A validation study of the Norwegian version of the Health Literacy Questionnaire: A robust nine-dimension factor model

ASTRID K. WAHL1, ÅSMUND HERMANSEN2, RICHARD H. OSBORNE3 & MARIE HAMILTON LARSEN1

1Department of Interdisciplinary Health Sciences, Institute of Health and Society, University of Oslo, Norway, 2Department of Social Work, Child Welfare and Social Policy, Faculty of Social Sciences, Oslo Metropolitan University, Norway, 3Centre for Global Health and Equity, Faculty of Health, Arts and Design, Swinburne University of Technology, Australia

Abstract
Objective: This study aimed to undertake a rigorous psychometric evaluation of the nine-scale Norwegian version of the Health Literacy Questionnaire (HLQ) based on data from a sample of people with psoriasis. Methods: Cross-sectional survey data were collected from 825 adults with psoriasis who previously participated in the Norwegian Climate Heliotherapy programme. To investigate the factorial validity of the Norwegian HLQ, confirmatory factor analyses were carried out using Stata. Results: A highly restricted model fit with no cross-loadings or correlated residuals was acceptable for three of the nine scales (‘Feeling understood and supported by health-care providers’, ‘Appraisal of health information’ and ‘Ability to find good health information’). After minor model adjustments of the other scales, one-factor models were acceptable. All scales showed acceptable internal consistency, with Cronbach’s alpha ranging from 0.71 to 0.87. Except for three items, all items had high to acceptable factor loadings. Conclusions: This study of the Norwegian HLQ replicates the original factor structure of the Australian HLQ, indicating the questionnaire has cogent and independent scales with good reliability. Researchers, programme implementers and policymakers could use the Norwegian version of the HLQ with confidence to generate reliable information on health literacy for different purposes.

Keywords: Confirmatory factor analysis, Health Literacy Questionnaire (HLQ), psoriasis, psychometric testing, validity testing

Introduction
Fifteen million deaths attributed to non-communicable diseases (NCDs) occur between the ages of 30 and 69 years, and people from all age groups are vulnerable to the risk factors that contribute to NCDs [1]. The UN Sustainable Development Goal 3.4 states that by 2030, the mortality rate from NCDs needs to be reduced by one third [2–4]. Hence, taking appropriate action is an urgent matter in order to reduce premature mortality related to NCDs. The World Health Organization (WHO) points to health literacy (HL) as an essential factor in the prevention and control of NCDs throughout all stages of life [1], and defines HL as ‘The personal characteristics and social resources needed for individuals and communities to access, understand, appraise and use information and services to make decisions about health. Health literacy includes the capacity to communicate, assert and enact these decisions’ [1,5].

Low HL has been reported to be associated with increased mortality, hospitalisation, lower use of preventive health-care services, poor adherence to described medication, difficulty communicating with health professionals and poorer knowledge about disease processes and self-management skills among people with NCDs. Furthermore, research also suggests that strengths and weaknesses in HL abilities may explain observed health inequalities among people of different races and with different educational attainment [6]. These results are based on the use of different HL measures with variation in content and...
psychometric robustness across populations and cultures. However, in order to describe the possible HL challenges within and across groups, cultures and populations accurately, and to develop and evaluate the effect of appropriate interventions, it is crucial to have high-quality HL instruments that have undergone modern and robust psychometric testing [7].

While many different HL measurements exist, a main feature of the measurements is whether they are subjective or objective in scope. In objective measurement, people are challenged by standardised test stimuli to measure an underlying trait. In subjective measurement, people self-report their responses to questions about their experiences and skills. Furthermore, HL measures can be characterised as uni- or multidimensional. However, despite promising recent work developing tools in tandem with definitions, the disconnect between definitions and what the tools measure has been a persistent conceptual stumbling block, which has led to several conundrums that will need to be solved for the field to progress in a coherent manner [7].

There is a need for robust HL instruments to be used for health promotion, the prevention of diseases and in disease management in research and practice. The Health Literacy Questionnaire (HLQ) is a generic, subjective, multidimensional instrument (nine domains) measuring HL, developed in Australia by Osborne et al. [6]. The questionnaire was developed using a ‘validity-driven’ instrument development approach and is based on the definition of HL from the WHO noted above. The tool seeks to detect a wide range of HL needs of people in communities and to be used for a variety of purposes, from describing the HL needs of the population in health surveys through to measuring outcomes of public health and clinical interventions designed to improve HL. The HLQ has been translated and adapted for application in many countries and underlies the work done across WHO National Health Literacy Demonstration Projects (NDPHL) across Europe and in other regions [8].

The HLQ has shown psychometric robustness in the original development process (English language) where nine conceptually distinct areas of HL are defined to assess the needs and challenges of a wide range of people and organisations. Across the European versions of the HLQ (France, Denmark, Germany and Slovakia), confirmatory factor analysis (CFA) has been performed and support the initial nine-factor model [9–13]. These diverse studies with disparate patient populations and health-care systems provide strong evidence that the HLQ captures multiple dimensions of HL. The HLQ was translated and culturally adapted into the Norwegian language following a standardised protocol [14]. Data from application of the HLQ in Norway are now available for validity testing, and the present paper describes the psychometric properties of the Norwegian HLQ using robust CFA procedures in a large population of patients with psoriasis.

Method

Design and sample

Data from a cross-sectional study of 825 adults with psoriasis who had previously participated in the Norwegian Climate Heliotherapy (CHT) programme was used to investigate the psychometric properties of the HLQ. The CHT is one of the therapeutic options available to Norwegian patients with moderate to severe psoriasis. A three-week multidisciplinary programme is provided in the Canary Islands (located in the Atlantic Ocean at 28°N, 16°W), which includes tailored sunlight ultraviolet (UV)B radiation, physical exercise, group discussions and comprehensive education.

A total of 1275 individuals (all patients participating in the CHT programme from 2011 to 2017) were invited to participate in the study. The data collection period was from March to August 2017. Invitations were sent by post, together with a consent form and the survey. At six weeks following the initial post, a reminder letter was sent. Of the 1275 individuals, 825 (65%) returned a completed survey.

In addition to HLQ data, the present paper draws on self-reported data on sociodemographic information (age, sex, marital status and education) and self-reported health status, co-morbidities, number of CHT treatments and duration of the disease. Reports on associations between these variables and HL, based on the same sample, are published elsewhere [15].

The Regional Committee for Medical Research Ethics for Southern Norway (ID 2016/1745) approved the study. In addition, the administrative leaders of the Section for Climate treatment and the Centre for Privacy and Information Security at Oslo University Hospital approved the study. The study was conducted in accordance with the Declaration of Helsinki.

The HLQ

The HLQ comprises 44 items representing nine independent HL domains. For further information on the development and the questionnaire, see Osborne et al. [6]. The domains and content descriptors are summarised in Table I.

Each domain comprises four to six items. The first five domains are scored using response options
indicating the level of agreement, ranging from 1 = strongly disagree to 4 = strongly agree. Four domains report on the capabilities, from 1 = cannot do or usually difficult to 5 = very easy. The domain scores are calculated as the average of the item scores, with higher scores indicating better HL.

**Data analysis**

Analyses were carried out using Stata v16 (StataCorp, College Station, TX). Descriptive analyses were carried out to describe the sample. Further, Cronbach’s alpha was calculated to investigate each scale’s internal consistency.

In order to investigate the factorial validity of the Norwegian version of the HLQ, CFAs were performed. As the structure of the original HLQ has been described previously [6], the factor structure was specified a priori. Consequently, confirmatory analyses were done exclusively. Overall, the amount of missing data per HLQ item was low (1.23–3.13%), with an average of 3.33% of respondents having missing values on one of the nine HLQ scales. CFA models were fitted in Stata v16. For model estimation, maximum likelihood was applied – an estimator that is appropriate for use with ordinal data as is the case with the HLQ. Model evaluation was based on chi-square tests for model fit and further model fit indices, including the root mean square error of approximation (RMSEA), the comparative fit index (CFI), the Tucker–Lewis index (TLI) and the standardised root mean square residual (SRMR). For model fit to be interpreted as ‘acceptable’, a RMSEA of <0.05 was considered a close fit, while a RMSEA and a SRMR of up to 0.08 were considered acceptable. Comparing the fit of a target model to the fit of an independent or null model, the CFI has a cut-off for good fit CFI.
of $\geq 0.90$. A TLI of 0.95 indicates the model of interest improves the fit by 95% relative to the null model, and the cut-off for good fit was set at TLI $\geq 0.95$. Furthermore, the correlations of residuals to improve model fit when fitting the nine one-factor models were considered. Correlated residuals $<0.2$ were considered acceptable when fitting the models [16,17]. Potential model adjustments were based on modification indices as provided in the Stata output using the ‘estat gof, stats (all)’ command.

To obtain a clearer idea of the data and potential problematic items, a full nine-factor model and nine one-factor models were fitted to the data. To test whether modifications, in terms of correlated within-factor residuals, led to significant model improvement, modification indices were obtained using the ‘estat mindices’ command in Stata.

Results
Sample characteristics are shown in Table II. The total sample consisted of 47.3% women. Age ranged from 21 to 83 years, with a mean age of 53.3 years ($SD=12.4$ years). About 61% of respondents had only formal education of up to 13 years, which is the universal benchmark in Norway.

The median disease duration was 27 years, ranging from 0 to 77 years. Mean self-assessed health was 3.33 on a scale ranging from 1 to 5, with a higher number indicating better health. The mean number of co-morbidities was 4.4 ($SD=2.5$).

Nine one-factor models were fitted to the data, and results are shown in Table III. The results are shown when fitting the Norwegian HLQ as a full nine-factor model with no cross-loadings and no correlated residuals, with $\chi^2$ ($866$)=$3351.98$, $p<0.001$; RMSEA=0.063 (90% confidence interval 0.060–0.065); CFI=0.849; TLI=0.835; SRMR=0.071. All factor loadings were high to acceptable (i.e. $>0.5$; see column ‘Standardised factor loading’ in Table III), except for three items which had somewhat low loading: ‘I can rely on at least one healthcare provider’ (0.420) in domain 1 (‘Feeling understood and supported by health-care providers’), ‘I have at least one person who can come to medical appointments with me’ (0.494) in domain 4 (‘Social support for health’) and ‘I ask health-care providers about the quality of the health information I find’ (0.486) in domain 5 (‘Appraisal of health information’; Table III). In order to understand these results further, the authors (A.K.W., Å.H. and M.H.L.) compared the original English items (mentioned above) with the Norwegian items, and examined them again with regard to meaning equivalence at the item and phrase level, with reference to the original English item intent. We could not detect any misunderstandings in the Norwegian wordings compared to the English content. One explanation could be that in Norway, it is somewhat uncommon to ask someone other than your GP for advice and guidance about personal health matters. Consequently, this may have resulted in somewhat lower observed factor loadings.

When fitting the one-factor models, correlated residuals were sequentially added to respective models, which improved each model fit significantly. Table IV shows the results of the CFA separately for the nine individual HLQ scales.

All factor loadings were high to acceptable (i.e. $>0.5$; see column ‘Standardised factor loading’ in Table III), except for three items which had somewhat low loading: ‘I can rely on at least one healthcare provider’ (0.420) in domain 1 (‘Feeling understood and supported by health-care providers’), ‘I have at least one person who can come to medical appointments with me’ (0.494) in domain 4 (‘Social support for health’) and ‘I ask health-care providers about the quality of the health information I find’ (0.486) in domain 5 (‘Appraisal of health information’; Table III). In order to understand these results further, the authors (A.K.W., Å.H. and M.H.L.) compared the original English items (mentioned above) with the Norwegian items, and examined them again with regard to meaning equivalence at the item and phrase level, with reference to the original English item intent. We could not detect any misunderstandings in the Norwegian wordings compared to the English content. One explanation could be that in Norway, it is somewhat uncommon to ask someone other than your GP for advice and guidance about personal health matters. Consequently, this may have resulted in somewhat lower observed factor loadings.

When fitting the one-factor models, correlated residuals were sequentially added to respective models, which improved each model fit significantly. Table IV shows the results of the CFA separately for the nine individual HLQ scales.

While model fit without modifications was acceptable for ‘Feeling understood and supported by health-care providers’, ‘Appraisal of health information’ and ‘Ability to find good health information’, one or more correlated residuals were observed in the remaining six scales. Correlated residuals were $<0.2$ for the majority of the model adjustments, with the exception of one model adjustment (–0.228) in domain 4 (‘Social support for health’), three adjustments...
(0.316, –0.273 and –0.205) in domain 7 (‘Navigating the health-care system’) and one adjustment (0.280) in domain 9 (‘Understanding health information well enough to know what to do’). After respective model adjustments, the one-factor models were acceptable. All nine HLQ scales showed satisfactory to good internal consistency, with Cronbach’s alpha ranging from 0.71 to 0.87 (see Table IV).

**Discussion**

This study of the Norwegian version of the HLQ supports the original Australian nine-dimension HLQ using robust CFA procedures (both full-model and one-model approaches). The model fit (RMSEA and SRMR) for the full nine-factor model is acceptable, although the CFI and TLI is just below 0.85, which is somewhat lower than the original Australian version [6]. However, results showed that after respective model adjustments for some of the scales, the factor models were acceptable. The factor loadings were high to acceptable for almost all items, and the Cronbach’s alpha showed acceptable internal consistency across all scales. Although some of the correlated residuals for model adjustments are just above 0.20, they are small and should not indicate a problem with
item-factor loadings. The proportion of non-response to items was small (<3.13%), suggesting that items were well understood and had acceptable content. Ahead of our study, the Norwegian version was translated, applying rigorous linguistic and cultural adaptation methods in order to produce a high-quality Norwegian version [14,18]. Consequently, the Norwegian version of the HLQ appears to be robust and might serve as a good foundation for the measurement of HL in Norwegian populations.

Our results are also in accordance with similar studies from other countries investigating the HLQ structure. Despite differences in samples and statistical programmes/procedures, the Danish, German, Slovakian and French versions of the HLQ all demonstrated comparable results [9–13] with regard to the nine-factor model. Despite some possible overlap between scale 7 (‘Navigating the health-care system’) and scale 8 (‘Ability to find good health information’) in part 2 of the questionnaire, the HLQ appeared relatively robust. The reason for the good fit found across studies is likely a result of the way the HLQ was developed, translated and adapted [6]. The thorough validity-driven development approach included concept mapping workshops and interviews to identify conceptually distinct domains. Questionnaire items were developed directly from consultation data following a strict process aiming to capture the full range of experiences of people currently engaged in health care through to people in the general population. The psychometric analysis included CFA and item response theory. Cognitive interviews were used to ensure that questions were understood as intended. Items were tested in a calibration sample from community health, home care and hospital settings and then in a replication sample comprising recent emergency department attendees [6].

When adapting the questionnaire into a language-specific context, such as Norwegian, the procedure of translation and adaption has been highly standardised across countries, including transparent translation steps, the use of item intents and cognitive interviews [6,14]. The development of the
Measurement of health literacy

HLQ reported in 2013 gave rise to global interest. In a follow-up study of the psychometric properties of the HLQ including respondents from a diverse range of community-based organisations, the instrument was found to be highly reliable. All HLQ scales were found to be homogenous, and the factor structure of the HLQ was replicated. With a small number of exceptions involving no invariance of factor loadings, strict measurement invariance was established across the participating organisations and sex, language background, age and educational level of respondents [19]. A study evaluating the measurement properties of the HLQ using Rasch analysis among older adults presenting to an emergency department after a fall showed that all nine scales of the HLQ were unidimensional, with good internal consistency [20]. Given this validity-driven approach, the HLQ is likely to be useful in surveys, intervention evaluation and studies of the needs and capabilities of individuals across countries, ages, sex, education and cultures.

Today, the HLQ is used in many studies across countries and health contexts, and gives valuable information on HL from a broader and more in-depth perspective compared to traditional and functional measures of HL [15,21–30]. Globally, directed by the WHO, and in many countries across the world, including Norway, HL has become an important key to solving health challenges in communities and countries. However, to be able to act on HL, policymakers and clinicians need meaningful and valid tools that generate data to identify HL needs, challenges and strengths across settings. Based on the extensive rigorous and ongoing validity testing behind the HLQ, there is mounting evidence that in Norway, the HLQ will provide policymakers, clinicians and researchers with good information for public health and policy decisions. As noted in the introduction, the instrument is useful for both health promotion and disease management settings, since the interpretation of the responses of HLQ can be salutogenic (higher HL is a promotional resource to be strengthened/maintained for health) and pathogenic (lower HL is associated with risks for disease and needs action). Health-care professionals or health workers may use the instrument to explore what HL supports a wide range of the target groups might need.

Validity testing of instruments such as the HLQ is comprehensive and ongoing work, including qualitative and quantitative techniques in order to build evidence. The present study focuses on psychometric properties of the Norwegian version of the HLQ and includes a large number of respondents (N=825), the use of robust CFA procedures and a well-developed original questionnaire. A weakness maybe the homogenous sample, including only individuals with psoriasis. While the patient group may not represent the broad Norwegian population, participants had an average of more than four other co-morbidities and may well be similar to common groups with chronic conditions. The response rate was reasonable high (65%), with good variation with regard to sociodemographic characteristics, and participants were situated in different parts of Norway. Although the present study supports the use of the HLQ in Norway, studies in other NCD groups and in the general population are warranted.

Practice implications

There is mounting evidence that supports the interpretation and use of the HLQ in Western cultures, including Norwegian and other European settings. Our study builds upon growing evidence of the measurement properties of HLQ and that it is a valuable tool for researchers, programme implementers and policymakers to use to explore HL in diverse populations.

Acknowledgements

Our appreciation goes to individuals responding to the questionnaire and to the Department of Climate Therapy at Oslo University Hospital for helping with recruitment.

Declaration of conflicting interests

The authors declared the following potential conflicts of interest with respect to the research, authorship and/or publication of this article: R.H.O. is one of the original authors of the HLQ.

Funding

The authors disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The University of Oslo, the Faculty of Medicine, Institute of Health and Society, Department of Interdisciplinary Health Sciences, Norway, supported this work. R.H.O. was funded in part through a National Health and Medical Research Council (NHMRC) of Australia Principal Research Fellowship #APP1155125.

ORCID iD

Astrid K. Wahl https://orcid.org/0000-0002-5545-5710

References

[1] Mikkelsen B, Williams J, Rakovac I, et al. Life course approach to prevention and control of non-communicable diseases. BMJ 2019;364:l257.
[2] GBD 2017 SDG Collaborators. Measuring progress from 1990 to 2017 and projecting attainment to 2030 of the health-related Sustainable Development Goals for 195 countries and territories: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018;392:2091–2138.

[3] Cao B, Bray F, Ilbawi A, et al. Effect on longevity of one-third reduction in premature mortality from non-communicable diseases by 2030: a global analysis of the Sustainable Development Goal health target. *Lancet Glob Health* 2018 6:e1288–96.

[4] NCD Countdown 2030 Collaborators. NCD Countdown 2030: worldwide trends in non-communicable disease mortality and progress towards Sustainable Development Goal target 3.4. *Lancet* 2018;392:1072–1088.

[5] Sørensen K and Pleasant A. Understanding the conceptual importance of the differences among health literacy definitions. *Stud Health Technol Inform* 2017;240:3–14.

[6] Osborne RH, Batterham RW, Elsworth GR, et al. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC Public Health* 2013;13:658.

[7] Nguyen TH. The state of science of health literacy measurement. *Stud Health Technol Inform* 2017;240:17–33.

[8] Bakker M, Putrik F, Debussche X, et al. Acting together - WHO National Health Literacy Demonstration Projects (NHLDPs) address health literacy needs in the European Region. *Public Health Panorama* 2019;5(2–3):123.

[9] Debussche X, Lenclume V, Balcou-Debussche M, et al. Characterisation of health literacy strengths and weaknesses among people at metabolic and cardiovascular risk: Validity testing of the Health Literacy Questionnaire. *SAGE Open Med* 2018;6:2050312118801250.

[10] Kolarcik P, Cepova E, Madarasova Geckova A, et al. Structural properties and psychometric improvements of the Health Literacy Questionnaire in a Slovak population. *Int J Public Health* 2017;62:591–604.

[11] Friis K, Lasgaard M, Osborne RH, et al. Gaps in understanding health and engagement with healthcare providers across common long-term conditions: a population survey of health literacy in 29,473 Danish citizens. *BMJ Open* 2016;6:e009627.

[12] Mandal HT, Kayser L, Norgaard O, et al. Cultural adaptation and validation of the Health Literacy Questionnaire (HLQ): robust nine-dimensional Danish language confirmatory factor model. *Springerplus* 2016;5:1232.

[13] Nolte S, Osborne RH, Dwinger S, et al. German translation, cultural adaptation, and validation of the Health Literacy Questionnaire (HLQ). *PLoS One* 2017;12:e0172340.

[14] Hawkins M and Osborne RH. *Questionnaire translation and cultural adaptation procedure*. Deakin University, Burwood, Australia, 2010.

[15] Larsen MH, Strumse YAS, Borge CR, et al. Health literacy: a new piece of the puzzle in psoriasis care? A cross-sectional study. *Br J Dermatol* 2019;180:1506–16.

[16] Browne M and Cudeck R. Alternative ways of assessing model fit. In: Bollen K and Long J (eds) *Testing structural equation models*. London: Sage, 1993.

[17] Hu L and Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;6:1–55.

[18] Urstad KH, Andenæs R, Wahl AK, et al. The Health Literacy Questionnaire (HLQ): initial validity testing in a Norwegian sample. *Health Lit Res Pract* 2020.

[19] Elsworth GR, Beauchamp A and Osborne RH. Measuring health literacy in community agencies: a Bayesian study of the factor structure and measurement invariance of the health literacy questionnaire (HLQ). *BMC Health Serv Res* 2016;16:508.

[20] Morris RL, Soh SE, Hill KD, et al. Measurement properties of the Health Literacy Questionnaire (HLQ) among older adults who present to the emergency department after a fall: a Rasch analysis. *BMC Health Serv Res* 2017;17:605.

[21] Freundlich Grygdaard M and Bager P. Health literacy levels in outpatients with liver cirrhosis. *Scand J Gastroenterol* 2018;53:1584–9.

[22] Bourne A, Peerbux S, Jessup R, et al. Health literacy profile of recently hospitalised patients in the private hospital setting: a cross sectional survey using the Health Literacy Questionnaire (HLQ). *BMC Health Serv Res* 2018;18:877.

[23] Cianfrocca C, Caponnetto V, Donati D, et al. The effects of a multidisciplinary education course on the burden, health literacy and needs of family caregivers. *Appl Nurs Res* 2018;44:100–6.

[24] Hosking SM, Brennan-Olsen SL, Beauchamp A, et al. Health literacy in a population-based sample of Australian women: a cross-sectional profile of the Geelong Osteoporosis Study. *BMC Public Health* 2018;18:876.

[25] Goodwin BC, March S, Zajdelwicz L, et al. Health literacy and the health status of men with prostate cancer. *Psychooncology* 2018;27:2374–81.

[26] Hosking SM, Brennan-Olsen SL, Beauchamp A, et al. Health literacy and uptake of anti-fracture medications in a population-based sample of Australian women. *Res Social Adm Pharm* 2018;14:846–50.

[27] Yiu AW and Bajorek BV. Health literacy and knowledge in a cohort of Australian patients taking warfarin. *Pharm Pract (Granada)* 2018;16:1080.

[28] Jessup RL, Osborne RH, Beauchamp A, et al. Differences in health literacy profiles of patients admitted to a public and a private hospital in Melbourne, Australia. *BMC Health Serv Res* 2018;18:134.

[29] Lim S, Beauchamp A, Dodson S, et al. Health literacy and fruit and vegetable intake in rural Australia. *Public Health Nutr* 2017;20:2680–4.

[30] Van Der Giag M, Van Der Heide I, Spreeuwenberg PMM, et al. Health literacy and primary health care use of ethnic minorities in the Netherlands. *BMC Health Serv Res* 2017;17:350.