OXFORD

Constructing a measure of health literacy in Sub-Saharan African countries

Heather F. McClintock[1],*, Julia M. Alber[2], Sarah J. Schrauben[3], Carmella M. Mazzola[1], and Douglas J. Wiebe[4]

[1]Department of Public Health, College of Health Sciences, Arcadia University, 450 S. Easton, Glenside, PA, 19038, USA, [2]Department of Kinesiology and Public Health, California Polytechnic State University, 1 Grand Ave, San Luis Obispo, CA 93407, USA, [3]Department of Medicine, University of Pennsylvania, 3400 Civic Center Boulevard, Philadelphia, PA 19104, USA and [4]Department of Biostatistics, Epidemiology & Informatics, Perelman School of Medicine, University of Pennsylvania, 423 Guardian Drive, Philadelphia, PA 19104, USA

*Corresponding author. E-mail: devriesmcclintockh@arcadia.edu

Summary

We sought to develop and evaluate a health literacy measure in a multi-national study and to examine demographic characteristics associated with health literacy. Data were obtained from Demographic Health Surveys conducted between 2006–15 in 14 countries in Sub-Saharan Africa. Surveys were the same in all countries but translated to local languages as appropriate. We identified eight questions that corresponded to the National Academy of Medicine (NAM) definition of health literacy. Factor analysis was used to extract one measure of health literacy. Logistic regression was employed to examine the relationship between demographic characteristics and health literacy. A total of 224 751 individuals between the ages of 15 and 49 years were included. The derived health literacy measure demonstrated good internal consistency (Cronbach's $\alpha = 0.72$) and good content validity. The prevalence of high health literacy overall was 35.77%; females 34.08% and males 39.17%; less than or equal to primary education 8.93%, some secondary education 69.40% and $\geq$complete secondary 84.35%. High health literacy varied across nations, from 8.51% in Niger to 63.89% in Namibia. This is the first known study to evaluate a measure of health literacy relying on the NAM definition utilizing a large sample from 14 countries in Sub-Saharan Africa. Our study derived a robust indicator of NAM-defined health literacy. This indicator could be used to examine determinants and outcomes of health literacy in additional countries.

Key words: health literacy, global health, assessment tool, Africa

INTRODUCTION

Recent global initiatives to understand and develop strategies to improve health literacy, namely the National Academy of Medicine (NAM) Roundtable on Health Literacy (Institute of Medicine, 2017) and the Evidence Report on Health Literacy by the Agency for Healthcare Quality and Research (AHQR) (Berkman *et al.*, 2011), indicate that few low or middle lower-income countries have objectively measured health literacy or its impact of health. The concept of health literacy was introduced as a derivation of literacy in the 1970s

(Simonds, 1974), with its primary purpose to serve as an important tool to maximize comprehension during clinical encounters and healthcare settings. In recent years, the meaning and purpose of health literacy have broadened in scope to incorporate a health promotion perspective. This perspective espouses that health literacy is an essential skill needed to navigate a complicated health care system as well as contextual factors outside of the clinical setting (Massey _et al._, 2012).

The NAM defines health literacy as the degree to which individuals have the capacity to obtain, interpret and understand basic health information and services needed to make appropriate health decisions (Institute of Medicine Board on Neuroscience and Behavioral Health and Committee on Health Literacy, 2004). Health literacy measures one's ability to understand and act on health information in an increasingly complex modern health care system. It encompasses proficiency in more than just reading ability but also writing, speaking and listening as well as computational abilities (numeracy) (Berkman _et al._, 2011). Educational attainment and literacy are among the strongest predictors of health status (Zimmerman _et al._, 2015). While education and literacy are related to health literacy, higher levels of education and literacy do not guarantee adequate health literacy to function and interact effectively within a health system. Given the importance of measuring health literacy in large populations and the accompanying challenges, the Demographic Health Survey (DHS) data represent an unprecedented and valuable opportunity to study health literacy on a global scale and to evaluate how it varies within and across populations as it relates to health.

While the NAM's definition of health literacy is recognized and accepted, a lack of development and evaluation of a valid and reliable measure that is obtainable using existing data sources in low or middle lower-income countries, has limited its application (Institute of Medicine, 2017). Approaches that have been employed involve the usage of proxies (e.g. educational attainment) or sole items, often as a part of other measures (e.g. school attendance, reading score), to evaluate health literacy (Berkman _et al._, 2011). As a result, there is currently a lack of comparability between estimates of health literacy within and between countries. Furthermore, the validity of many measures in relation to NAM's definition of health literacy remains unclear.

A recent study by Schrauben and Wiebe (Schrauben and Wiebe, 2017) developed a measure of health literacy using factor analysis methods to derive a single indicator. Its validity was assessed in a large population-based sample of residents in Zambia. This was accomplished by using data from the Demographic Health Surveys (DHS) program, which administers items on its questionnaire that correspond to the NAM's definition of health literacy. The DHS Program, administered by the United States Agency for International Development (USAID), began in 1984 and has conducted surveys in over 90 countries worldwide. In this study, we seek to build on this prior work by expanding the application of this measure from a single country to a multi-national study in which we assess this derived measure in 14 countries in Sub-Saharan Africa. This will facilitate within and between country comparisons and the examination of health literacy as well as associated factors. The specific role of health literacy as a distinct and separate construct has yet to be examined in most low and middle lower-income countries. We sought to develop and test the internal consistency and content validity of a health literacy scale in a multi-national study, and to examine demographic characteristics associated with health literacy.

MATERIALS

In this cross-sectional, community-based study we sought to construct a health literacy measure using DHS survey data from 14 countries in Sub-Saharan Africa. The evaluation of data from multiple countries facilitates comparisons between countries and has implications for widespread application. We also sought to examine respondent characteristics associated with health literacy.

Study sample

We obtained the most recent (2006–15) DHS data from the following 14 developing countries in Sub-Saharan Africa: Cameroon, Democratic Republic of the Congo, Ethiopia, Ghana, Guinea, Ivory Coast, Lesotho, Rwanda, Niger, Namibia, Sierra Leone, Swaziland, Togo and Zambia. Participation by household in the DHS was randomly selected. Surveys were verbally administered in the language of preference by a gender-matched team member. DHS surveys are funded and administered by USAID and partners. Details of the survey design, protocols, operations and objectives can be found elsewhere (United States Agency for International Development, 2006).

Conceptualization and construction of a measure of health literacy

The definition of health literacy established by NAM includes the capacity to interpret, obtain, understand and make appropriate health decisions. We identified

eight questions within the DHS that corresponded to the NAM definition. For instance, capacity to interpret was assessed by questions about whether primary school was finished and the ability to read a whole or portion of a sentence in the DHS. Capacity to obtain was assessed by questions regarding reading a magazine, listening to the radio or watching television. Capacity to understand was evaluated using questions about learning family planning information from a magazine, radio or television in the last month. Making appropriate health decisions was assessed by two questions. The first question asked if participants had knowledge of a place to get an HIV test. The second question asked if participants had knowledge of a place to acquire condoms. More information on the selection of the items can be found in our earlier study (Schrauben and Wiebe, 2017). The DHS questions utilized can be found in the Supplementary data file.

A single factor was derived from these eight survey questions using factor analysis. The purpose of factor analysis is to reduce the information contained in an original number of items to a smaller set of new composite measures (factors) with a minimum loss of information (Nunnally, 1989). This factor represented health literacy as defined by NAM. Internal consistency of the eight items was assessed using Cronbach's α. To allow for easier conceptual comparisons between health literacy items and relevant parameters, the resulting final score was dichotomized as high versus low. The eight items were summed after weighting each item by its factor loading obtained from a factor analysis using oblique rotation (Harman, 1976). This approach enabled us to assess how question responses related to each other and to extract a single indicator of health literacy which weighted the contribution of each item based on the strength of its loading on a factor representing health literacy or a construct within health literacy. The single derived score was a continuous variable which was standardized with a mean of 0 and standard deviation of 1. Next, tertiles of this score were created and observations were categorized into two groups: high health literacy (top tertile) and low health literacy (bottom two tertiles). Thus, the final health literacy scale was dichotomous with persons in the highest group representing high health literacy (top tertile) and others as not having high health literacy (lower two tertiles) were labeled as having low health literacy.

Validity

In order to examine the validity of our derived measure of health literacy we examined it in relation to a variable indicating whether or not an individual had comprehensive knowledge of HIV/AIDS, assessed by a binary variable. This approach has been previously employed in the DHS, Millennium Development Goals and Sustainable Development Goals (United States Agency for International Development, 2006; United Nations, 2015).

Statistical analysis

Demographic characteristics and health literacy items were summarized using weighted data. Then the chi-squared test and logistic regression were employed to examine the association between health literacy and respondent characteristics. In the logistic regression models initially, crude associations were assessed by running unadjusted models for each demographic variable and health literacy. After this, the model was run to simultaneously adjust for all covariates in the model. Conventional procedures were employed to examine diagnostics such as outliers, indices of model fit and multi-collinearity (Hosmer *et al.*, 1991; Archer and Lemeshow, 2006). Listwise deletion was employed to address missing data and not more than 8% of the cases were excluded in this analysis. Statistical outliers were not detected and sampling weights were employed in the analysis.

Sensitivity analysis

A lack of an existing standardized measure of health literacy led us to use three approaches to examine the sensitivity of our derived dichotomous measure. The first approach involved evaluating the relationship of our original continuous standardized variable obtained from our factor analysis with respondent characteristics in both unadjusted and adjusted ordinary least squares (OLS) regression. In our second strategy, we created a composite score in which each of the eight survey items were scored as one point if an affirmative response was given. This continuous variable was then assessed in relation to our independent variables (respondent characteristics) using OLS. In our third approach, another composite scores was created in which media items were given a value of half a point, instead of one point. This was done because three items were included representing media exposure and we sought to downgrade its role on our examination. This variable was also assessed continuously in relation to our independent variables using OLS. Lastly, we dichotomized these continuous variables previously described (in our second and third approaches) by their median. These new variables were modeled in logistic regression analysis with our independent variables. We also then created tertiles from these variables and then subsequently dichotomized them based on the lower tertile (versus top two tertiles) or top

tertile (versus lower two tertiles). All analysis were conducted using Stata 14 (College Station, TX, USA).

RESULTS

Study sample
This analysis included 224 751 individuals aged of 15 and 49 years old. In all, 150 358 (66.88%) of the respondents were women. Approximately half (49.89%) of respondents were married and over a third (39.90%) lived in an urban setting. Over two-thirds of respondents (68.34%) had comprehensive knowledge of HIV/AIDS. The sample was comprised of residents in 14 African countries with the highest proportion of participants residing in Zambia (13.63%) and the least in Cameroon (3.11%). Approximately a quarter of the sample (26.89%) had no education, one-fifth of the sample (22.84%) had incomplete primary school and over a half (50.27%) had primary school or greater education. Respondent characteristics and survey items from which the measures of health literacy were derived are reported in Table 1.

Deriving a measure of health literacy (factor analysis)
The internal consistency of the eight survey questions selected to represent health literacy as defined by the NAM was good (Cronbach's $\alpha = 0.72$). The measure of health literacy was comprised of three factors (eigenvalues were >1 for 3 factors) which accounted for 61.20% of the variance (13.29, 34.16, and 14.44%, respectively). The first factor contained primary school education (loading = 0.89), reading a whole sentence or part (loading = 0.87) and reading a magazine (loading = 0.49). The second factor consisted of listening to the radio (loading = 0.76), watching television (loading = 0.51) and learning family planning information from the media (loading = 0.79). The third factor consisted of knowing a place to get condoms (loading = 0.67) and an HIV test (loading = 0.85). Subsequent to deriving one factor and dichotomizing it to be employed as our primary measure of health literacy, we evaluated this variable in relation to the survey questions from which it was obtained. Participants categorized as having high health literacy responded affirmatively to an average of 5.63 of the 8 survey questions, while participants categorized as having low health literacy responded affirmatively to an average of 2.69 of the 8 survey questions. Overall, 35.77% of respondents had high health literacy and 64.23% had low health literacy. Over a third (39.17%) of women had high health literacy and 34.08% of males had high health literacy.

Determinants of health literacy
Respondents categorized as aged 20–24 years of age had the highest proportions of persons with high health literacy (43.95%) while respondents categorized as 45–49 had the lowest proportion of persons who had high health literacy (25.52%). Statistically significant differences were found for all relationships between dichotomized health literacy (high versus low) and respondent characteristics. Figure 1 shows health literacy by respondent characteristics.

Logistic regression models, both crude and adjusted, that examined demographic characteristics associated with health literacy are shown in Table 2. Respondents who were female, within older age groups, completed secondary or greater education, completed partial secondary education, lived in an urban setting, or were in a higher category of wealth were more likely to have high health literacy in models adjusted for potentially influential covariates. Those who were currently married were less likely to have high health literacy in adjusted models than those who were not currently married.

Sensitivity analysis
Our alternative scoring of health literacy (continuously and with different categorizations) showed a similar relationship to respondent characteristics as our binary measure of health literacy (high versus low) when evaluated in regression models. The significance of these relationships was the same as determined with our binary measure.

DISCUSSION
Given a movement towards worldwide collaboration for the assessment of health determinants and outcomes, constructing a robust measure to feasibly examine health literacy in developing countries is essential for the advancement of global health. Recent global initiatives conducted by the NAM and AHQR (Berkman et al., 2011; Institute of Medicine, 2017) highlight a critical need to develop a psychometrically sound measure which may be employed in low and middle lower-income countries. We derived a measure of health literacy using an easily accessible, large, international data source for 14 countries in Sub-Saharan Africa. The measure has acceptable internal consistency and good content validity as it was comprised of the elements defined by the NAM. The construction of this measure will allow for the comparison and assessment of health literacy providing the foundation for interventions to understand and improve health literacy. To the best of our

Table 1: Characteristics of adult DHS respondents between 2006 and 2015 in 14 Sub-Saharan African countries ($n = 224\,751$)

Demographics	%
Female	66.88
Age groups	
15–19	21.92
20–24	18.35
25–29	17.04
30–34	13.81
35–39	12.12
40–44	9.20
45–49	7.57
Marital status	
Not married	50.11
Married	49.89
Urban/rural	
Urban	39.90
Rural	60.10
Country	
Cameroon	3.11
Congo DR	11.24
Ethiopia	12.80
Ghana	5.96
Guinea	3.82
Ivory Coast	6.35
Lesotho	4.18
Namibia	5.67
Niger	5.93
Rwanda	7.35
Sierra Leone	9.96
Swaziland	3.99
Togo	6.00
Zambia	13.63
Educational attainment	
No education	26.89
Incomplete primary	22.84
Complete primary	9.28
Incomplete secondary	28.67
Complete secondary	6.85
Higher	5.47
Wealth index	
Poorest	18.31
Poorer	17.53
Middle	18.72
Richer	20.37
Richest	25.07
Health literacy	
Primary school education	50.50
Able to read whole sentence or part	62.29
Read magazine at least once per week	14.16
Listen to radio at least once per week	46.73
Watch TV at least once per week	31.96

(continued)

Table 1: (Continued)

Demographics	%
Heard family planning information from magazine, radio, or TV in the last few months	44.38
Knows place to get an AIDS test	77.32
Knows place to get condoms	41.26
Knowledge of HIV/AIDS	
Reduce chance of HIV by always using condoms during sex	84.52
Reduce change of HIV by only having one sex partner	89.03
A healthy-looking person can have AIDS	83.04
Cannot get HIV from a mosquito bite	69.19
Cannot get AIDS from witchcraft or supernatural means	78.78
Comprehensive knowledge of HIV/AIDS[a]	68.32

Analysis and results presented based on weighted data.
[a]Variable representing correct knowledge of four out of the 5 variables items in the knowledge of HIV/AIDS sub-group.

knowledge, this is the first study to derive a measure of health literacy across countries Sub-Saharan Africa.

In operationalizing and quantifying the construct of health literacy, it is logical that health literacy varies within a population and that for an individual to benefit from health literacy, a certain level of health literacy must be attained. In other words, a threshold must be surpassed to achieve any notable benefits. Thus, we evaluated health literacy as dichotomous (high or low health literacy) making this construct interpretable and measurable utilizing existing data sources for further work. Given that health literacy was constructed using an objective, weighted and intuitive manner we anticipate and hope that this measure can be replicated for use in other studies. Furthermore, in our assessment of demographic characteristics, this measure was related to hypothesized correlates providing further evidence of its validity and application in real world settings.

A majority of respondents, approximately two out of the three, had low health literacy. Demographic characteristics associated with high health literacy included higher educational attainment, being unmarried, living in an urban setting and having greater wealth. These findings are consistent with prior work, showing that persons who are educated, living in urban areas and are wealthier tend be more literate (Stuebing, 1997; Kickbusch, 2001; Yanhong, 2006). Education, in particular, has been documented as among the strongest predictors of health outcomes. Risk factors, morbidity and mortality across countries of all income levels has shown
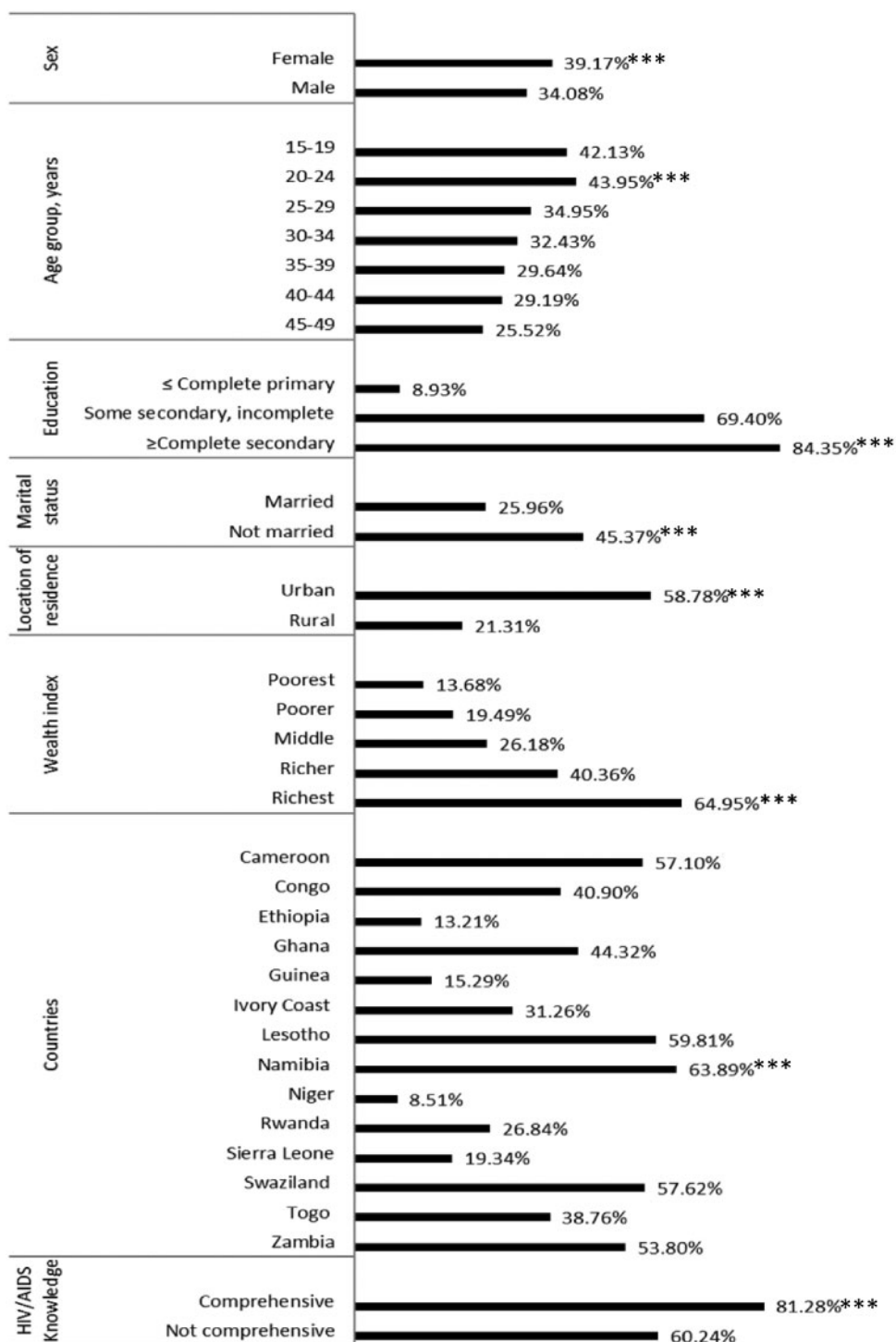
Fig. 1: Health literacy (high) by characteristics in 14 sub-Saharan African countries between 2006 and 2015. ***$p<0.001$ based on chi-square tests of equivalence across subgroups by each demographic category. Note: Variable representing HIV/AIDS knowledge corresponds to four out of the five variables items in the knowledge of HIV/AIDS sub-group in Table 1.

Table 2: Association between demographic characteristics and health literacy in 14 Sub-Saharan African countries from unadjusted and adjusted logistic regression models

	Unadjusted high health literacy OR (95% CI)	Adjusted high health literacy OR (95% CI)
Sex		
Female	0.80 (0.77, 0.83)	1.10 (1.05, 1.16)
Male	Ref	Ref
Age group, years		
45–49	0.47 (0.45, 0.50)	0.96 (0.89, 1.04)
40–44	0.57 (0.54, 0.59)	1.06 (0.98, 1.13)
35–39	0.58 (0.55, 0.61)	1.02 (0.96, 1.09)
30–34	0.66 (0.63, 0.69)	1.07 (1.01, 1.14)
25–29	0.74 (0.71, 0.77)	0.96 (0.92, 1.02)
20–24	1.08 (1.04, 1.17)	0.99 (0.94, 1.04)
15–19	Ref	Ref
Education		
≥Complete secondary	54.93 (51.20, 58.93)	31.50 (29.50, 33.63)
Some secondary complete	23.12 (22.06, 24.22)	18.30 (17.46, 19.18)
≤Complete primary	Ref	Ref
Marital status		
Married	0.42 (0.41, 0.44)	0.78 (0.75, 0.81)
Not married	Ref	Ref
Location of residence		
Urban	5.26 (4.92, 5.63)	1.69 (1.56, 1.82)
Rural	Ref	Ref
Wealth index		
Richest	11.69 (10.73, 12.73)	2.66 (2.42, 2.91)
Richer	4.27 (3.95, 4.62)	1.67 (1.54, 1.81)
Middle	2.24 (2.10, 2.39)	1.26 (1.18, 1.35)
Poorer	1.53 (1.44, 1.63)	1.18 (1.10, 1.26)
Poorest	Ref	Ref

strong relationships with educational attainment. A mother's educational attainment is associated with child well-being and survival. A systematic analysis of 175 countries found that among 8.2 million fewer deaths in children younger than 5 years between 1970 and 2009, 4.2 million (51.2%) could be attributed to increased educational attainment in women of reproductive age (Gakidou *et al.*, 2010). In industrialized nations, the gradient between life expectancy and education has steepened over time (Olshansky *et al.*, 2012). Recent findings suggest that progress towards reductions in mortality in accordance with sustainable development goals (SDGs) has been achieved in parts of Africa, but that many countries are unlikely to achieve goals for 2030 (Golding *et al.*, 2017). Educational initiatives are a critical component of achieving SDGs, however, the role of health literacy, as a distinct and separate construct, could be crucial as well. In this study, while persons with higher levels of education were more likely to be health literate; many persons at higher levels of education had low health literacy. The evaluation of health

literacy may provide additional insight into the mechanisms and factors that influence health and well-being. Targeted interventions focused specifically on health literacy may be needed to result in substantial improvements in health indicators.

In our unadjusted models females, were significantly less likely to have high health literacy than males. However, after adjustment females were more likely to have high health literacy than males. Low levels of education for females, which were then accounted for by adjustment in the analysis likely explain this change. This finding is consistent with other work in which women have higher health literacy than men (Kutner *et al.*, 2006) which may be due in part to the differing nature of their interaction with the health care system. Women may have more comprehensive and frequent engagement with health care provision due to their need for reproductive services related to pregnancy and birth, as well as preventive care such as contraceptives and cervical screening (Sen and Ostlin, 2008). Furthermore, women use more health care services in general, seek

treatment more often, and have higher rates of morbidity than men (Bertakis *et al.*, 2000; Lee *et al.*, 2015). This increased exposure to the health care systems may result in improved health literacy in comparison with men.

This is the first known study to demonstrate acceptable internal consistency and good construct validity of a measure of health literacy derived from the NAM definition using data from a large ongoing survey conducted in multiple countries in Sub-Saharan Africa. An evidence-based review conducted by the AHQR evaluated health literacy interventions and the influence of literacy on health outcomes and disparities (Berkman *et al.*, 2011). It found that over 90% of the studies included in this review based on their objective measurement of health literacy were conducted in high or middle upper income countries highlighting a need for health literacy research in low and middle lower-income countries. With global initiatives recognizing and calling for work to measure and address the issue of low health literacy, our derived measure may provide the foundation for examining both the burden and role of health literacy in low and lower-middle income countries. DHS datasets provide an immediate opportunity to examine health literacy in over 90 countries worldwide enabling researchers to answer many important health literacy-related questions. These include the investigation of the significance and role of health literacy as well as its association with health behaviors and outcomes.

Limitations

The limitations of this work deserve attention. In the absence of a single objective measure for comparison, we sought to develop a measure that could be constructed using existing data sources and was derived from the NAM definition. Given the limited number of items available through the DHS data, our measure may lack precision and applicability across all aspects of health. It should be noted that the content of item is limited in terms of health context (i.e. does not provide information across multiple health behaviors or conditions). In the absence of the ability to collect primary data across multiple countries, this measure can allow research to assess health literacy in countries where little information is currently available. We selected questions from DHS data based on the NAM's definition of health literacy. While our results revealed acceptable internal consistency, there were three factors that arose from the factor analysis differing from the four domains within NAM's definition. We were limited in item selection as the sample was taken from an existing database of items. However, the purpose of the study was to use an existing database to try to capture the concept of health literacy. Researchers who are interested

in differentiating between sub-domains of the NAM definition of health literacy may seek to collect new data that contain additional questions that are more comprehensive. Some researchers have called for a renewed definition of health literacy that would be multidimensional and incorporate both individual- and system-level skills and abilities (Pleasant *et al.*, 2016). Future research should examine the data using multi-dimensional or hierarchical models as well as the inclusion of other items. Our work is also limited in that the data we utilized was not collected with the purpose of the examination of health literacy.

CONCLUSION

This study provides support for using items from an existing international survey (DHS) to capture NAM's definition of health literacy. While our findings suggest that this measure has strong properties, demonstration of its reliability, validity and utility in different regions, cultures and across a range of populations is needed. Application of our measure of health literacy may provide additional insight into the burden of low health literacy. Based on this, modifiable factors may be identified for targeted intervention strategies. Understanding the burden of health literacy and associated factors (e.g. risk behaviors, health outcomes) is critical for implementing appropriate and effective global health promotion initiatives.

ETHICAL APPROVAL

This study was conducted using secondary data that contained no participant identifiers. Survey procedures received ethical approval from country-specific Institutional Review Boards (IRBs) and the Ethics Committee of ICF. Review by the host country IRB ensures compliance with country-specific laws and norms. Review by the ICF IRB ensures compliance with the US Department of Health and Human Services regulations for the protection of human subjects (45 CFR 46).

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Health Promotion International* online.

FUNDING

REFERENCES

Archer, K. J. and Lemeshow, S. (2006) Goodness-of-fit test for a logistic regression model fitted using survey sample data. *The Stata Journal: Promoting Communications on Statistics and Stata*, **6**, 97–105.

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., Viera, A., Crotty, K. *et al.* (2011) Health literacy interventions and outcomes: an updated systematic review. *Evidence Report/Technology Assessment (Full Report)*, (199), 1–941.

Bertakis, K. D., Azari, R., Helms, L. J., Callahan, E. J. and Robbins, J. A. (2000) Gender differences in the utilization of health care services. *The Journal of Family Practice*, **49**, 147–152.

Gakidou, E., Cowling, K., Lozano, R. and Murray, C. J. (2010) Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis. *The Lancet*, **376**, 959–974.

Golding, N., Burstein, R., Longbottom, J., Browne, A. J., Fullman, N., Osgood-Zimmerman, A. *et al.* (2017) Mapping under-5 and neonatal mortality in Africa, 2000-15: a baseline analysis for the sustainable development goals. *The Lancet*, **390**, 2171–2182.

Harman, H. H. (1976) *Modern Factor Analysis*. University of Chicago Press, Chicago, IL.

Hosmer, D. W., Taber, S. and Lemeshow, S. (1991) The importance of assessing the fit of logistic regression models: a case study. *American Journal of Public Health*, **81**, 1630–1635.

Institute of Medicine. (2017) *Roundtable on Health Literacy*. National Academies of Science, Engineering, and Medicine, Washington, DC.

Institute of Medicine Board on Neuroscience and Behavioral Health and Committee on Health Literacy. (2004) *Health Literacy: A Prescription to End Confusion*. National Academies Press, Washington, DC.

Kickbusch, I. S. (2001) Health literacy: addressing the health and education divide. *Health Promotion International*, **16**, 289–297.

Kutner, M., Greenburg, E., Jin, Y., Paulsen, C. and National Center for Educational Statistics and American Institutes for Research. (2006) *The Health Literacy of America's Adults: Results from the 2003 National Assessment of Adult Literacy*, National Center for Education Statistics, Washington, DC.

Lee, H. Y., Lee, J. and Kim, N. K. (2015) Gender differences in health literacy among Korean adults: do women have a higher level of health literacy than men? *American Journal of Men's Health*, **9**, 370–379.

Massey, P. M., Prelip, M., Calimlim, B. M., Quiter, E. S. and Glik, D. C. (2012) Contextualizing an expanded definition of health literacy among adolescents in the health care setting. *Health Education Research*, **27**, 961–974.

Nunnally, J. C. (1989) *Fundamentals of Factor Analysis. Psychometric Theory*. McGraw-Hill, New York, NY.

Olshansky, S. J., Antonucci, T., Berkman, L., Binstock, R. H., Boersch-Supan, A., Cacioppo, J. T. *et al.* (2012) Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Affairs (Project Hope)*, **31**, 1803–1813.

Pleasant, A., Rudd, R. E., O'Leary, C., Paasche-Orlow, M. K., Allen, M. P., Alvarado-Little, W. *et al.* (2016) *Considerations for a New Definition of Health Literacy*. National Academy of Medicine, Washington, DC.

Schrauben, S. J. and Wiebe, D. J. (2017) Health literacy assessment in developing countries: a case study in Zambia. *Health Promotion International*, **32**, 475–481.

Sen, G. and Ostlin, P. (2008) Gender inequity in health: why it exists and how we can change it. *Global Public Health*, **3(Suppl 1)**, 1–12.

Simonds, S. K. (1974) Health education as social policy. *Health Education Monographs*, **2(Suppl 1)**, 1–10.

Stuebing, K. W. (1997) Maternal schooling and comprehension of child health information in urban Zambia: is literacy a missing link in the maternal schooling-child health relationship? *Health Transition Review*, **7**, 151–171.

United Nations. (2015) *United Nations Sustainable Development Goals*. https://www.un.org/sustainabledevelopment/health/ (last accessed 3 July 2018).

United States Agency for International Development. (2006) *Guide to DHS Statistics: Demographic and Health Surveys Methodology*. United States Agency for International Development, Calverton, MD, USA.

Yanhong, Z. (2006) Urban-Rural Literacy Gaps in Sub-Saharan Africa: the Roles of Socioeconomic Status and School Quality. *Comparative Education Review*, **50**, 581–602.

Zimmerman, E. B., Woolf, S. H. and Haley, A. (2015) *Understanding the Relationship Between Education and Health: A Review of the Evidence and an Examination of Community Perspectives. Population Health: Behavioral and Social Science Insights*. Agency for Health-care Research and Quality, Washington, DC, pp. 347–384.