PROJECT MUSE®

Validity of a Computational Linguistics-Derived Automated Health Literacy Measure Across Race/Ethnicity: Findings from The ECLIPPSE Project

Dean Schillinger, Renu Balyan, Scott Crossley, Danielle McNamara, Andrew Karter

Journal of Health Care for the Poor and Underserved, Volume 32, Number 2, May 2021 Supplement, pp. 347-365 (Article)

Published by Johns Hopkins University Press

DOI: <https://doi.org/10.1353/hpu.2021.0067>

➔ *For additional information about this article*

<https://muse.jhu.edu/article/789674>

Validity of a Computational Linguistics-Derived Automated Health Literacy Measure Across Race/Ethnicity: Findings from The ECLIPPSE Project

Dean Schillinger, MD
Renu Balyan, PhD
Scott Crossley, PhD
Danielle McNamara, PhD
Andrew Karter, PhD

Abstract: Limited health literacy (HL) partially mediates health disparities. Measurement constraints, including lack of validity assessment across racial/ethnic groups and administration challenges, have undermined the field and impeded scaling of HL interventions. We employed computational linguistics to develop an automated and novel HL measure, analyzing >300,000 messages sent by >9,000 diabetes patients via a patient portal to create a Literacy Profiles. We carried out stratified analyses among White/non-Hispanics, Black/non-Hispanics, Hispanics, and Asian/Pacific Islanders to determine if the Literacy Profile has comparable criterion and predictive validities. We discovered that criterion validity was consistently high across all groups (c-statistics 0.82–0.89). We observed consistent relationships across racial/ethnic groups between HL and outcomes, including communication, adherence, hypoglycemia, diabetes control, and ED utilization. While concerns have arisen regarding bias in AI, the automated Literacy Profile appears sufficiently valid across race/ethnicity, enabling HL measurement at a scale that could improve clinical care and population health among diverse populations.

Key words: Health literacy, communication, validation study, artificial intelligence, machine learning, diabetes, health disparities, computational linguistics.

Limited health literacy (HL) is associated with untoward and costly health outcomes that contribute to health disparities.¹ Limited HL has been found to be more common among minority groups in the U.S., including among non-Hispanic Blacks,

DEAN SCHILLINGER is affiliated with the University of California San Francisco and the Division of Research, Northern California Kaiser Permanente. **RENU BALYAN** is affiliated with the State University of New York Old Westbury and Arizona State University. **SCOTT CROSSLEY** is affiliated with Georgia State University. **DANIELLE MCNAMARA** is affiliated with Arizona State University. **ANDREW KARTER** is affiliated with the Division of Research, Northern California Kaiser Permanente. Please address all correspondence to: Dean Schillinger, Division of General Internal Medicine, Building 10, Ward 13, San Francisco General Hospital, 1001 Potrero Ave, San Francisco, CA 94110; Email: Dean.Schillinger@ucsf.edu.

Hispanics, and Asian/Pacific Islander subgroups.²⁻³ Poor communication exchange is an important mediator in the relationship between limited HL and health outcomes.⁴⁻⁶ Patient-physician communication is a fundamental pillar of care that influences patient satisfaction and health outcomes,⁷ notably so in diabetes mellitus.⁸ Limited HL impedes physician-patient communication, as well as imparts a barrier to patients' learning and understanding across numerous communication domains.⁹⁻¹² Health literacy interventions can improve outcomes among diverse populations and, in some cases, have been shown to reduce disparities.¹³ However, measurement constraints, including the time required to administer HL instruments and lack of validation across racial/ethnic subgroups, have limited internal and external validity and impeded scaling potentially effective interventions.

How best to measure patient HL—and whether or not HL measures are detecting true differences in capacities and skills in marginalized populations—can be problematic and controversial. Despite the importance of HL as a contributor to health disparities by race/ethnicity, to our knowledge, no study has compared the performance of a HL measure across the most common racial/ethnic groups in the U.S., either with respect to criterion validity or predictive validity.

Electronic patient portals are an increasingly popular channel for patients and providers to communicate via secure messaging, offering the possibility of employing computational linguistics to estimate patient HL. While individuals of minority status and with limited HL have historically been shown to be less likely to use the patient portal, engagement rates are steadily rising and disparities in portal access are rapidly narrowing.¹⁴ Because “big data”—in this case data derived from patients' written secure messages sent via patient portals—are increasingly available, we recently employed computational linguistics and machine learning to develop a novel HL measure, analyzing language from ~300,000 secure messages sent by ~9,000 ethnically diverse patients with diabetes via an integrated health system's portal. This artificial intelligence (AI) approach harnesses big linguistic data to enable the automated generation of a HL measure which we called the Literacy Profile. This automated process led to the creation of a Literacy Profile with a high level of accuracy against a gold standard.¹⁵ Furthermore, the Literacy Profile was associated with patterns that mirror previous research in terms of its relationship with patient sociodemographics, ratings of physician communication, and a range of diabetes-related health outcomes. Thus, the Literacy Profile provides a novel health IT tool that could be harnessed to enable tailored communication support and other targeted interventions with the potential to reduce HL-related disparities.

Given the fact that few, if any, established HL measures have been assessed with respect to their cross-cultural validity, and given the concerns that have recently arisen regarding bias in applying AI technology in health care settings¹⁶ (including concerns over automating and propagating existing biases¹⁷) we present the first study to validate a HL measure across the most common racial/ethnic subgroups in the U.S. We previously have found that the automated Literacy Profile is strongly correlated with key demographic variables, such as race/ethnicity and educational attainment. Furthermore, we have shown that the Literacy Profile has high criterion validity with respect to a reference standard of health literacy, as well as significant predictive validity with respect to health outcomes.¹⁸⁻¹⁹ Our current objective is to determine whether the

Literacy Profile has sufficient validity across races/ethnicities to justify applying and scaling it in practice at a scale that could improve clinical care and population health among diverse populations.

Methods

Data sources and participants. This study is part of the NLM-funded ECLIPPSE Project (Employing Computational Linguistics to Improve Patient-Physician email Exchange), and a detailed review of methods used to develop the Literacy Profile and results and implications of this work can be found in prior reports.^{14–15,18–19} Briefly, our sampling frame included over one million secure messages (SMs) exchanged between diabetes patients and providers between 2006 and 2015 at Kaiser Permanente Northern California (KPNC), a fully integrated health system that provides care to ~4.4 million patients and supports a well-developed and mature patient portal (kp.org). We selected diabetes patients for our study because more than 30 million U.S. adults are living with diabetes,²⁰ and one quarter to one third of them has limited HL skills.^{4,21} In addition, diabetes is a chronic disease in which (a) the quality of communication has been shown to influence health outcomes,⁸ (b) patient portal use is commonly used to enable inter-visit communication, and (c) engagement in secure messaging has been shown to be associated with salutary outcomes.²²

The ECLIPPSE project derived its sample from over 20,000 patients who completed a 2005–2006 survey as part of the NIH-funded Diabetes Study of Northern California (DISTANCE).^{8,23–24} DISTANCE oversampled minority subgroups to assess the role of socio-demographic factors on quality and outcomes of care. The average age of the DISTANCE study population at the time was 56.8 (± 10); 54.3% were male; and 18.4% Hispanic, 16.9% Black/non-Hispanic, 22.8% White/non-Hispanic, and 30.8% were Asian/Pacific Islander and 11.0% Other. Race/ethnicity was measured based on patient self-report as previously described.²⁴ Variables were collected from questionnaires completed via telephone, on-line, or paper and pencil (62% response rate). Details of the DISTANCE Study have been reported previously.^{22,24}

We first extracted all SMs ($N=1,050,577$) exchanged from 01/01/2006 through 12/31/2015 between DISTANCE diabetes patients and all clinicians from KPNC's patient portal. Members have been able to use the patient portal, kp.org, since 1999, with the SM feature enabled since 2005; the patient portal was only available in English during the study period. For the current analyses, only those SMs that a patient sent to his or her primary care physician were included. We excluded all SMs from patients who did not have matching DISTANCE survey data; were written in a language other than English; or were written by proxy caregivers (determined by the KP.org proxy check-box or by a validated NLP algorithm²⁵). The final ECLIPPSE dataset used for the assessment of validity of the Literacy Profile by race/ethnicity consisted of >300,000 SMs sent by 9,527 patients to their primary care physicians.

This study was approved by the KPNC and UCSF Institutional Review Boards. All analyses involved secondary data and all data were housed on a password-protected secure KPNC server that could only be accessed by authorized researchers.

Health literacy reference standard. We generated HL scores based on expert rat-

ings of the quality of patients' SMs. These ratings were carried out on a subset of the ECLIPSE sample comprising aggregated secure messages written by 512 patients, purposively sampled to represent a balance of self-reported HL, as well as a range of age, race/ethnicity and socio-economic status.¹⁵ A HL scoring rubric was used to assess holistically the HL of the patients based on the content of their SMs, adapting an established rubric used to score the writing abilities of high school students entering college.^{15,26} An ordinal scale ranging from 1 to 6 assessed the extent to which patients' SMs demonstrated mastery of written English, organization, and focus, and a varied, accurate, and appropriate health vocabulary to enable clear access to the health-related content and ideas the patient wanted to express to their physician.¹⁵ Because of limited relevance to the construct of HL, we removed parts of the rubric related to length, developing point of views, and discourse-related elements important in argumentative writing including the use of examples, reason, and evidence. Two raters with advanced degrees in linguistics and experience in HL research were trained twice on 25 separate, aggregated SMs not included in the 512 messages used in the final analysis. After reaching a satisfactory inter-rater reliability measured using the weighted Kappa ($>.70$), raters independently scored the 512 messages. Secure messages were categorized into two groups: limited HL (scores <4 , $n = 200$) and adequate health HL (scores ≥ 4 , $n = 312$).

Natural language processing (NLP) tools and the Literacy Profile. The linguistic features we examined were derived from the patients' SMs using several NLP tools that measure different facets of language.¹⁵ Prior research that has indicated that lexical features related to word choice, discourse features, and sentence structure are strong predictors of writing quality.^{26–28} To capture these features, we used three NLP tools that derive linguistic features related to lexical sophistication, text cohesion, and syntactic complexity, which we briefly describe here. These included: (1) the Tool for the Automatic Analysis of Lexical Sophistication (TAALES),^{29–30} a computational tool that incorporates over 100 classic and newly developed indices of lexical sophistication. These indices measure word frequency, lexical range, n -gram frequency and proportion, academic words and phrases, word information, lexical and phrasal sophistication, and age of exposure. The tool also reports on a number of word information and psycholinguistic scores derived from databases which calculate number of word associations per word and the number of phonological neighbors a word has (i.e., how many words sound similar to the word in question) and lexical decision response times for words (i.e., how long does it take to decide a word is a word versus a non-word). (2) The Tool for the Automatic Analysis of Cohesion (TAACO),^{31–32} which incorporates a number of classic and recently developed indices related to text cohesion. This tool has features for content and function words and provides linguistic counts for both sentence and paragraph markers of cohesion. It calculates sentence and paragraph overlap indices (i.e., local and global cohesion) and a variety of connective indices. For example, argument overlap is a count of arguments that are shared between sentences and paragraphs. (3) The Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC), which measures large and fine-grained clausal and phrasal indices of syntactic complexity and use-based frequency/contingency indices of syntactic sophistication.^{33–34} At the clausal level, TAASSC measure features such as the number of passive auxiliary verbs and adjective complements per clause. At the phrasal level,

TAASSC calculates features such as determiners per nominal phrase and dependents per nominal subject. In addition, TAASSC reports on features related to verb argument criteria (VAC) including the frequency of VAC and the attested lemmas per VAC as found in reference corpora taken from sections (e.g., magazine or newspaper) of the Corpus of Contemporary American English.

Using the patients' SMs, we applied NLP and machine learning techniques to develop a Literacy Profile for predicting patients' expert-rated HL. A set of eight linguistic indices, including lexical decision latencies, age of exposure, word naming response times, academic word lists, bigrams association strength and dependency structures were used as independent variables to predict human ratings of HL from the purposively sampled subset of 512 SMs described above. Additional details related to the development and experimental design of the Literacy Profile have been previously reported.¹⁵

Assessing performance of the Literacy Profile by race/ethnicity. We assessed the performance of the Literacy Profile using the supervised machine learning classification algorithm known as support vector machine (SVM). In a supervised machine learning model, the algorithm learns from a labeled dataset, providing an answer key that the algorithm can use to classify unseen data and evaluate its accuracy. Using a randomly allocated split sample approach, we then measured the discriminatory performance results across the entire sample using the c-statistic (area under the receiver operator [ROC] curves). We previously found that the Literacy Profile performed well in its ability to discriminate between those with limited vs. adequate expert-rated HL, with a c-statistic of 0.87.^{15,18} For the current study, we carried out a set of stratified analyses, separately measuring discriminatory performance for each common racial/ethnic group (excluding the "Other/Multiethnic" category, N=1,053) to enable cross-group validity comparisons, as well as validity comparisons with the overall sample.

Assessing predictive validity for the Literacy Profile by race/ethnicity. We then examined associations between the HL classifications generated by the Literacy Profile and known health outcome correlates of HL among the total sample (N=9,527).¹⁸ Outcomes included suboptimal patient-provider communication,⁹⁻¹² using an adapted version of the most HL-relevant item from the 4-item CAHPS survey:⁹ "In the last one year, how often have your physician and health care providers explained things in a way that you could understand?" We also examined the extent to which the Literacy Profile was associated with four diabetes-related outcomes previously found to be associated with HL. These included poor adherence to cardio-metabolic medications based on continuous medication gaps (CMG),³⁵⁻³⁶ a validated measure based on percent time with insufficient medication supply; poor diabetes (glycemic) control (HbA1c $\geq 9\%$); and ≥ 1 clinically relevant hypoglycemic episode (a patient safety event related to diabetes treatment and self-management skills).³⁷ To be consistent with the prior literature on HL and diabetes outcomes, HbA1c reflected the value collected adjacent to the first SM sent, while CMG and hypoglycemia were measured the year before the first SM. The occurrence of one or more hypoglycemia-related ED visits or hospitalizations in the year prior was based on a validated algorithm that uses specific diagnostic codes.³⁸ Finally, we explored the relationship between the Literacy Profile and emergency room utilization data 12 months prior to the first SM date. For all analyses, we examined bivariate associations using a two-sided p-value at the .05 level. Categorical variables

such as adherence, HbA1c levels, hypoglycemia, and ED visits were analyzed using chi-square analysis. In our prior research, we found that the Literacy Profile had predictive validity in the overall sample: those patients whose Literacy Profiles were indicative of limited HL reported worse communication scores (i.e., that their physician and health care providers were less likely to explain things in a way that they could understand), worse medication adherence, higher rates of poor diabetes control, higher prevalence of severe hypoglycemic events, and higher ED use¹⁸ ($p < .05$ for all associations). For the current study, we carried out a set of stratified analyses, separately measuring associations between the Literacy Profile and each outcome for each racial/ethnic group to enable cross-group comparisons with respect to predictive validity, as well as comparisons with the overall sample. Because our power would be more limited due to the smaller samples sizes contained in each racial/ethnic subgroup, our interest was in (a) determining if the direction and extent of the associations observed in the entire sample were maintained in the stratified analyses, as well as (b) establishing whether the statistical significance of these associations observed in the entire sample were also achieved. Finally, we looked for interactions between HL, race/ethnicity, and health outcomes, using a cutoff of $< .20$ as representing a potentially significant interaction.

Results

With respect to criterion validity, we observed high performance of the automated Literacy Profile with respect to expert-rated HL for all racial/ethnic groups under study, with c-statistics of > 0.82 for all groups and only minor variance between them (see Figure 1).

Applying the automated Literacy Profile algorithm to the full sample ($N=9,527$)

Figure 1. Discriminatory performance results of the Literacy Profile with respect to expert-rated health literacy, across racial and ethnic sub-groups, using the c-statistic (area under the receiver operator [ROC] curves).

generated rates of limited HL across race/ethnicity that varied in a manner consistent with prior HL research.²⁻³ Specifically, the prevalence of limited HL was 29.6% among White/non-Hispanic patients (total N= 2,797), 39.5% among Black/non-Hispanic (total N=1,409), 46.8% among Hispanic (total N=1,374), and 39.1% among Asian/Pacific Islanders (total N=2,894).

With respect to predictive validity, in our analysis of the relationships between HL—as measured by the automated Literacy Profile—and health outcomes in the entire ECLIPPSE sample, we observed that patients with limited HL compared with those with high HL, demonstrated statistically significantly worse physician communication, medication adherence, hypoglycemia, diabetes control, and ED utilization (Table 1).

In our race/ethnicity-stratified analyses of the relationships between HL—as measured by the automated Literacy Profile—and health outcomes in the ECLIPPSE sample, we observed statistically significant relationships between: (1) HL and communication among Hispanic and among Asian/Pacific Islanders; (2) HL and adherence among Black/non-Hispanics and among Hispanics; (3) HL and diabetes control among White/non-Hispanics, Hispanics, and Asian/Pacific Islanders; and (4) HL and ED visits among White/non-Hispanics (see Table 2).

We observed no interactions ($p < .20$) among HL, race/ethnicity, and any of the health outcomes. Further, while not always significant, the point estimates for the odds ratios

Table 1.

RELATIONSHIPS BETWEEN THE HEALTH LITERACY MEASURE (AUTOMATED LITERACY PROFILE) AND HEALTH OUTCOMES IN THE ECLIPPSE SAMPLE (N=9,527)

Outcome	Health Literacy as Measured by Literacy Profile	
		(%)
Poor Communication	Limited	18.4
	Adequate	12.7
	p-value	<.001
Poor Adherence	Limited	38.8
	Adequate	29.8
	p-value	<.001
≥1 Severe Hypoglycemia	Limited	4.5
	Adequate	3.5
	p-value	.02
Poor diabetes control	Limited	19.6
	Adequate	15.0
	p-value	<.001
>1 ED visit	Limited	40.0
	Adequate	36.1
	p-value	.01

Table 2.
RELATIONSHIPS BETWEEN THE HEALTH LITERACY MEASURE (AUTOMATED LITERACY PROFILE) AND HEALTH OUTCOMES AMONG THE ECLIPSE SAMPLE, STRATIFIED BY RACE/ETHNICITY

Outcome	Health Literacy (Literacy Profile)	White/NH (%) (N=2797)	Black/NH (%) (N=1409)	Hispanic (%) (N=1374)	Asian/PI (%) (N=2894)
Poor Communication	Limited	10.5	14.1	24.4	26.7
	Adequate	8.4	11.5	16.8	17.2
	p-value	.18	.33	.03	<.001
Poor Adherence	Limited	29.7	51.9	59.7	29.7
	Adequate	26.1	37.2	36.4	25.5
	p-value	.24	.007	<.001	.13
≥1 Severe Hypoglycemia	Limited	3.9	6.7	5.1	3.2
	Adequate	2.9	5.2	3.7	2.9
	p-value	.26	.33	.30	.71
Poor diabetes control	Limited	17.9	22.1	29.6	13.8
	Adequate	12.1	20.9	22.0	10.7
	p-value	.002	.73	.03	.05
ED visits (Mean & SD)	Limited	40.3	45.8	37.4	27.2
	Adequate	30.9	42.7	34.4	25.1
	p-value	.005	.59	.53	.41

observed in the overall sample were similar for all health outcomes across all racial/ethnic groups (see Figure 2).

Discussion

To our knowledge, this is the first study to evaluate rigorously the performance of a HL measure across the largest racial/ethnic groups in the U.S. Furthermore, our study not only compared performance of a HL measure in terms of its criterion validity (the degree to which a new measure is associated with other measures of the same construct), but also attempted to compare predictive validity (the degree to which a new measure is associated with indicators of other constructs, based on prior research or established theory). Assessing the validity of HL measures across races/ethnicities is especially important, given the disproportionate burden that limited HL places on vulnerable populations, the role that HL plays as a contributor to health disparities, and the checkered history of literacy measurement (and mismeasurement) in the U.S., especially as a means to oppress Black Americans.^{1,39–40} Finally, our study is particularly relevant and novel in that the HL measure that we were examining represents the product of a set of AI-based methods that, to date, have not been applied to the

Figure 2. Odds ratios (95% CI) for outcomes, comparing limited health literacy to adequate health literacy, for each racial/ethnic group, and for total ECLIPPSE sample (N=9,527).

measurement of HL. Specifically, the Literacy Profile was created by employing a form of computational linguistics that brings together natural language processing with machine learning. While the application of AI in health care is rapidly expanding, concerns have been raised regarding whether applications generated via machine learning perform well across diverse populations,¹⁶ and whether AI might actually perpetuate biases due to race and ethnicity.¹⁷ One explanation for the finding that the automated Literacy Profile had sufficient validity across race/ethnicity may have to do with the broad diversity of the sample from which the measure was generated—one that included significant representation from the four racial and ethnic groups under study.¹⁶ This likely enabled the machine learning process to generate an algorithm that is applicable to diverse populations.

Our research suggests that the automated Literacy Profile has consistently high levels of criterion validity across races and ethnicities in an insured U.S. population, with c-statistics ranging from 0.82–0.89. We believe these findings have significant external validity, as the study took place in an integrated health system that delivers care to an insured population that is not only racially and ethnically diverse but that is largely representative in terms of socioeconomic status, with the exception of the extremes of income.⁴¹ While limited HL is more concentrated in safety-net health care settings, it is still common in this fully insured population. Kaiser Permanente Northern California cares for a sizable Medicaid population, and over one third of their diabetes patients have limited HL.

Our study also found that the relationships between HL—as measured by the automated Literacy Profile—and a range of diabetes outcomes demonstrated similar patterns across races/ethnicities. This should provide motivation for additional translational and implementation research involving the Literacy Profile, enabling HL measurement at a scale that could improve clinical care and population health among diverse populations.

Among other next steps, future research should examine how well the Literacy Profile performs in other settings, such as safety-net settings, and among English speakers for whom English is a second language. We had insufficient power to address this latter question within racial/ethnic subgroups, as only 2.6% of our total sample reported “always,” “often,” or “sometimes” having difficulties speaking English.

Generating accurate information on a diverse population’s HL or on an individual patient’s HL in an efficient and automated fashion opens new avenues that could improve health services delivery and population management. The value of our approach is that it could obviate the need to measure patients’ HL one patient at a time; the effort required to operationalize the automated system could provide economies of scale. The automated Literacy Profile has the potential to enable health systems (a) to efficiently determine whether quality of care and outcomes vary by patient HL; (b) to inform clinicians to enable improvements in individual-level care; and (c) to identify populations and/or individual patients who may be at risk of miscommunication in order to target and deliver tailored health communications and self-management support.

In 2012, the National Academy of Medicine defined the attributes of *health literate health care organizations*, calling for health systems to measure the extent to which quality and outcomes differ across patient HL level so that systems can take steps to reduce HL-related disparities and track the success of quality improvement efforts.⁴² However, to date, no measure of HL has been available to enable such comparisons. Furthermore, prior studies have demonstrated that clinicians often overestimate the HL status of their patients.⁴³ However, when their patients have been screened, primary care physicians have been shown to be receptive to this information and, once they have learned that a patient has limited HL, physicians have been shown to engage in a range of communication behaviors that can promote better comprehension and adherence. The translational implications of the research on physician behavior has been limited due, in part, to the lack of efficient and scalable measures of HL, as well as physicians’ reports that in order for them to best respond, they would need additional system-level support. Finally, research has shown that HL-appropriate communication interventions can disproportionately benefit those with limited HL skills or narrow HL-related disparities in such conditions such as diabetes, heart failure, asthma, and end-of-life care.^{11,44–48}

Translation of this research into real-world settings, however, has been hampered by the inability to scale the identification of limited HL so as to target those most in need. Health systems are increasingly interested in incorporating predictive analytics as a means of risk stratifying and targeting care. Harnessing big (linguistic) data by using natural language processing and machine learning approaches to categorize HL also opens up possibilities for enhancing population management. Not doing so in population management interventions has been shown to amplify HL-related disparities.⁴⁹

In this study, we assessed the comparative validity of a novel HL measure that was generated from computational linguistic analyses of patients' written language across race and ethnicity. The ECLIPPSE Project is the first attempt to measure HL by assessing patients' own original written content, specifically written communications to their physicians. Notably, studies in the field of general literacy have shown that literacy-related production (e.g., writing skill) is highly correlated with literacy-related comprehension (e.g., reading skill), providing a rationale for harnessing patients' SMs to assess HL. Evidence from the general literacy field suggests that individuals' ability to write is also strongly associated with other domains of literacy, linguistic competence, and problem-solving capacities.^{50–53} However, HL is a multifaceted construct that includes not only the ability of patients to communicate information but also the ability to process, comprehend, and act on health information that they receive. A more comprehensive measure of patients' HL would include not just communication ability, but also patients' ability to read and understand specific health topics, critically appraise and execute health instructions, including verbal instructions, and effectively problem-solve based on a foundation of health-related knowledge.⁵⁴

Additionally, while the model developed here is a strong indicator of patients' unidirectional communicative ability via online health portals (specifically using SMs), much health communication is not written. To capture this variance, future studies should also collect data from spoken exchanges between patients and physicians.⁵⁵ Relatedly, while our objective was to measure patients' HL, we acknowledge that assessing the linguistic content of only one actor in a communication exchange limited our ability to evaluate communication exchanges and seek evidence (or absence) of comprehension. Nevertheless, our findings that a model of HL derived from expert ratings of patient SMs was predictive of patient reports of poor receptive communication suggests that limited HL as determined by the Literacy Profile may be a marker for less interactive and lower-quality bidirectional communication. While our work harnessed SM exchange to estimate HL, there is reason to believe that the Literacy Profile measure may also be a marker of skills in other communication contexts, not just written digital communication. For example, prior research has found that limited literacy is correlated with greater difficulties with oral/aural communication.^{56–58} That we found this measure to be associated with multiple health outcomes and that it can have clinical consequences¹⁵ further supports the notion that this form of communicative HL may be a marker of more general health communication challenges.

There are a number of additional limitations to our study. First, while the patterns we observed between HL and health outcomes were fairly consistent across races/ethnicities, not every relationship for each group was statistically significant. Insofar as the stratified analyses we presented were bivariate in nature, assessing the effect of HL on outcomes independent of other factors was beyond the scope of the current study (and our study was not designed to explore causal effects).⁵⁹ Second, while our patient sample was large and diverse, and while we studied a very large number of SMs, we only were able to analyze those patients who had engaged in SM, likely excluding patients with severe HL limitations or other barriers to portal use. However, in a separate analysis from the ECLIPPSE Study, we have found that patients with limited HL are accelerating in their use of patient portals and SM relative to those with adequate HL.⁶⁰ Between 2006

and 2015, the proportion of those with limited HL who used the portal to engage in two or more SM threads increased nearly 10-fold (from 6% to 57%), compared with a fivefold increase among those with adequate HL (13% to 74%). Prior research has also found that portal use historically has been lower for minority subgroups,^{61–63} but our recent research suggests that such disparities are also narrowing.⁵⁵ Between 2006 and 2015, the proportion of those who engaged in two or more SM threads increased to a greater extent among Black/non-Hispanics (from 6.5% to 56%) and Hispanics (from 5.3 to 56%) than it did among White/non-Hispanics (from 15% to 77%).

Third, while our study advances the field of HL measurement by virtue of demonstrating that an automated HL measure can be derived from linguistic analyses of patients' own written language, we also recognize that the use of expert ratings of patients' SM quality and content is not free from the risk of bias. Health disparities are produced and perpetuated by multilevel forces operating at the individual, family, health system, community, and public policy levels that mutually reinforce each other to produce injustice and perpetuate inequity. The problem of cultural hegemony in literacy assessment, and the untoward downstream effects of related mismeasurement, has been well elucidated in the field of social psychology.⁶⁴ Nearly all HL measures have such limitations when it comes to using them as the gold standard for the development of a novel measure. A recent review of HL research measures found that at least 200 unique measures have been created and employed, with most measures (52%) requiring paper and pencil responses, and some measures (12%) requiring more than 15 minutes to administer. Of the 200, 128 (64%) measured general HL, and 76 (38%) measured disease or content-specific HL.¹ Thirty-one (15.5%) assessed pronunciation, 25 (12.5%) assessed conceptual knowledge test, and 43 (21.5%) assessed comprehension. It is likely that most conventional HL assessments are bounded by cultural and linguistic assumptions derived from the dominant, majority population, making assessments of each of these domains potentially subject to bias. While our study provides a degree of reassurance that the performance of the automated Literacy Profile has sufficient criterion and predictive validity across racially and ethnically diverse groups to justify its use more broadly, more research is needed to assess patient HL in a comprehensive, holistic, and unbiased manner, and to expand the assessment of reliability and validity across sub-groups of interest in order to avoid misattributing health disparities to limited HL. While our expert raters were blinded to patients' names and demographic characteristics, and while each had received implicit bias training, it is possible that their ratings were influenced by their perceptions of the races/ethnicities of the authors of the SMs they rated. We intend to carry out additional work on the Literacy Profile, assessing whether expert raters who are racially and ethnically concordant with the patients whose SMs are being rated generate similar HL scores to those who are discordant. In addition, we are undertaking qualitative analyses of a purposive sample of SM exchanges to understand whether patients' race/ethnicity and/or racial/ethnic concordance between patient and physician influence SM exchange, and whether any differences are moderated by patient HL. Finally, while the performance of the automated Literacy Profile overall and across races/ethnicities appears more than adequate, the fact that we used linguistic indices developed and validated before email exchange became so prevalent may have limited the accuracy of our categorization of HL.

Conclusions. Limited health literacy (HL) is associated with worse health and can serve as both a mediator and moderator of health disparities related to race and ethnicity. While some HL interventions have been shown to improve outcomes and reduce disparities, measurement constraints, including lack of validity assessment across racial/ethnic groups and administration challenges, have undermined the field and impeded scaling of interventions. In this study, we employed computational linguistics to develop a novel HL measure, analyzing language from more than 300,000 messages sent by 9,527 ethnically diverse diabetes patients via a patient portal. This AI approach harnessed big linguistic data to estimate HL, applying machine learning to a gold standard of expert ratings of a purposive sample of messages to create what we have called the Literacy Profile and then applying this tool to categorize the HL of the entire sample. In our previous research, we demonstrated that the Literacy Profile performed well in discriminating between high and low HL and was predictive of a range of diabetes-related health outcomes. In the current study, we carried out stratified analyses to determine if the Literacy Profile has comparable criterion and predictive validities with respect to physician communication, medication adherence, severe hypoglycemia, poor glycemic control ($A1c >9\%$) and ED utilization among White/non-Hispanics, Black/non-Hispanics, Hispanics, and Asian/Pacific Islanders. We discovered that criterion validity of the Literacy Profile was consistently high across White-NH, Black-NH, Hispanic, and Asian/Pacific Islander groups. Furthermore, across racial/ethnic groups, we observed that the proportion of patients with low HL who had worse outcomes was consistently higher than that of high HL, using indicators of process, behavioral, metabolic, safety, and health care utilization outcomes. This is the first study to validate a HL measure across the most common racial/ethnic sub-groups in the U.S. While concerns have arisen regarding bias in AI, automated Literacy Profiles appear sufficiently valid across races/ethnicities, enabling HL measurement at a scale that could improve clinical care and population health among diverse populations.

In sum, an automated Literacy Profile could provide an efficient means to identify subpopulations of diverse patients with limited HL and assist health systems in their journeys to become more health literate health care organizations.⁶⁵ Employing such a scalable, automated measure of HL has the potential to enable health systems (a) to determine efficiently whether quality of care and health outcomes vary by patient HL; (b) to identify populations and/or individual patients at risk of miscommunication in order to target and deliver tailored health communications and self-management support interventions; and (c) to inform clinicians in order to promote improvements in individual-level care. In view of this, our research to develop an automated method for HL assessment that performs well across races/ethnicities represents a significant accomplishment with potentially broad clinical and population health benefits in the context of health services delivery. As secure messaging is rapidly accelerating in health systems nationwide, and as it becomes a standard of care as a vehicle to enhance patient-provider communication,⁶⁰ we believe our study provides important additional rationale to encourage the use of the Literacy Profile to advance health equity.

Acknowledgments

Major funding for this study was supported by grants from the National Library of Medicine (NLM R01 LM012355) and the National Institute of Diabetes and Digestive and Kidney Diseases (P30 DK092924).

References

- Schillinger D. The intersections between social determinants of health, health literacy, and health disparities. *Stud Health Technol Inform*. 2020 Jun 25;269:22–41.
- Kutner, M., Greenburg, E., Jin, Y, et al. The health literacy of America's adults: results from the 2003 national assessment of adult literacy. National Center for Education Statistics. 2006;483.
- Lee HY, Rhee TG, Kim NK, et al. Health literacy as a social determinant of health in Asian American immigrants: findings from a population-based survey in California. *J Gen Intern Med*. 2015;30(8):1118–24.
<https://doi.org/10.1007/s11606-015-3217-6>
PMid:25715993 PMCID:PMC4510223
- Bailey SC, Brega AG, Crutchfield TM, et al. Update on health literacy and diabetes. *Diabetes Educ*. 2014 Sep–Oct;40(5):581–604.
<https://doi.org/10.1177/0145721714540220>
PMid:24947871 PMCID:PMC4174500
- Bauer AM, Schillinger D, Parker M, et al. Health literacy and antidepressant medication adherence among adults with diabetes: the diabetes study of Northern California (DISTANCE). *J Gen Intern Med*. 2013 Sep;28(9):1181–7.
<https://doi.org/10.1007/s11606-013-2402-8>
PMid:23512335 PMCID:PMC3744297
- Karter, AJ, Parker MM, Duru OK, et al. Impact of a pharmacy benefit change on new use of mail order pharmacy among diabetes patients: the Diabetes study of Northern California (DISTANCE). *Health Serv Res*. 2015 Apr;50(2):537–59.
<https://doi.org/10.1111/1475-6773.12223>
PMid:25131156 PMCID:PMC4329275
- Brach C, Keller D, Hernandez LM, et al. Ten attributes of health literate health care organizations. *NAM Perspectives*. Washington, DC: National Academy of Medicine, 2012 Jun 7.
- Stewart MA. Effective physician-patient communication and health outcomes: a review. *CMAJ*. 1995 May 1;152(9):1423.
- Ratanawongsa N, Karter AJ, Parker MM, et al. Communication and medication refill adherence: the Diabetes Study of Northern California. *JAMA Intern Med*. 2013 Feb 11;173(3):210–8.
<https://doi.org/10.1001/jamainternmed.2013.1216>
PMid:23277199 PMCID:PMC3609434
- Schillinger D, Bindman A, Wang F, et al. Functional health literacy and the quality of physician-patient communication among diabetes patients. *Patient Educ Couns*. 2004 Mar 1;52(3):315–23.
[https://doi.org/10.1016/S0738-3991\(03\)00107-1](https://doi.org/10.1016/S0738-3991(03)00107-1)
- Castro CM, Wilson C, Wang F, et al. Babel babble: physicians' use of unclarified medical jargon with patients. *Am J Health Behav*. 2007 Sep–Oct;31(1):S85–95.
<https://doi.org/10.5993/AJHB.31.s1.11>

12. Schillinger D, Piette J, Grumbach K, et al. Closing the loop: physician communication with diabetic patients who have low health literacy. *Arch Intern Med*. 2003 Jan 13;163(1):83–90.
<https://doi.org/10.1001/archinte.163.1.83>
PMid:12523921
13. 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 Jan;70(1):102–10.
<https://doi.org/10.1016/j.pec.2007.09.008>
PMid:17997264 PMCID:PMC2745943
14. DeWalt, D. A., Schillinger, D., Ruo, B., et al. (2012). Multisite randomized trial of a single-session versus multisession literacy-sensitive self-care intervention for patients with heart failure. *Circulation*. 2012 Jun 12;125(23):2854–62.
<https://doi.org/10.1161/CIRCULATIONAHA.111.081745>
PMid:22572916 PMCID:PMC3400336
15. Schillinger, D., McNamara, D., Crossley, S., et al. (2017). The next frontier in communication and the ECLIPPSE study: bridging the linguistic divide in secure messaging. *J Diabetes Res*. 2017
<https://doi.org/10.1155/2017/1348242>
PMid:28265579 PMCID:PMC5318623
16. Crossley SA, Balyan R, Lui J, et al. Developing and testing automatic models of patient communicative health literacy using linguistic features: findings from the ECLIPPSE study. *Health Commun*. 2020 Mar 2;1–11.
<https://doi.org/10.1080/10410236.2020.1731781>
PMid:32114833
17. Liu Y, Chen PC, Krause J, et al. How to read articles that use machine learning: users' guides to the medical literature. *JAMA*. 2019 Nov 12;322(18):1806–16.
<https://doi.org/10.1001/jama.2019.16489>
PMid:31714992
18. Parikh RB, Teeple S, Navathe AS. Addressing bias in artificial intelligence in health care. *JAMA*. 2019 Nov 22;322(24):2377–8.
<https://doi.org/10.1001/jama.2019.18058>
PMid:31755905
19. Schillinger D, Balyan R, Crossley SA, et al. Employing computational linguistics techniques to identify limited patient health literacy: findings from the ECLIPPSE study. *Health Serv Res*. 2020 Sep 23.
<https://doi.org/10.1111/1475-6773.13560>
PMid:32966630 PMCID:PMC7839650
20. Balyan R, Crossley SA, Brown W, et al. Using natural language processing and machine learning to classify health literacy from secure messages: the ECLIPPSE study. *PLoS One*. 2019 Feb 22;14(2):e0212488.
<https://doi.org/10.1371/journal.pone.0212488>
PMid:30794616 PMCID:PMC6386302
21. Centers for Disease Control and Prevention. Diabetes report card 2017. Atlanta, GA: US Dept of Health and Human Services, 2018.
22. Schillinger D, Grumbach K, Piette J, et al. Association of health literacy with diabetes outcomes. *JAMA*. 2002 Jul 24–31;288(4):475–82.
<https://doi.org/10.1001/jama.288.4.475>
PMid:12132978

23. Harris LT, Haneuse SJ, Martin DP, et al. Diabetes quality of care and outpatient utilization associated with electronic patient-provider messaging: a cross-sectional analysis. *Diabetes Care*. 2009 Jul;32(7):1182–7.
<https://doi.org/10.2337/dc08-1771>
PMid:19366959 PMCID:PMC2699712
24. Kanaya AM, Adler N, Moffet HH, et al. Heterogeneity of diabetes outcomes among Asians and Pacific Islanders in the US: the diabetes study of northern California (DISTANCE). *Diabetes Care*. 2011 Apr;34(4):930–7.
<https://doi.org/10.2337/dc10-1964>
PMid:21350114 PMCID:PMC3064053
25. Moffet HH, Adler N, Schillinger D, et al. Cohort profile: the diabetes study of Northern California (DISTANCE)-objectives and design of a survey follow-up study of social health disparities in a managed care population. *Int J Epidemiol*. 2009 Feb;38(1):38–47.
<https://doi.org/10.1093/ije/dyn040>
PMid:18326513 PMCID:PMC2635421
26. Semere W, Crossley S, Karter AJ, et al. Secure messaging with physicians by proxies for patients with diabetes: findings from the ECLIPSE Study. *J Gen Intern Med*. 2019 Aug 19;1–7.
27. Crossley SA, Kyle K, McNamara DS. To aggregate or not? linguistic features in automatic essay scoring and feedbacks systems. *Grantee Submission*. 2015;8(1).
28. Crossley SA, & McNamara DS. Say more and be more coherent: how text elaboration and cohesion can increase writing quality. *J Writ Res*. 2016 Feb;7(3):351–370.
<https://doi.org/10.17239/jowr-2016.07.03.02>
29. McNamara, DS, Crossley SA, Roscoe RD, et al. A hierarchical classification approach to automated essay scoring. *Assessing Writing*. 2015;23:35–59.
<https://doi.org/10.1016/j.asw.2014.09.002>
30. Kyle K, Crossley S, Berger C. The tool for the automatic analysis of lexical sophistication (TAALES): version 2.0. *Behav Res Methods*. 2018 Jun;50(3):1030–46.
<https://doi.org/10.3758/s13428-017-0924-4>
PMid:28699123
31. Kyle, K, & Crossley SA. Automatically assessing lexical sophistication: indices, tools, findings, and application. *TESOL Quarterly*. 2015 Dec;49(4):757–86.
<https://doi.org/10.1002/tesq.194>
32. Crossley SA, Kyle K, McNamara DS. The tool for the automatic analysis of text cohesion (TAACO): automatic assessment of local, global, and text cohesion. *Behav Res Methods*. 2016 Dec;48(4):1227–1237.
<https://doi.org/10.3758/s13428-015-0651-7>
PMid:26416138
33. Crossley SA, Kyle K, Dascalu M. The tool for the automatic analysis of cohesion 2.0: integrating semantic similarity and text overlap. *Behav Res Methods*. 2019 Feb;51(1):14–27.
<https://doi.org/10.3758/s13428-018-1142-4>
PMid:30298264
34. Kyle K. Measuring syntactic development in L2 writing: fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication (doctoral dissertation). Atlanta, GA: Georgia State University, 2015. Available at: http://scholarworks.gsu.edu/alesl_diss/35.

35. Kyle K, Crossley SA. Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *Mod Lang J*. 2018 Feb 16;102(2):333–349.
<https://doi.org/10.1111/modl.12468>
36. Steiner JF, Koepsell TD, Fihn SD, et al. A general method of compliance assessment using centralized pharmacy records: description and validation. *Medical Care*. 1988 Aug 1;814–23.
<https://doi.org/10.1097/00005650-198808000-00007>
PMid:3398608
37. Steiner JF, Prochazka AV. The assessment of refill compliance using pharmacy records: methods, validity, and applications. *J Clin Epidemiol*. 1997 Jan 1;50(1):105–16.
[https://doi.org/10.1016/S0895-4356\(96\)00268-5](https://doi.org/10.1016/S0895-4356(96)00268-5)
38. Sarkar U, Karter AJ, Liu JY, et al. Hypoglycemia is more common among type 2 diabetes patients with limited health literacy: the diabetes study of northern California (DISTANCE). *J Gen Intern Med*. 2010 Sep 1;25(9):962–8.
<https://doi.org/10.1007/s11606-010-1389-7>
PMid:20480249 PMCID:PMC2917655
39. Ginde AA, Blanc PG, Lieberman RM, et al. Validation of ICD-9-CM coding algorithm for improved identification of hypoglycemia visits. *BMC Endocr Disord*. 2008 Apr 1;8(1):4.
<https://doi.org/10.1186/1472-6823-8-4>
PMid:18380903 PMCID:PMC2323001
40. Harris VJ. African-American conceptions of literacy: a historical perspective. *Theory Pract*. 1992 Sep 1;31(4):276–86
<https://doi.org/10.1080/00405849209543554>
41. Goldman D. The modern-day literacy test: felon disenfranchisement and race discrimination. *Stanf Law Rev*. 2004; 57: 611.
42. Moffet HH, Adler N, Schillinger D, et al. lCohort profile: the diabetes study of northern California (DISTANCE)-objectives and design of a survey follow-up study of social health disparities in a managed care population. *Int J Epidemiol*. 2009 Feb;38(1): 38–47.
<https://doi.org/10.1093/ije/dyn040>
PMid:18326513 PMCID:PMC2635421
43. Brach C, Keller D, Hernandez LM, et al. Ten attributes of health literate health care organizations. *NAM Perspectives*. Washington, DC: National Academy of Medicine, 2012 Jun 7.
44. Seligman HK, Wang FF, Palacios JL, et al. Physician notification of their diabetes patients' limited health literacy: a randomized, controlled trial. *J Gen Intern Med*. 2005 Nov 1;20(11):1001–7.
<https://doi.org/10.1111/j.1525-1497.2005.00189.x>
PMid:16307624 PMCID:PMC1490250
45. DeWalt DA, Baker DW, Schillinger D, et al. A multisite randomized trial of a single-versus multi-session literacy sensitive self-care intervention for patients with heart failure. *J Gen Intern Med*. 2011 May 1;26:S57–S58.
46. Sheridan SL, Halpern DJ, Viera AJ, et al. Interventions for individuals with low health literacy: a systematic review. *J Health Commun*. 2011 Sep 30;16(3):30–54.
<https://doi.org/10.1080/10810730.2011.604391>
PMid:21951242
47. Machtinger EL, Wang F, Chen LL, et al. A visual medication schedule to improve

- anticoagulation control: a randomized, controlled trial. *Jt Comm J Qual Patient Saf.* 2007 Oct 1;33(10):625–35.
[https://doi.org/10.1016/S1553-7250\(07\)33072-9](https://doi.org/10.1016/S1553-7250(07)33072-9)
48. Paasche-Orlow MK, Riekert KA, Bilderback A, et al. Tailored education may reduce health literacy disparities in asthma self-management. *Am J Respir Crit Care Med.* 2005 Oct 15;172(8):980–6.
<https://doi.org/10.1164/rccm.200409-1291OC>
PMid:16081544 PMCID:PMC2718412
49. Sudore RL, Schillinger D, Katen MT, et al. Engaging diverse English-and Spanish-speaking older adults in advance care planning: the PREPARE randomized clinical trial. *JAMA Intern Med.* 2018 Dec 1;178(12):1616–25.
<https://doi.org/10.1001/jamainternmed.2018.4657>
PMid:30383086 PMCID:PMC6342283
50. Karter AJ, Parker MM, Duru OK, et al. Impact of a pharmacy benefit change on new use of mail order pharmacy among diabetes patients: the diabetes study of northern California (DISTANCE). *Health Serv Res.* 2015 Apr;50(2):537–59.
<https://doi.org/10.1111/1475-6773.12223>
PMid:25131156 PMCID:PMC4329275
51. Allen LK, Dascalu M, McNamara DS, et al. Modeling individual differences among writers using readerbench. In: *Proceedings of the 8th International Conference on Education and New Learning Technologies (EduLearn)*, Jul 4–6, 2016:5269–5279. Barcelona, Spain: IATED, 2016.
52. Allen LK, Snow EL, Jackson GT, et al. Reading components and their relation to writing. *L'Année psychologique.* 2014;114(4):663–691.
<https://doi.org/10.4074/S0003503314004047>
53. Crossley SA, Allen L, Snow E, et al. Incorporating learning characteristics into automatic essay scoring models: what individual differences and linguistic features tell us about writing quality. *Journal of Educational Data Mining.* 2016;8(2):1–19.
54. Schoonen R. Are reading and writing building on the same skills? The relationship between reading and writing in L1 and EFL. *Reading and Writing.* 2019 Mar;32(3):511–535.
<https://doi.org/10.1007/s11145-018-9874-1>
55. Nutbeam D. (2009). Defining and measuring health literacy: what can we learn from literacy studies? *Int J Public Health.* 2009;54:303.
<https://doi.org/10.1007/s00038-009-0050-x>
PMid:19641847
56. Harrington KE, Valerio MA. A conceptual model of verbal exchange health literacy. *Patient Educ Couns.* 2014 Mar;94(3):403–410.
<https://doi.org/10.1016/j.pec.2013.10.024>
PMid:24291145 PMCID:PMC3944213
57. Nouri SS, Rudd RE. Health literacy in the “oral exchange”: an important element of patient-provider communication. *Patient Educ Couns.* 2015 May;98(5):565–571.
<https://doi.org/10.1016/j.pec.2014.12.002>
PMid:25620074
58. Schonlau M, Martin L, Haas A, et al. Patients' literacy skills: more than just reading ability *J Health Commun.* 2011 Nov;16(10):1046–1054.
<https://doi.org/10.1080/10810730.2011.571345>
PMid:21916699 PMCID:PMC3213295

59. Koch-Weser S, Rudd RE, DeJong W. Quantifying word use to study health literacy in doctor-patient communication. *J Health Commun.* 2010 Sep;15(6):590–602.
<https://doi.org/10.1080/10810730.2010.499592>
PMid:20812122 PMCID:PMC2933931
60. Rasu RS, Bawa WA, Suminski R, et al. Health literacy impact on national healthcare utilization and expenditure. *Int J Health Policy Manag.* 2015 Aug 17;4(11):747–755.
<https://doi.org/10.15171/ijhpm.2015.151>
PMid:26673335 PMCID:PMC4629700
61. Cemballi AG, Karter, AJ, Schillinger D, et al. Descriptive examination of secure messaging in a longitudinal cohort of diabetes patients in the ECLIPPSE study. *J Am Med Inform Assoc.* 2020 Nov 24;ocaa281.
<https://doi.org/10.1093/jamia/ocaa281>
PMid:33236117
62. Lyles CR, Karter AJ, Young BA, et al. Provider factors and patient-reported healthcare discrimination in the diabetes study of California (DISTANCE). *Patient Educ Couns.* 2011 Dec;85(3):e216–24.
<https://doi.org/10.1016/j.pec.2011.04.031>
PMid:21605956 PMCID:PMC3178668
63. Lyles CR, Allen JY, Poole D, et al. “I want to keep the personal relationship with my doctor”: understanding barriers to portal use among African Americans and Latinos. *J Med Internet Res.* 2016 Oct;18(10):e263.
<https://doi.org/10.2196/jmir.5910>
PMid:27697748 PMCID:PMC5067358
64. Perzynski AT, Roach MJ, Shick S, et al. Patient portals and broadband internet inequality. *J Am Med Inform Assoc.* 2017;24(5):927–32.
<https://doi.org/10.1093/jamia/ocx020>
PMid:28371853 PMCID:PMC6259664
65. Chege M. Literacy and hegemony: critical pedagogy vis-à-vis contending paradigms. *International Journal of Teaching and Learning in Higher Education* 2009;21(2): 228–238.
66. Brach C, Keller D, Hernandez LM, et al. Ten attributes of health literate health care organizations. *NAM Perspectives.* Washington, DC: National Academy of Medicine, 2012 Jun 7.