Pr2s were lower than 0.7 (Q1-*High Calorie*, Q8-*Palliative care*, Q13-*Direction*, Q15-*Tumor spread*, Q16-*Generic drugs*, Q17-*Survival rate*, Q19-*Smoking risk*, and Q23-*Metastasized*).

Fig. 1 Probability of correct answer (*N* = 500)
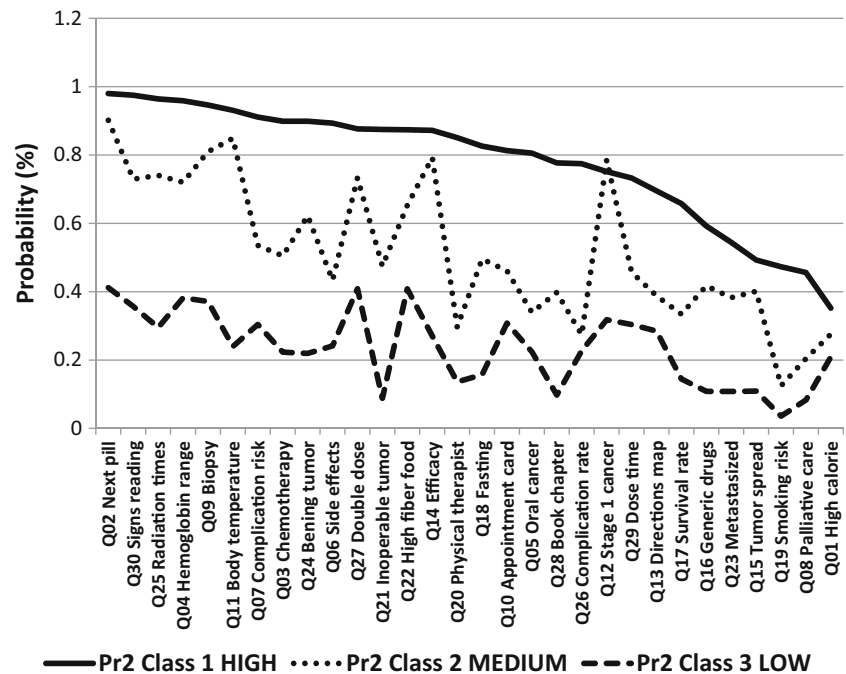
- Seventeen items could not be used to adequately classify respondents in the MEDIUM scoring class because their Pr2s were between 0.7 and 0.3 (Q3-*Chemotherapy*, Q5-*Oral cancer*, Q6-*Side effects*, Q7-*Risk of complication*, Q10-*Appointment location*, Q13-*Direction*, Q15-*Tumor spread*, Q16-*Generic drugs*, Q17-*Survival rate*, Q18-*Fasting*, Q20-*Physical therapist*, Q21-*Inoperable tumor*, Q22-*High fiber food*, Q23-*Metastasized*, Q24-*Benign tumor*, Q28-*Book chapter*, and Q29-*Dose time*).
- Eleven items could not be used to adequately classify respondents in the LOW scoring class because their Pr2s were higher than 0.3 (Q2-*Next pill*, Q4-*Hemoglobin range*, Q7-*Risk of complication*, Q9-*Biopsy*, Q10-*Appointment location*, Q12-*Stage 1 cancer*, Q22-*High fiber food*, Q25-*Radiation treatment*, Q27-*Double dose*, Q29-*Dose time*, and Q30-*Map reading*)

Separation among Classes

As indicated earlier, two different criteria were used to identify high separation among classes: three times the differences in standard errors of Pr2 (SE criterion) and the estimated odds ratios (ORs) or likelihood of members in one class answering an item correctly, compared to members in the reference class. When using the SE criterion, six items differentiated between the three classes (Q18-*Fasting*, Q21-*Inoperable tumor*, Q24-*Benign tumor*, Q25-*Radiation treatment*, Q28-*Book chapter*, and Q30-*Map reading*). When using the OR criterion, six items differentiated between the three classes (Q02-*Next pill*, Q18-*Fasting*, Q21-*Inoperable tumor*, Q25-*Radiation*

treatment, Q28-*Book chapter*, and Q30-*Map reading*). Deleting the Q02-*Next pill* and Q24-*Benign tumor* items because they did not meet both the SE and OR criteria, the remaining five items met both criteria indicating high separation among the three classes.

When we combined both separation criteria and examined each possible pair of classes independently, the results indicated that

- All items except one (Q1-*High calorie*) clearly separated HIGH and LOW classes
- Fourteen items clearly separated the HIGH and MEDIUM classes (Q3-*Chemotherapy*, Q4-*Hemoglobin range*, Q5-*Oral cancer*, Q6-*Side effects*, Q7-*Risk of complication*, Q10-*Appointment location*, Q18-*Fasting*, Q19-*Smoking risk*, Q20-*Physical therapist*, Q21-*Inoperable tumor*, Q25-*Radiation treatment*, Q26-*Complication rate*, Q28-*Book chapter*, and Q30-*Map reading*)
- Fourteen items clearly separated the MEDIUM and LOW classes (Q2-*Next pill*, Q9-*Biopsy*, Q11-*Body temperature*, Q12-*Stage 1 cancer*, Q14-*Efficacy*, Q15-*Tumor spread*, Q16-*Generic drugs*, Q18-*Fasting*, Q21-*Inoperable tumor*, Q23-*Metastasized*, Q24-*Benign tumor*, Q25-*Radiation treatment*, Q28-*Book chapter*, and Q30-*Map reading*)

Combined Consideration of within Class Homogeneity and Separation between Classes

We combined the three criteria, one measure of within-class homogeneity (probability of correct answers above 0.70 or

below 0.30), and two measures of separation between classes (odds ratios >5.0 or <0.2 and items with differences of three times the SE in either direction) to identify the best items that clearly classify individuals into the three classes. In summary, we found that there was no one single item meeting all three criteria that clearly allows for the separation of participants into these three classes with high within-class homogeneity. Table 1 summarizes the results of the analyses of within-class homogeneity and separation among the three classes for each of the 30 items in the CHLT-30-DKspa.

Although five items were able to clearly separate the classes (Q18-*Fasting*, Q21-*Inoperable tumor*, Q25-*Radiation treatment*, Q28-*Book chapter*, and Q30-*Map reading*), none of these items were able to statistically define homogenous classes. While Q18-*Fasting*, Q21-*Inoperable tumor*, and Q28-*Book chapter* did not perform well for the MEDIUM class, Q25-*Radiation treatment* and Q30-*Map reading* did not perform well for the LOW class. Similarly, only three items had high within-class homogeneity for all three classes (Q11-*Body temperature*, Q14-*Efficacy*, and Q26-*Complication rate*), but none of them could clearly separate the three classes: Q11-*Body temperature* and Q14-*Efficacy* could not separate the HIGH class from the MEDIUM class while Q26-*Complication rate* could not clearly separate the MEDIUM class from the LOW class.

Four items clearly separated the HIGH from the MEDIUM class and also had high within-class homogeneity (Q4-*Hemoglobin range*, Q25-*Radiation treatment*, Q26-*Complication rate*, and Q30-*Map reading*). Three of the four items were related to numeracy skills. Q4-*Hemoglobin range* required participants to review data and make a decision, Q25-*Radiation treatment* required participants to make a simple mathematic operation (multiply), and Q26-*Complication rate* required participants to understand and calculate a percentage.

Two items clearly separated the MEDIUM from the LOW class and also had high within-class homogeneity (Q11-*Body temperature* and Q14-*Efficacy*). While Q11-*Body temperature* required respondents to review data about fever classification and make a decision if the person should go or not go to the doctor (numeracy skills), the Q14-*Efficacy* assessed respondents' ability to interpret accurately the meaning of the word "efficacy."

Eleven items clearly separated the HIGH and LOW classes (high separation among classes and high homogeneity between classes). These items met the three criteria and could be used to identify individuals with HIGH vs. LOW cancer health literacy skills (Q3-*Chemotherapy*, Q5-*Oral cancer*, Q6-*Side effects*, Q11-*Body temperature*, Q14-*Efficacy*, Q18-*Fasting*, Q20-*Physical therapist*, Q21- *Inoperable tumor*, Q24-*Benign tumor*, Q26-*Complication rate*, and Q28-*Book chapter*). The 11 items in this shorter tool, referred to now as the CHLT-11-DKspa, assess a comprehensive set of literacy skills. Four questions focus on numeracy skills (Q5-*Oral cancer*, Q11-*Body temperature*, Q18-*Fasting*, and Q26-*Complication rate*), and seven questions focus on reading comprehension, vocabulary, and analysis skills (Q3-*Chemotherapy*, Q6-*Side effects*, Q14-*Efficacy*, Q20-*Physical therapist*, Q21- *Inoperable tumor*, Q24-*Benign tumor*, and Q28-*Book chapter*). There was a significant difference in percent correct answers between high predicted class and low predicted class individuals for all questions except Q1-*High Calorie*. Items with significant differences for all the three classes are shown on Table 1.

Discussion

We used LCA to understand the heterogeneity of Spanish-speaking participants' cancer health literacy levels based on their responses to a new Spanish-language measure of CHL. The goal was to identify groups of participants who were similar in levels while different from other literacy level groups. Examining fit indices for models with varying numbers of latent classes, the three-class solution demonstrated the best fit, with one class clearly scoring higher in the CHL items, another scoring lower, and one class in the middle. Results indicate that both the CHLT-30-DKspa and the CHLT-11-DKspa can be used to assess CHL in individuals who are primarily Spanish-speaking. Healthcare providers and organizations that want to prioritize interventions for Spanish-speaking patients with the poorest levels of CHL may benefit especially from the shorter version CHLT-11-DKspa to identify these individuals.

In the original analyses of the English version of the CHLT-30, it was concluded that a two-latent class model allowed for the separation of participants into HIGH and LOW cancer health literacy levels [13]. They found six items that clearly separated the HIGH and LOW classes and referred to this subset of items as the CHLT-6 (Q4-*Hemoglobin range*, Q9-*Biopsy*, Q12-*Stage 1 cancer*, Q20-*Physical Therapist*, Q21-*Inoperable Tumor*, and Q25-*Radiation Treatment*). However, our findings differed and provide support for a three-latent class model that classifies participants into HIGH, MEDIUM, and LOW cancer health literacy levels. One should exercise caution when comparing results from our study with the original CHLT-30 validation study because of important differences between the studies. In our study, participants had lower levels of education (59.4% had a high school education or less compared to 30% in the original study). The tool used in this study included a DK response option for all items that was not an option in the original study; this eliminated inflation of scores based on guesses [25]. Finally, we included education, region of origin, age, and age × education as covariates in the analyses, while these

Table 1 Summary analysis of class homogeneity and separation for the Spanish version of the Cancer Health Literacy Test (CHLT-30-DKspa), $N = 500$

Item[a,b]	High homogeneity within classes[c]			High separation among classes[c]					
	Probability of Correct Answers (PR2)[d]			HIGH vs MEDIUM		MEDIUM vs LOW		HIGH vs LOW	
	HIGH	MEDIUM	LOW	3SE[e]	OR[f]	3SE[e]	OR[f]	3SE[e]	OR[f]
Q01 High calorie	N	Y	Y	N	N	N	N	N	N
Q02 Next pill	Y	Y	N	N	Y	Y	Y	Y	Y
Q03 Chemotherapy[g]	Y	N	Y	Y	Y	N	N	Y	Y
Q04 Hemoglobin range[g]	Y	Y	N	Y	Y	N	N	Y	Y
Q05 Oral cancer	Y	N	Y	Y	Y	N	N	Y	Y
Q06 Side effects	Y	N	Y	Y	Y	N	N	Y	Y
Q07 Risk of complication	Y	N	**N**	Y	Y	N	N	Y	Y
Q08 Palliative care	N	Y	Y	Y	N	N	N	Y	Y
Q09 Biopsy	Y	Y	N	N	N	Y	Y	Y	Y
Q10 Appointment location	Y	N	N	Y	Y	N	N	Y	Y
Q11 Body temperature	Y	Y	Y	N	N	Y	Y	Y	Y
Q12 Stage 1 cancer	Y	Y	N	N	N	Y	Y	Y	Y
Q13 Direction	**N**	N	Y	Y	N	N	N	Y	Y
Q14 Efficacy	Y	Y	Y	N	N	Y	Y	Y	Y
Q15 Tumor spread	N	N	Y	N	N	Y	Y	Y	Y
Q16 Generic drugs[g]	N	N	Y	N	N	Y	Y	Y	Y
Q17 Survival rate	N	N	Y	Y	N	N	N	Y	Y
Q18 Fasting[g]	Y	N	Y	Y	Y	Y	Y	Y	Y
Q19 Smoking risk	N	Y	Y	Y	Y	N	N	Y	Y
Q20 Physical therapist	Y	**N**	Y	Y	Y	N	N	Y	Y
Q21 Inoperable tumor[g]	Y	N	Y	Y	Y	Y	Y	Y	Y
Q22 High fiber food[g]	Y	N	N	Y	N	N	N	Y	Y
Q23 Metastasized	N	N	Y	N	N	Y	Y	Y	Y
Q24 Benign tumor[g]	Y	N	Y	Y	N	Y	Y	Y	Y
Q25 Radiation treatment[g]	Y	Y	**N**	Y	Y	Y	Y	Y	Y
Q26 Complication rate	Y	Y	Y	Y	Y	N	N	Y	Y
Q27 Double dose[g]	Y	Y	N	N	N	N	N	Y	Y
Q28 Book chapter[g]	Y	N	Y	Y	Y	Y	Y	Y	Y
Q29 Dose time	Y	N	**N**	Y	N	N	N	Y	Y
Q30 Map reading[g]	Y	Y	N	Y	Y	Y	Y	Y	Y

CHLT, Cancer Health Literacy Test, *DKspa*, Don't Know Spanish version, *PR2* Probability of Correct Answers, *OR* odd ratios; *SE*, standard errors, *Y* yes, *N* no

[a] Stem for all of the items = Question # + abbreviated description used in Dumenci et al [13]

[b] The 11 items in bold are a subset of the CHLT-30-DKspa and are referred as the CHLT-11-DKspa

[c] A Y for YES and N for NO are used to indicate whether or not the criterion is met

[d] Selection criterion: Pr2s above 0.70 or below 0.30

[e] Selection criterion: positive differences in three times standard errors (3SE) of Pr2 among the classes

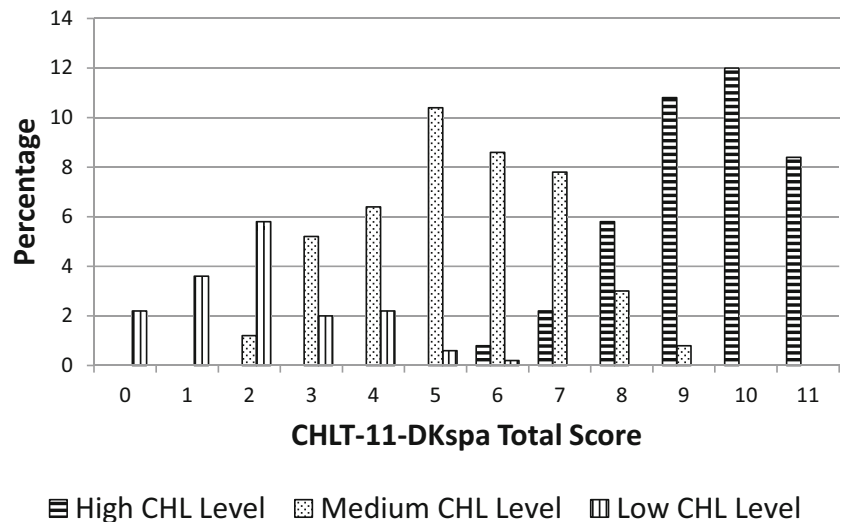[f] Selection criterion: odd ratios (ORs) above 5 or below 0.2

[g] Pr2s for all three groups are significantly different from one another ($p < 0.05$). Bonferroni Correction used for multiple comparisons

covariates were not included in the original CHLT-30 study.

In our study, it was easy to identify participants at the extremes, that is, participants with HIGH or LOW level of CHL. However, participants in the middle were more difficult to classify (Fig. 2); nonetheless, this group can be identified as anyone not belonging to the HIGH or LOW groups. Our results identified 11 items that clearly

Fig. 2 Cancer health literacy (CHLT-11-DKspa) scores by predicted class ($N = 500$)

Conclusions and Recommendations

In conclusion, we were able to identify an 11-item short-form measure, the CHLT-11-DKspa, which can be used in clinical practice to differentiate between Spanish-speaking Latinos with high and low levels of CHL. This measure could help identify a subset of patients who may require health-related instructions in oral or video, rather than written formats and literacy improvement interventions. Although the CHLT-30-DKspa can be used to measure a Spanish-speaking individual's CHL, no single item performed well enough to identify and classify, in a statistically rigorous way, individuals with HIGH, MEDIUM, and LOW cancer health literacy levels. Most likely, this is due to the comprehensive set of skills that make up CHL that cannot be captured by a single item measure. As stated in the Institute of Medicine's report, "If there are no data, there is no problem. If there is no problem, there is no action" [14, p93]. Although health literacy can seem invisible (latent), it is present among patients seeking care, affects patients' understanding and compliance to treatments, and is becoming increasingly critical to assess in the context of the growing complexity of health care services provided by organizations [26].

separate the HIGH class from the LOW class, referred to as the CHLT-11-DKspa (Table 1, items in bold). The CHLT-11-DKspa offers a brief Spanish-language measure that can be used in clinical settings to identify individuals with low CHL, who might benefit the most from interventions to improve their health literacy and associated outcomes. Note that both questions which clearly separated the MEDIUM class from the LOW class (Q11-*Body temperature* and Q14-*Efficacy*) are among these 11 items. The items assess a comprehensive set of CHL characteristics, including numeracy, reading comprehension, vocabulary, and analysis skills.

Scores on the CHLT-11-DKspa could be used to categorize the CHL of Spanish-speaking patients (Fig. 2): those with CHLT-11-DKspa total scores between 0 and 2 could be considered in the LOW class, 5–6 scores in the MEDIUM class, and 9–11 scores in the HIGH class. Being conservative, we recommend classifying in the lower category those patients in-between two classes, that is, 3–4 scores in the LOW and 7–8 scores in the MEDIUM class.

Important limitations of this study need to be considered. First, the CHLT-30-DKspa is a new tool that was validated in a pilot project conducted among Spanish-speaking Latinos and includes 30 items that measure prose, document, and quantitative literacy [25]. Considering that only 11 items demonstrated utility for classifying individuals with HIGH and LOW levels of CHL; including additional items might improve the tool and its ability to discriminate between literacy levels. Second, although LCA results suggested a three-class model; the fact that no item was able to clearly identify and classify individuals with MEDIUM cancer health literacy levels suggests that a third class was not empirically supported.

References

1. Kutner M, Greenberg E, Jin Y, Paulsen C (2006). The health literacy of America's adults: results from the 2003 National Assessment of Adult Literacy (NCES 2006–483). U.S. Department of Education. Washington, DC: National Center for Education Statistics (NCES)

2. Zarcadoolas C, Greer DS, Pleasant A (2006) Advancing health literacy: a framework for understanding and action. Wiley, San Francisco

3. Institute of Medicine (2004) Health literacy: a prescription to end confusion. National Academies Press, Washington, DC

4. Gazmararian JA, Williams MV, Peel J, Baker DW (2003) Health literacy and knowledge of chronic disease. Patient Educ Couns 51(3):267–275

5. Shaw SJ, Armin J, Torres CH et al (2012) Chronic disease self-management and health literacy in four ethnic groups. J Health Commun 17(sup3):67–81

6. Heijmans M, Waverijn G, Rademakers J et al (2015) Functional, communicative and critical health literacy of chronic disease patients and their importance for self-management. Patient Educ Couns 98(1):41–48

7. U.S. Census Bureau (2015) Facts for features: Hispanic Heritage Month. Release Number CB15-FF.18, September 14, 2015. From http://www.census.gov/newsroom/facts-for-features/2015/cb15-ff18.html. Accessed 19 July 2017

8. Koskan A, Friedman DB, Hilfinger M, DeAnne K (2010) Health literacy among Hispanics: a systematic research review (1992–2008). Hispanic Health Care Int 8(2):65–76. doi:10.1891/1540-4153.8.2.65

9. American Cancer Society (2012) Cancer facts & figures for Hispanics/Latinos 2012–2014. American Cancer Society, Atlanta, 40 p

10. Buki LP, Yee WK, Weiterschan K, Lehardy E (2016) Essential qualitative inquiry in the development of a cancer literacy measure for immigrant women. Qual Health Res 26(5):640–648

11. Betancourt H, Flynn PM, Riggs M, Garberoglio C (2010) A cultural research approach to instrument development: the case of breast and cervical cancer screening among Latino and Anglo women. Adv Access Publ 25(6):991–1007

12. William KP, Templin TN (2013) Bringing the real world to psychometric evaluation of cervical cancer literacy assessments with black, Latina, and Arab women in real-world settings. J Cancer Educ 28: 738–743

13. Dumenci L, Matsumaya RK, Riddle DL et al (2014) Measurement of cancer health literacy and identification of patients with limited cancer health literacy. J Health Commun Int Perspect 19(sup2): 205–224

14. Institute of Medicine Roundtable on Health Literacy (2009) Measures of Health Literacy: Workshop Summary. Washington (DC): National Academies Press; 2009. Available from: https://www.ncbi.nlm.nih.gov/books/NBK45384/, doi:10.17226/12690

15. Collins LM, Lanza ST (2010) Latent class and latent transition analysis: with applications in the social, behavioral, and health sciences. Wiley, New York, 330 p

16. McCutcheone AL (1987) Latent class analysis, Sage University paper series on quantitative applications in the social sciences, no 07–064. Sage, Newbury Park

17. Nylund KL, Asparouhov T, Muthén BO (2007) Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equ Model 14(4):535–569

18. Dean M, Raftery AE (2010) Latent class analysis variable selection. Ann Inst Stat Math 62:11–35

19. Linzer DA, Lewis JB (2011) poLCA: an R package for polytomous variable latent class analysis. J Stat Soft 42(10):1–29

20. Muthén LK, Muthén BO (2012) Mplus user's guide, 7th edn. Muthén & Muthén, Los Angeles

21. Nylund-Gibson K, Hart SR (2014: Chapter 21) Latent class analysis in prevention science. In: Sloboda Z, Petras H (eds) Defining Prevention Science. Springer, New York, pp 493–512

22. Masyn K (2013). Applied Latent Class Analysis: A Workshop. Texas Tech University 2013. From http://www.depts.ttu.edu/immap/Masyn_LCAWorkshop_Dec2013_TTU.pdf. Accessed 19 July 2017

23. R Development Core Team (2011) R: a language and environment for statistical computing. Vienna, Austria: The R Foundation for Statistical Computing, 2011

24. Wurpts IC, Geiser C (2014) Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. Front Psychol 5:920. doi:10.3389/fpsyg.2014.00920

25. Echeverri M, Anderson D, Nápoles AM (2016) Cancer Health Literacy Test-30-Spanish (CHLT-30-DKspa), a new Spanish-language version of the Cancer Health Literacy Test (CHLT-30) for Spanish-speaking Latinos. J Health Commun Int Perspect 21(1):69–78. doi:10.1080/10810730.2015.1131777

26. Center for Health Care Strategies (2013) Health Literacy Fact Sheets. From http://www.chcs.org/media/CHCS_Health_Literacy_Fact_Sheets_2013.pdf. Accessed 19 July 2017