



State of the Science of Health Literacy Measures, 1975-2013

Tam Nguyen¹, PhD, MSN/MPH, RN; Hae-Ra Han², RN, PhD; Cindy Cao¹, BSN; Miyong T. Kim³, RN, PhD, FAAN
¹Boston College William F. Connell School of Nursing; ²Johns Hopkins University School of Nursing; ³University of Texas, Austin School of Nursing



Correspondence: Tam H. Nguyen, email: tam.nguyen@bc.edu
Funding & Support: NINR F31 NR010992, BC Undergraduate Research Honors Program, BD Research Incentive Grant

BACKGROUND & PURPOSE

The growing prevalence and consequences of limited health literacy has sparked a proliferation of measures to accurately quantify the problem.

The purpose of this systematic review is to synthesize the current state of the science in health literacy measures.

Specifically, this review aims to:

- Assess how measures conceptualize health literacy
- Whether they measure general or content specific health literacy
- The measures' psychometric properties, and
- The measures' utility in assessing health literacy in ethnic minority populations, a sub-group that has been known to be at high risk for limited health literacy.

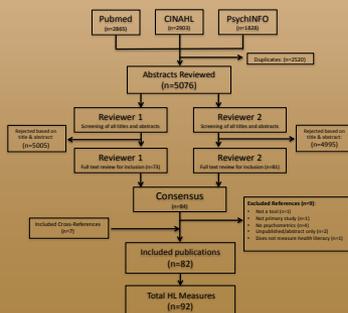
METHODS

PubMed, PsycINFO and CINHAL data-bases were searched for health literacy measurement studies from 1975 to September 2013. The following search strategy was used in PubMed, corresponding to terms and synonyms related to "health literacy" and "measures":

(health literacy [mh] OR "health literacy" OR (health [tiab] AND (literate [tiab] OR literacy [tiab]) OR numeracy [tiab]) AND (measur* [tw] OR language tests [mh] OR psychometrics [mh] OR tool* [tiab] OR survey* [tiab] OR instrument* [tiab] OR questionnaire* [tiab] OR screen* [tw] OR assessment [tiab]))

The search strategies used in CINHAL and PsycINFO were built to reflect similar MeSH and text-word terms used in the PubMed search [APPENDIX 1].

Two authors independently reviewed the titles and abstracts of all publications. After agreement was reached, full length articles were reviewed using a standard extraction form.



RESULTS

The search yielded 82 publications that met the inclusion criteria, resulting in **92 unique HL measures** (some publications reported on multiple measures)

Conceptualization of HL

- Many definitions of HL have been used. A common thread among earlier measures was that they defined HL as a set of *individual capacities* that allows a person to gain skills needed to function in the healthcare setting (i.e. print literacy, oral literacy, numeracy, listening, speaking, and cultural/conceptual knowledge), as well as the *application of critical skills* needed to actively engage in healthcare (i.e., finding, understanding, evaluating, and communicating information)
- More recent measures define HL as an interaction between individual abilities and factors at the personal, healthcare system and broader community level

General vs. Content Specific HL

- Of the 92 measures identified, 47 assessed general HL, and 45 assessed content/context-specific HL. The high number of content/context-specific measures speaks to the idea that HL can differ across different contexts

Psychometric properties

- The reported reliability was generally strong across most measures (α or KR-20 = 0.70-0.98), with the exception of one (the 4-item ANQ, $\alpha = 0.57$).
- The convergent validity of most measures was positive and significant ($r = 0.32-0.98$). However, the concurrent and predictive validities were often not tested, non-significant or only weakly associated.
- Majority of the measures (n=81) used classical test theory (CTT) for reliability testing and/or item reduction, while a small but growing number reported using item response theory (IRT) either alone or adjunctively (n= 11) with CTT.

HL & minority populations

- Most study samples included predominately White (n=43), Black (n=38), and/or Hispanic (n=16) Americans.
- From 2010, an increase number of measures were developed for or tested with a wider range of minority populations
- Understanding the phonetic structure of various languages is important before directly translating measures

CONCLUSIONS & IMPLICATIONS FOR PRACTICE/RESEARCH

- Despite the incredible number of measures available, a "gold standard" tool that can be used with the general population, within a specific health context, or a specific ethnic group has not been established.
- A clear conceptual basis for the measurement of HL was lacking
- Consistent with recommendations from several health literacy professional groups, there is a need to: (1) develop a unifying conceptual definition of health literacy, and (2) to "clinically re-engineer" health literacy measures.
- What is envisioned is a core set of items that can be universally used, with add-on modules that address targeted health topics. This family of health literacy measures can potentially capitalize on the Computerized Adaptive Testing (CAT) platform, given the large pool of existing measures available. Administering items on a CAT platform allows different items to be given to different people (based on their health literacy level), while still making it possible to compare scores across groups.
- As the science of HL measurement advances, the limitations of current measures should not act as a barrier to their continued use. The simplicity of some measures makes them quite useful in certain settings. For example, word-recognition tests such as the REALM are preferred over more complex measures such as the TOFHLA when the goal is assessing HL in populations with limited English proficiency

STRENGTHS & LIMITATIONS

- A rigorous systematic approach was used to identify existing measures of HL, which yielded over 4 times the number of measures than previous reviews
- However, it is always possible that the search terms and methods used were not sensitive enough to capture everything that is available in the literature
- To address this concern, we made use of careful cross-referencing to ensure the inclusion of as many relevant HL measures as possible

REFERENCES

- Please see handout for reference of all HL measure