Validation of a health literacy screening tool (REALM) in a UK Population with coronary heart disease

S. Y. Ibrahim1, F. Reid2, A. Shaw1, G. Rowlands1, G. B. Gomez3, M. Chesnokov4, M. Ussher2
1Institute of Primary Care and Public Health, Faculty of Health and Social Care, London South Bank University, 103 Borough Road, London SE1 0AA, UK
2Division of Community Health Sciences, St. George’s, University of London, UK
3Division of Epidemiology, Public Health and Primary Care, Imperial College, London, UK
4Institute of Psychiatry, King’s College London, UK
Address correspondence to Saima Y. Ibrahim, E-mail: ibrahis5@lsbu.ac.uk

ABSTRACT

Background Health literacy (HL) has been recognized as an important public health issue in other developed countries such as the US. There is currently no HL screening tool valid for use in the UK. This study aimed to validate a US-developed HL screening tool (the Rapid Estimate for Adult Literacy in Medicine; REALM) for use in the UK against the UK's general literacy screening tool (the Basic Skills Agency Initial Assessment Test, BSAIT).

Methods A cross-sectional survey involving 300 adult patients admitted to hospital for investigation of coronary heart disease were given the REALM and BSAIT tools to complete as well as specific questions considered likely to predict HL. These questions relate to the difficulty in understanding medical information, medical forms or instructions on tablets, frequency of reading books and whether the participant's job involves reading.

Results The REALM was significantly correlated with the BSAIT (r = 0.70; P < 0.001), and significantly related to seven of the eight questions likely to be predictive of HL.

Conclusions This study has shown that the REALM has face, criterion and construct validity for use as an HL screening tool in the UK, in research and in everyday clinical practice. Further studies are needed to assess the prevalence of low HL in a wider population and to explore the links that may exist between low HL and poor health in the UK.

Keywords health literacy, literacy, patient needs, poor health, screening tools

Introduction

As the UK’s healthcare system develops, there are increased demands for patients to be able to read and understand often complex information in leaflets, medication labels, appointment slips, consent forms and other health-related materials. Those with low literacy are likely to find difficulty in understanding this information.1–3

General literacy has been defined as ‘using printed and written information to function in society, to achieve one’s goals and to develop one’s knowledge and potential’.4 National surveys show that 16% of the population in the UK have low general literacy skills.5 Health literacy (HL) has been recognized as a domain of literacy and can be defined as ‘the degree to which individuals have the capacity to obtain, process and understand basic health information and services needed to make appropriate health decisions’.6 Therefore, HL is more than just the ability to read and write, but relates to the successful application of literacy skills in a health context.

Extensive research in the US has highlighted the impact of low HL both on the individual's health and on the healthcare system. HL has been shown to act as an independent risk factor for poor health.7 Low HL can act as a barrier to achieving or maintaining good health through
difficulties with reading and understanding medical forms and leaflets, making and keeping appointments and with following medication instructions. These difficulties may be exacerbated by the social stigma patients with low literacy experience, which can act as a barrier to gaining social support and advice. Patients with low HL are at a higher risk of poor health and hospitalization resulting in higher healthcare costs.

The Rapid Estimate of Adult Literacy in Medicine (REALM) has been developed in the US as a screening tool for HL. It is a word recognition and pronunciation tool designed for use in healthcare settings and has been shown to correlate highly with measures of HL such as the Test of Functional Health Literacy in Adults (TOFHLA). Although the REALM has been used in the UK before, in three different studies, it has never been validated for use in a UK population. This raises a concern as both language and healthcare systems in the US differ significantly from those in the UK. Validating an HL screening tool in the UK would allow researchers to study the impact of HL and healthcare providers to identify individuals with limited HL, and hence tailor care to patients’ needs.

Low HL is likely to become a particular problem if people become ill or are at an increased risk of illness. For example, those with coronary heart disease (CHD) are expected to read, understand and act on complex health information and medication regimes, as well as consent to invasive investigations. HL has been demonstrated to be an important issue in cardiovascular disease (CVD), but we were unable to find supporting research carried out in the UK focusing on HL and CHD, and this study focuses on this issue. The importance of CHD in public health in the UK makes this disease a suitable choice for investigating HL. This study aimed to validate the REALM for use in the UK in a group of patients with CHD, as a first step to building the evidence base on the impact of low HL on public health in the UK.

Methods

The study entailed a cross-sectional design involving a questionnaire-based survey.

Participants

Consecutive adult patients with CHD admitted to five cardiology wards at a teaching hospital in south London were considered for recruitment to the study. Patients considered to be too ill or distressed to participate by senior ward staff were not approached, and patients who were approached but did not meet the inclusion criteria due to sight impairment, hearing impairment, cognitive difficulties, inability to consult in English or physical inability to write, were excluded by the research nurse. Those <18 years of age were also excluded. Recruitment took place until 300 completed interviews were obtained from November 2005 to January 2007.

Measures

The REALM involves subjects reading out loud a list of 66 medical words arranged in an increasing order of difficulty. The REALM score is calculated by awarding one for each correctly pronounced word and nil for each mispronounced or skipped word. It takes 2–3 min to complete. A score of 59 or less is defined as indicating low HL by the creators of the REALM, while a score of 60 or more indicates adequate HL.

The Basic Skills Assessment Initial Test (BSAIT) (literacy version 1), developed and validated in the UK, provides a measure of general literacy. It takes ~20 min for subjects to self-complete a nine-page workbook of 72 questions. The BSAIT tests reading comprehension using a range of everyday scenarios such as understanding a café menu.

Patient demographics (age, gender, marital status, employment, ethnicity and age of leaving full-time education) were recorded. Ethnicity was coded according to the 2001 Census groups of White, Mixed, Asian, Black and Chinese. Age at leaving full-time education was recoded into those who left school at or below the school leaving age and those who left later, using data on school leaving ages from the Department for Education and Skills (England).

Six additional questions provided variables with which a measure of HL would be expected to correlate. Five of these additional questions were derived from the English arm of the International Literacy Survey conducted in 1996. These questions relate to the difficulty in understanding medical forms or instructions on tablets, frequency of reading books and whether the patient’s job involves reading. The question on needing help to read information or fill in forms from doctors, nurses or hospitals, was designed by the research team but was not validated.

Procedure

Ward staff indicated which patients could be approached by the research nurse, who briefed each patient about the study and gave them the patient information sheet. Written consent was obtained from all patients who agreed to participate. The interview process lasted approximately one hour and took place by the patient’s bedside. Demographic
and other questions were read out loud to the patients and the research nurse recorded the participants’ responses. The REALM was presented on a single sheet and given to participants to read the words out to the research nurse. The research nurse went through the practice questions of the BSAIT workbook with each patient, after which the patients completed the rest of the workbook independently as required by the BSAIT procedure.

Analysis
Data were analysed using the Statistical Package for the Social Sciences (SPSS) software, version 14.0.

Both the criterion and construct validity were assessed. The criterion validity (measurement against a gold standard) was assessed in two ways. First, the REALM was correlated against the BSAIT. As both the REALM and BSAIT scores were negatively skewed, Spearman’s rho was used. Secondly, the REALM was compared with four alternative measures of HL relating to patients’ perception of their ability to read and understand health-related information, and to complete medical forms. The percentage of subjects with low HL (REALM ≤ 59) was compared across the ordered categories of these four measures using the chi-squared test for trend.

The construct validity (the extent to which the test predicted other variables in ways which fit with the underlying construct) was evaluated by assessing the ability of the REALM to predict HL from patients’ reported levels of education, employment status, frequency of reading a book and whether their job involved reading. The percentage of subjects with low HL (REALM ≤ 59) was compared across categorical variables using the chi-squared test (or Fisher’s exact test if numbers were small), and across ordered categories using the chi-squared test for trend.

Results
A total of 300 patients participated in this study (Fig. 1). The majority of the participants were white males, over 60 years and were married or living with a partner (Table 1).

The mean (SD) score for the REALM was 62.1 (6.1) and the lowest score was 19. REALM scores were negatively skewed, with 84 participants (28%) scoring the maximum of 66. At a cut-off point of 59 or lower on the REALM, 19% (n = 57) of the sample were found to be in the low literacy range. REALM scores were not significantly associated with age, gender or marital status (Table 1). However, ethnicity was found to be significantly related to REALM scores (P < 0.001).

Face validity
The REALM requires subjects to recognize and correctly pronounce 66 medical words of increasing complexity, and hence has intuitive face validity as a measure of HL. There was a clear inverse relationship between correct pronunciation and word length, measured by the number of syllables per word, adding further to the credibility of the REALM (Table 2). The number of respondents, who gave a correct response to the ‘worst performing word’, decreased as the number of syllables increased.

Criterion validity
As there is no ‘gold standard’ measure of HL in the UK, the REALM was validated against a general literacy test, the BSAIT. BSAIT scores were negatively skewed. The mean (SD) score for the BSAIT was 65 (7.7) and the lowest score was 29. Forty-three participants scored the maximum of 72 on the BSAIT. Ninety-six subjects (32%) scored 64 or less on the BSAIT, and were therefore classified as having low general literacy. The correlation between the REALM and BSAIT was moderately strong at r = 0.70 (P < 0.001) (Fig. 2).

Participants were asked four questions which provided possible alternative measures of HL. Comparison of REALM scores with patients’ perception of their ability to read and understand health-related information, and ability to complete medical forms showed a statistically significant relationship for three of these four questions. Table 3 shows a trend for decreasing levels of HL with perception of increasing difficulty in reading or completing medical forms and in
understanding medical leaflets. Although lower levels of HL were observed for those who sometimes had difficulty understanding instructions on tablets from the chemist, this association was not significant.

Construct validity
The construct validity was assessed by the association of education, employment status and reading (frequency of reading a book and whether participant’s job involved reading) with REALM scores (Table 4). Higher rates of poor HL were observed for subjects who left school at the minimum leaving age, who are unemployed, whose job does not involve reading and who never or only sometimes read a book. These associations were all statistically significant, and in the expected direction, reflecting the underlying HL construct.

Discussion
Main finding of this study
This study confirms that the REALM is a valid HL screening tool for use in a UK setting. The REALM has face validity as a tool that presents medical words in order of increasing difficulty. Additionally, the REALM has criterion validity.
First, it is highly correlated with a measure of general literacy developed in the UK. The level of correlation (0.70) found is what might be expected for two tests measuring similar but not identical constructs. Secondly, it is statistically related to patients’ perceived difficulty in completion of medical forms and the degree of help required to read or complete medical forms. Finally, the REALM has construct validity as people with low REALM scores are more likely to have left school at the minimum school leaving age, to report ‘reading a book’ less often, to be unemployed and to be in a job that does not involve reading.

Ninety-six participants (32%) of our sample population were detected to have low general literacy as they scored 64 or less on the BSAIT. This contrasts with the most recent findings from the Skills for Life Survey, when 16% of the UK population were detected to have low general literacy. There may be several reasons for this; first, this may reflect the fact that the study population, as might be expected for people with CHD, was older than the general population and that older people have lower literacy skills. Another possibility is that people with lower literacy skills are more likely to have chronic disorders than the general population. This is supported by the findings of the Skills for Life Survey, where people with lower general literacy skills had lower ratings of self-reported general health and were more likely to report chronic illness and disability.

Our study shows a prevalence of low HL, as indicated by the REALM scores, of 19%. If, as in the US, low HL is found to be a predictor of poorer health, low HL is thus likely to be an important public health issue in the UK.

**What is already known on this topic?**

It is known that a large proportion of the population in the UK have low general literacy and numeracy skills. Research in the US has recognized HL as a contributor to poor health. The prevalence of low general literacy and numeracy skills in the UK raises the possibility that low HL may be a public health issue in the UK. The importance of CHD as a
public health issue\textsuperscript{19} makes it an important disease area in which to start investigating the impact of low HL. In the US, the REALM has been validated as a quick and easy screening tool and has been tested across different populations including primary care patients, university clinic patients and prison inmates.\textsuperscript{13} The REALM has also been validated against other US-developed HL measures such as the TOFHLA.\textsuperscript{12–14}

**What this study adds**

We have validated the REALM for use in a UK setting. This has implications for both research and practice. Researchers in the UK can now investigate the links between low HL and a variety of physical and psychological health problems. Investigation of the links between low HL and demonstrable health outcomes, such as the risk of developing CHD and the risk of having poorer disease outcomes, will enable an accurate assessment of both the physical costs of low HL to individuals and families, and the financial costs to the economy through higher hospital admissions and treatment costs.

The speed and ease of the administration of the REALM means that it also has potential as a screening tool that can be used in everyday clinical practice. Screening patients for low HL has some potential advantages in everyday clinical practice; identification of patients with low HL enables them to be given additional time and resources to support them to develop skills and to understand health information. The ease of administration of the REALM (2–3 min) renders it short enough for this purpose. It may be, however, that the stigma associated with low HL might negatively impact patient self-esteem or the patient–practitioner relationship.

Proxy indicator questions, which can be easily applied to large populations on routine registration forms, are potentially a less stigmatizing screening option. In this study we found that employment status, education attainment, reading a book and whether or not an individual’s job involves reading were statistically significantly associated with REALM scores. However, the use of these questions could result in people being ‘labelled’ as having low HL without being aware of being evaluated. Further research is required to assess which measurement options are best for use in everyday practice and to understand more about patients’ views on stigma and acceptability.

**Limitations of this study**

At the time of the study, there were no measures of HL validated for use in the UK. The REALM was therefore validated against the best ‘gold standard’ available, a test of general literacy (BSAIT). It is likely that although there will be a large overlap between health and general literacy, they reflect different skill sets. Indeed it is possible to hypothesize that people with low general literacy skills who develop health problems may develop high HL skills in ‘their’ disease area; this is a possible explanation for the difference in prevalence of low general literacy (32%) and HL (19%) found in our study.

The BSAIT tests a range of skills and presents questions in varying formats and given the age range of this population sample (40–80 years), they may not have been familiar with the style of questions presented to them in the workbook (such as multiple choice questions). The generalization of the findings from this study to other populations is questionable as the focus was on a population with CHD, who, by definition will differ from the general population as they are older and have a chronic disease. Assessment of low HL in the general population, and of the links between low HL and the prevalence and severity of chronic disorders, would require further research.

We found that REALM scores were highly negatively skewed; only seven words wrong out of 66 was sufficient for a classification of ‘low HL’. It can, thus, be argued that many of the 66 words are ‘redundant’ and add nothing to the screening tool. Reducing the length of the REALM to those words with the most power to discriminate between those with low and adequate HL would make the REALM quicker to administer and, thus, more practical for use in everyday clinical practice.

Our study had a high rate of individuals who declined participation (42%), which could have been because of individuals not being able to commit themselves to an hour long interview for reasons such as being discharged, having visitors or awaiting procedures. There is also a concern that participants may have declined to take part in the study because they were unable to read. We therefore feel it is likely that the prevalence of those undergoing investigations of CHD with low HL may be higher than 19%.

Although the REALM will identify individuals at risk of having low literacy, it should be remembered that the REALM is only a screening tool and, thus, not a definitive HL measure. The REALM only tests recognition and pronunciation of medical words and, thus, has limitations as a measure of the full concept of HL.

Since this study was undertaken, the US-developed TOFHLA \textsuperscript{[13]} has been amended for use in the UK and has been shown to have construct validity in a UK population, by being able to predict healthy lifestyle choices (eating, smoking) and better levels of self-reported health.\textsuperscript{22} The TOFHLA, however, takes longer to administer than the REALM (12 versus 3 min), thus potentially reducing its usefulness in research and in practice.
Another measure of HL, also designed in the US, is The Newest Vital Sign (NVS). The NVS presents an American nutrition label which requires individuals to answer specific questions relating to it. It takes 3 min to administer and therefore has potential for use in everyday clinical practice. Unlike the REALM and TOFHLA, it has no ‘ceiling effect’, meaning it can discriminate between different HL levels at the upper end of the skill range. The NVS, thus, has potential as both a research and practice tool; however, it does require validation for use in the UK.

All of the above measures evaluate HL as a basic skill; none measure emerging concepts in empowerment and active citizenship as described by international experts such as Kickbusch and Nutbeam. Further research is needed to develop tools to measure HL in this wider context.

Conclusion

The REALM is now validated for use in the UK as a screening tool to help identify individuals at risk of low HL. This study has shown that the REALM has face, criterion and construct validity for use as a screening tool in research and in everyday clinical practice. Further research in a UK setting is needed into this important public health issue.

Funding

The study was funded by STaRNet London. STaRNet London is funded by the Department of Health.

Acknowledgements

We would like to thank all the staff on participating cardiology wards for allowing us to conduct our study. The local research ethics committee gave their approval to conduct this study.

References