Appendix 1. A systematic review protocol

Measuring the Quality of Child and Adolescent Health Literacy Instruments: A Systematic Review

Shuaijun Guo[1*], Rebecca Armstrong[1], Elizabeth Waters[1], Thirunavukkarasu Sathish[1], Sheikh M Alif[1], Geoffrey R Browne[1], Xiaoming Yu[2*]

[1] *School of Population and Global Health, The University of Melbourne, Melbourne, Australia*

[2] *Institute of Child and Adolescent Health, School of Public Health, Peking University, Beijing, China*

* Corresponding author email: gshj1986@gmail.com yxm@bjmu.edu.cn

Background

Health literacy research has been a growing interest by researchers across the globe. The term *'health literacy'* was first used in 1974 in the proceedings of a health education conference discussing health education as a social policy issue affecting the healthcare system, mass communication and the education system (1, 2). However, few references were found regarding health literacy in the literature until 1992 (3). Since 1992, health literacy has been broadly studied both in clinical and public health contexts. In clinical settings, health literacy is typically 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*' by the Institute of Medicine (IOM) in America (4). In such circumstances, health literacy is a derivative concept from literacy and numeracy skills, which is often used as a risk factor that needs to be identified and appropriately managed for patients and health professionals (5). Accordingly, health literacy measurement tools and 'screening aids' for clinicians are developed to assess patient literacy levels, and help health professionals to tailor health information for better communication with their patients (6). From the public health perspective, health literacy is defined and accepted by World Health Organization (WHO) as '*the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health*' (7). This understanding of health literacy identifies it as a broad concept, which is seen as a personal asset to enable individuals to take more control over their health and determinants of health (5). With a different understanding of the concept, health literacy measures vary in a different way. Although health literacy measurement varies and is still

being debated (1, 8-10), there is consistent evidence showing health literacy is of potential importance and considered as a public health goal internationally. A recent WHO report pointed out that poor health literacy skills were associated with riskier behaviours, poorer health status, less self-management and longer hospitalization and more health costs (11).

Based on a preliminary search of health literacy, there were more interests in studies focusing on adult health literacy than adolescent health literacy. However, previous research studies suggested that poor health literacy was a prevalent problem in adolescents. In Australia, the 2006 National Health Literacy Survey reported that 67.6% of adolescents aged 15 to 19 years old did not attain the minimum skills required to deal with health information and service in everyday life (12). Compared with adult health literacy, there are several reasons for the potential importance of adolescent health literacy: 1) adolescents are future mainstream and independent healthcare consumers, a health literate person can contribute to less health care costs, better health status compared to that is not health literate (13); 2) adolescents are at a critical stage of development characterised by physical, emotional and cognitive changes, attempting to prepare for independence but lacking the adequate ability of reasoning and decision-making. Therefore, improving their health literacy skills could support sound health decisions in future (14, 15); 3) low health literacy has been demonstrated to associate with high levels of health-risk behaviors (16, 17) and low levels of health-promoting behaviors for adolescents (18); 4) enhancing health literacy through school-based interventions has great potential for improving students' access to and interpretation of health information (19). Adolescents spend most of their daily time in school, which means they can receive health education and learn how to improve healthy lifestyles and related skills through this setting (20, 21).

Health literacy is more challenging to understand for adolescents than that for adults. Researchers may have different understandings and underlying constructs when using the same definition. That is why there are such a large number of measurement tools of health literacy currently (22, 23), along with some newly-developed health literacy instruments (24). According to Mancuso (1), it is recommended to use specific assessment tools for a specific age group in a specific context. Studies measuring childhood and adolescent health literacy have been a research focus, particularly in the past five years (23). Ormshaw *et al.* (23) conducted a systematic review on measuring childhood and adolescent health literacy in 2011. They found 16 studies that were involved with health literacy measures in children and adolescents. The authors also identified 13 health topics and nine underlying components from existing health literacy instruments. However, the authors did not critically appraise health literacy indices explicitly regarding their validity and reliability. More importantly, the

authors did not assess the methodological quality of each included study. This may undermine the persuasiveness of its conclusion. To fill this knowledge gap, we aim to conduct a systematic review that examines studies' methodological quality and examine reliability and validity of each health literacy instrument, thus providing researchers with unbiased information about which instruments have good psychometric properties. The '*COnsensus-based Standards for the selection of health status Measurement INstruments*' (COSMIN) group has recently developed as a critical appraisal tool (a checklist) to evaluate the methodological quality of studies on measurement properties of health measurement instruments (25). These measurement properties are divided into three domains: reliability, validity, and responsiveness (26). According to the COSMIN checklist, it is possible and scientific to critically appraise and compare psychometric properties of health literacy instruments for children and adolescents.

In this protocol, our target population is adolescent. According to the definition of the WHO, adolescents are those people aged 10 to 19 years and young people aged 10-24 years (27, 28). Given that the term '*adolescent*', '*child*', '*youth*' and '*young people*' is closely related, and Erikson (29) reckoned that children between the ages of 6 and 12 years could learn, compete and co-operate with others, we define our target group as those aged 6-24 years old.

Objectives of the review

This review aims to identify which health literacy instruments have good psychometric properties for children and adolescents. Specifically, there are three objectives:

1) To examine the methodological quality of included studies that aim to measure health literacy in children and adolescents;

2) To examine the measurement properties (i.e. reliability; validity; responsiveness) of health literacy instruments in children and adolescents;

3) To compare the overall rating of measurement properties between each health literacy instrument used in children and adolescents.

Search strategy

Database and search terms

As the term '*health literacy*' was first coined in 1974, articles published from 1st,January 1974 to 30th May 2014 in all languages will be searched. Search strategies will be first designed and then be consulted with two librarian experts. Articles indexed in the following seven databases: Medline, Pubmed, Embase, PsycINFO, CINAHL, ERIC and Cochrane

Library will be searched. The search key terms are '*health literacy*' and '*assessment*' according to previously published studies (1, 23, 30, 31). Age group for '*child, adolescent and young adult*' will be defined in the database settings. The synonyms are listed in **Appendix Table 1**. These synonyms are connected by '*or*' and search strategies are completed by '*and*'.

Appendix Table 1 Searching terms in databases

Key term (1)	Key term (2)
health literacy	health literacy measur*
health AND literacy AND education	health literacy assess*
	health literacy evaluat*
	health literacy instrument*
	health literacy tool*

Other sources of literature

Searching other sources to identify relevant research including:

- Reference lists of identified studies;
- Reference lists of previous systematic reviews on health literacy (1, 23, 30-33).

Eligibility criteria for inclusion and exclusion

According to the guidelines recommended by Cochrane Handbook for systematic reviews (34), inclusion criteria will be addressed regarding population, intervention, comparison, outcome and study design (PICOS):

Inclusion criteria-Participants

The target group should be children and/or adolescents, any age from 6 to 24 years of age.

Inclusion criteria-Interventions and Comparators

As interventional studies are not our interest in this review, it is not applicable to set out guidelines for interventions and comparators

Inclusion criteria-Outcomes

The included studies must be involved with health literacy assessment for children and adolescents, that is, the study should specify the term '*health literacy*', and studies are included if they report on at least one or more attributes of the three measurement properties: 1) reliability; 2) validity; and 3) responsiveness.

Inclusion criteria-Study design

The article should be research-based and peer-reviewed paper including study aim, methods, and results. Also, the study aim should focus on health literacy instrument development or validation.

Exclusion criteria

Studies will be excluded if they are: 1) not focusing on the target group; 2) not focusing on the health literacy instrument development or tool validation; 3) not research-based and peer-reviewed papers including editorials, comments and letters; 4) not reporting findings or results regarding any one of the measurement properties.

Study selection

Search records will be kept including the names of databases searched, keywords, search timeframe, and the search results. All the electronic search results will be initially inputted into the bibliography software of EndNote X7 (Thomson Reuters, New York, NY), and other sources of literature results will be summarised in the print paper. This screening process will follow the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement (35). One reviewer will screen studies by titles and abstracts. Secondly, full copies of articles identified will be obtained for thorough screening according to the inclusion criteria by two reviewers independently. Any disagreements in reviewer selections will be resolved at a meeting.

Quality assessment

The methodological quality of each included study will be assessed by two reviewers independently using the COSMIN checklist (25). The checklist consists of nine boxes with 5-18 items concerning methodological standards for how each measurement property should be assessed. Four response options for each item of the COSMIN checklist are defined, representing '*excellent*', '*good*', '*fair*' and '*poor*' quality. An overall score for the methodological quality of a study will be determined for each measurement property separately, by taking the lowest rating of any items in a box ('*worst score counts*') (36). Discrepancies arise between the reviewers will be resolved through discussion, if necessary with a third independent person.

Data extraction

Data extraction will be performed along with the assessment of methodological quality using

the COSMIN checklist (25). In addition, information on the interpretability (e.g. norm scores, floor-ceiling effects, minimal important change of the instruments), generalisability (e.g. characteristics of the study population and sampling procedure), respondent and administrative burden, and forms of administration will be also collected because they are important characteristics of a measurement instrument (26, 37). The data will be entered in an electronic form. Where possible, authors of the original studies will be contacted to obtain essential missing or additional data. Two reviewers will independently extract the data. Consensus should be reached afterward, if necessary with a third independent person.

Data synthesis

The results of the quality of health literacy instruments will be assessed using Terwee's quality criteria (38), to see whether the results of the measurement attributes are '*positive*', '*negative*', or '*indeterminate*'. To summarise the overall ratings of the measurement properties of one health literacy instruments by different authors, the synthesis will be performed by combining the results of the quality of health literacy instruments, the results of methodological quality of health literacy measurement studies and the consistency of their results. The possible overall rating for a measurement property is '*positive*', '*indeterminate*', or '*negative*', accompanied by levels of evidence, similarly as was proposed by the Cochrane Back Review Group (39, 40). One reviewer will perform the data synthesis and a second reviewer will check the synthesised results. Discrepancies of the results will be resolved by discussion.

References

1. Mancuso JM. Assessment and measurement of health literacy: An integrative review of the literature. Nurs Health Sci 2009;11(1):77-89.
2. Simonds SK. Health education as social policy. Health Educ Monogr. 1974;2(1_suppl):1-10.
3. Speros C. Health literacy: concept analysis. J Adv Nurs. 2005;50(6):633-40.
4. Nielsen-Bohlman L, Panzer AM, Kindig DA. Health literacy: a prescription to end confusion. Washington DC, USA: The National Academies Press; 2004.
5. Nutbeam D. The evolving concept of health literacy. Soc Sci Med. 2008;67(12):2072-8.
6. Parker RM, Williams MV, Weiss BD, Baker DW, Davis TC, Doak CC, et al. Health literacy: report of the Council on Scientific Affairs. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. JAMA. 1999;281(6):552-7.
7. Nutbeam D. Health promotion glossary. Health Promot Int. 1998;13(4):349-64.
8. Mancuso JM. Health literacy: a concept/dimensional analysis. Nurs Health Sci. 2008;10(3):248-55.
9. Higgins JW, Begoray D, MacDonald M. A Social Ecological Conceptual Framework for Understanding Adolescent Health Literacy in the Health Education Classroom. Am J Commun Psychol. 2009;44(3-4):350-62.
10. Zarcadoolas C, Pleasant A, Greer DS. Understanding health literacy: an expanded model. Health Promot Int. 2005;20(2):195-203.
11. Kickbusch I, Pelikan JM, Apfel F, Tsouros A. Health literacy: the solid facts. Copenhagen, Denmark: WHO Regional Office for Europe, 2013.
12. Australian Bureau of Statistics. 4233.0-Health Literacy, Australia, 2006. Canberra: Australian Bureau of Statistics; 2008 Jun 25 [cited 2014 Mar 23]; Available from:

http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4233.02006?OpenDocument.
13. Ghaddar SF, Valerio MA, Garcia CM, Hansen L. Adolescent health literacy: the importance of credible sources for online health information. J Sch Health. 2012;82(1):28-36.
14. Bacon JL. Adolescent sexuality and pregnancy. Current Opinion in Obstetrics and Gynecology. 2000;12(5):345-7.
15. Paus T. Mapping brain maturation and cognitive development during adolescence. Trends Cogn Sci. 2005;9(2):60-8.
16. Miles, SB, Stipek, D. Contemporaneous and Longitudinal Associations Between Social Behavior and Literacy Achievement in a Sample of Low-Income Elementary School Children. Child Dev. 2006;77(1):103-17.
17. Conwell, LS, O'Callaghan, MJ, Andersen, MJ, Bor, W, Najman, JM, Williams, G. Early adolescent smoking and a web of personal and social disadvantage. J Paediatr Child Health. 2003;39(8):580-5.
18. Chang LC. Health literacy, self-reported status and health promoting behaviours for adolescents in Taiwan. J Clin Nurs. 2011;20(1-2):190-6.
19. Gazmararian, JA, Curran, JW, Parker, RM, Bernhardt, JM, DeBuono, BA. Public health literacy in America: an ethical imperative. Am J Prev Med. 2005;28(3):317-22.
20. Brey RA, Clark SE, Wantz MS. Enhancing health literacy through accessing health information, products, and services: An exercise for children and adolescents. J Sch Health. 2007;77(9):640-4.
21. Kickbusch I. Health literacy: an essential skill for the twenty-first century. Health Educ. 2008;108(2):101-4.
22. Wu A, Begoray D, Macdonald M, Wharf Higgins J, Frankish J, Kwan B, et al. Developing and evaluating a relevant and feasible instrument for measuring health literacy of Canadian high school students. Health Promot Int. 2010;25(4):444-52.
23. Ormshaw MJ, Paakkari LT, Kannas LK. Measuring child and adolescent health literacy: a systematic review of literature. Health Educ. 2013;113(5):433-55.
24. National Institutes of Health (NIH). Office of Behavior and Social Sciences Reasearch. Research Underway in Health Literacy Supported by NIH. 2014. Bethesda, Maryland: NIH; 2014 [cited 2014 Mar 3]; Available: https://obssr-archive.od.nih.gov/scientific_areas/social_culture_factors_in_health/health_literacy/research-underway-health.aspx.
25. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2010;19(4):539-49
26. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-45.
27. World Health Organization. Health topics: Adolescent health. Geneva, Switzerland: WHO; 2014 [cited 2014 Mar 7]; Available from: http://www.who.int/topics/adolescent_health/en/.
28. World Health Organization Media Centre. Young people: health risks and solutions, Fact sheet N°345. Geneva, Switzerland: WHO Media Centre; 2014 [cited 2014 Mar 7]; Available: http://www.who.int/mediacentre/factsheets/fs345/en/.
29. Erikson EH. Childhood and society. 2nd ed. New York, USA: W. W. Norton & Company; 1963.
30. Jordan JE, Osborne RH, Buchbinder R. Critical appraisal of health literacy indices revealed variable underlying constructs, narrow content and psychometric weaknesses. J Clin Epidemiol. 2011;64(4):366-79.
31. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med. 2011;155(2):97-107.
32. Sanders LM, Federico S, Klass P, Abrams MA, Dreyer B. Literacy and child health: a systematic review. Arch Pediatr Adolesc Med 2009;163(2):131-40.
33. DeWalt DA, Hink A. Health literacy and child health outcomes: a systematic review of the literature. Pediatrics. 2009;124(Supplement 3):S265-S74.
34. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1. 0 [updated Mar 2011]. London, UK: The Cochrane Collaboration; 2011. Available from: www.cochrane-handbook.org.
35. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264-9.
36. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2012;21(4):651-7.
37. Lohr KN. Assessing health status and quality-of-life instruments: attributes and review criteria. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2002;11(3):193-205.
38. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42.

39. Furlan AD, Pennick V, Bombardier C, van Tulder M. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine. 2009;34(18):1929-41.
40. Van Tulder, M, Furlan, A, Bombardier, C, Bouter, L, Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine. 2003;28(12):1290-9.

Appendix 2. Search strategy for seven databases

This section has two parts for SEARCH STRATEGY. The first part focuses on the timeline of 1974 to 2014. The second part focuses on the timeline of May 2014 to Jan 2018.

Part 1:

1 MEDLINE (Web of Science) search strategy

MEDLINE database was searched using the Web of Science interface on 16/05/2014 for the period 1974 to 2014.

Basic search:

Set	Results	
# 1	500	MeSH HEADING: (health literacy) OR ((TITLE: (health literacy) OR MeSH HEADING:exp: (Health Literacy)) AND (TITLE: (education) OR MeSH HEADING:exp: (Educational Status) OR MeSH HEADINGS:exp: (/education) OR MeSH HEADING:exp: (Teaching) OR MeSH HEADING:exp: (Educational Status) OR MeSH HEADING:exp: (Education))) Refined by: MeSH HEADINGS: (ADOLESCENT OR YOUNG ADULT OR CHILD) Indexes=MEDLINE Timespan=1974-2014
# 2	3,880	TOPIC: ((((health) literacy assess* OR health literacy measur*) OR health literacy evaluat*) OR health literacy instrument*) OR health literacy tool*) Indexes=MEDLINE Timespan=1974-2014
# 3	352	#2 AND #1 Indexes=MEDLINE Timespan=1974-2014

2 PubMed search strategy

PubMed database was searched (Advanced search) on 16/05/2014 for the period 1974 to 16/05/2014.

Set	Results	
# 1	4910	Search (health literacy[MeSH Terms]) OR (health AND education AND literacy[Title/Abstract]) Sort by: PublicationDate
# 2	3248385	Search (child* OR adolescent* OR student* OR youth OR young people OR teen* OR young adult[Title/Abstract]) Sort by: PublicationDate Because if we select age group including child, adolescent, and young adult, the newest papers such as published in 2014 will not be included, the reason maybe the database doesn't update properly. So we use these terms to identify.
# 3	1887	Search (health literacy assess* OR health literacy measur* OR health literacy evaluat* OR health literacy instrument* OR health literacy tool*) Sort by: PublicationDate
# 4	581	Search ((((health literacy[MeSH Terms]) OR (health AND education AND literacy[Title/Abstract]))) AND ((health literacy assess* OR health literacy measur* OR health literacy evaluat* OR health literacy instrument* OR health literacy tool*))) AND ((child* OR adolescent* OR student* OR youth OR young people OR teen* OR young adult[Title/Abstract])) Filters: Publication date from 1974/01/01 to 2014/05/16 Sort by: PublicationDate

3 EMBASE (Ovid) search strategy

EMBASE database was searched using Ovid interface on 16/05/2014 for the period 1974 to current.

Using **.mp** as searching terms (Advanced Search):

Set	Results	
#1	6060	("health literacy" or (health and literacy and education)).mp.
#2	6043	limit 1 to yr="1974 -Current"
#3	671	limit 2 to (school child <7 to 12 years> or adolescent <13 to 17 years>)
#4	170	(health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*).mp.
#5	170	limit 4 to yr="1974 -Current"
#6	18	3 and 5

4 PsycINFO (EBSCO) search strategy

PsycINFO database was searched using EBSCO interface on 16/05/2014 for the period January 1974 to May 2014.

Advanced Search:

Set	Results		
#1	786	health literacy OR (health AND literacy AND education)	Limiters - Published Date: 19740101-20140531; Age Groups: School Age (6-12 yrs), Adolescence (13-17 yrs), Young Adulthood (18-29 yrs) Search modes - Boolean/Phrase
#2	133	health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*	Limiters - Published Date: 19740101-20140531; Age Groups: School Age (6-12 yrs), Adolescence (13-17 yrs), Young Adulthood (18-29 yrs) Search modes - Boolean/Phrase
#3	133	(health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*) AND (S1 AND S2)	Search modes - Boolean/Phrase

5 CINAHL (EBSCO) search strategy

CINAHL database was searched using EBSCO interface on 16/05/2014 for the period January 1974 to May 2014.

Advanced Search:

Set	Results		
#1	437	health literacy OR (health AND education AND literacy)	Limiters - Published Date: 19740101-20140531; Age Groups: Child: 6-12 years, Adolescent: 13-18 years Search modes - Boolean/Phrase
#2	63	health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*	Limiters - Published Date: 19740101-20140531; Age Groups: Child: 6-12 years, Adolescent: 13-18 years Search modes - Boolean/Phrase
#3	63	(health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*) AND (S1 AND S2)	Search modes - Boolean/Phrase

6 ERIC (EBSCO) search strategy

ERIC database was searched using EBSCO interface on 16/05/2014 for the period January 1974 to May 2014.

Advanced Search:

Set	Results		
#1	59	health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*	Limiters - Date Published: 19740101-20140531 Search modes - Boolean/Phrase
#2	2,250	health literacy OR (health AND education AND literacy)	Limiters - Date Published: 19740101-20140531 Search modes - Boolean/Phrase
#3	59	S1 AND S2	Search modes - Boolean/Phrase

7 The Cochrane Library search strategy

The Cochrane Library database was searched on 30/05/2014 for the period January 1974 to May 2014.

Set	Results	Sub-database
#1	4	Cochrane Reviews: There are 4 results from 8483 records for your search on 'health literacy in Title, Abstract, Keywords and child* OR adolescent* OR student* OR teen* OR youth OR young adult OR young people in Title, Abstract, Keywords and health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool* in Title, Abstract, Keywords , Publication Date from 1974 to 2014 in Cochrane Reviews'
#2	114	Trials: There are 114 results from 789657 records for your search on 'health literacy in Title, Abstract, Keywords and child* OR adolescent* OR student* OR teen* OR youth OR young adult OR young people in Title, Abstract, Keywords and health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool* in Title, Abstract, Keywords , Publication Date from 1974 to 2014 in Trials'
#3	2	Methods Studies: There are 2 results from 15764 records for your search on 'health literacy in Title, Abstract, Keywords and child* OR adolescent* OR student* OR teen* OR youth OR young adult OR young people in Title, Abstract, Keywords and health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool* in Title, Abstract, Keywords , Publication Date from 1974 to 2014 in Methods Studies'
#4	120	

PART 2:

The above seven databases were searched using similar rationale as describe before for the timeframe of May 17 2014 to Jan 31 2018.

MEDLINE was searched using the Web of Science interface on 17/02/2018 for the period 2014 to 2018.

Basic search:

Set	Results	
# 5	**35**	#4 AND #3 Indexes=MEDLINE Timespan=2014-2018
# 4	**14,198**	MeSH MAJOR TOPIC:exp: ((((((child*) OR adolescent*) OR student*) OR youth) OR young people) OR teen*) OR young adult) Indexes=MEDLINE Timespan=2014-2018
# 3	**1,779**	#2 AND #1 Indexes=MEDLINE Timespan=2014-2018
# 2	**3,482**	((((TOPIC: (health literacy assess*) OR TOPIC: (health literacy measur*)) OR TOPIC: (health literacy instrument*)) OR TOPIC: (health literacy tool*)) ORTOPIC: (health literacy evaluat*)) Indexes=MEDLINE Timespan=2014-2018
# 1	**2,654**	((MeSH HEADING:exp: (health literacy) OR MeSH MAJOR TOPIC:exp: (health literacy)) OR TITLE: (health literacy)) OR MeSH MAJOR TOPIC: ((health) AND education) AND literacy) Indexes=MEDLINE Timespan=2014-2018

Pubmed was searched (Advanced search) on 17/02/2018 for the period 2014 to 31/01/2018.

Set	Results	
#6	**26**	Search (((((((health literacy[MeSH Terms]) OR health literacy[Title/Abstract]) OR (health[Title/Abstract] AND education[Title/Abstract] AND literacy[Title/Abstract])) OR (health[Title/Abstract] AND education[Title/Abstract] AND numeracy[Title/Abstract]))) AND ((health literacy assess*[Title/Abstract] OR health literacy measur*[Title/Abstract] OR health literacy evaluat*[Title/Abstract] OR health literacy instrument*[Title/Abstract] OR health literacy tool*[Title/Abstract]))))) AND ((child*[Title/Abstract] OR adolescent*[Title/Abstract] OR student*[Title/Abstract] OR youth[Title/Abstract] OR young people[Title/Abstract] OR teen*[Title/Abstract] OR young adult[Title/Abstract])) Filters:Publication date from 2014/05/16 to 2018/01/31
#5	48	Search (((((((health literacy[MeSH Terms]) OR health literacy[Title/Abstract]) OR (health[Title/Abstract] AND education[Title/Abstract] AND literacy[Title/Abstract])) OR (health[Title/Abstract] AND education[Title/Abstract] AND numeracy[Title/Abstract]))) AND ((health literacy assess*[Title/Abstract] OR health literacy measur*[Title/Abstract] OR health literacy evaluat*[Title/Abstract] OR health literacy instrument*[Title/Abstract] OR health literacy tool*[Title/Abstract]))))) AND ((child*[Title/Abstract] OR adolescent*[Title/Abstract] OR student*[Title/Abstract] OR youth[Title/Abstract] OR young people[Title/Abstract] OR teen*[Title/Abstract] OR young adult[Title/Abstract]))
#4	288	Search (((((health literacy[MeSH Terms]) OR health literacy[Title/Abstract]) OR (health[Title/Abstract] AND education[Title/Abstract] AND literacy[Title/Abstract])) OR (health[Title/Abstract] AND education[Title/Abstract] AND numeracy[Title/Abstract]))) AND ((health literacy assess*[Title/Abstract] OR health literacy measur*[Title/Abstract] OR health literacy evaluat*[Title/Abstract] OR health literacy instrument*[Title/Abstract] OR health literacy tool*[Title/Abstract]))
#3	288	Search (health literacy assess*[Title/Abstract] OR health literacy measur*[Title/Abstract] OR health literacy evaluat*[Title/Abstract] OR health literacy instrument*[Title/Abstract] OR health literacy tool*[Title/Abstract])
#2	1636528	Search (child*[Title/Abstract] OR adolescent*[Title/Abstract] OR student*[Title/Abstract] OR youth[Title/Abstract] OR young people[Title/Abstract] OR teen*[Title/Abstract] OR young adult[Title/Abstract])
#1	8495	Search (((health literacy[MeSH Terms]) OR health literacy[Title/Abstract]) OR (health[Title/Abstract] AND education[Title/Abstract] AND literacy[Title/Abstract])) OR (health[Title/Abstract] AND education[Title/Abstract] AND numeracy[Title/Abstract])

EMBASE was searched using Ovid interface on 17/02/2018 for the period 2014 to current.

Using **.mp** as searching terms (Basic Search):

Set	Results	
#1	11966	("health literacy" or (health and literacy and education)).mp.
#2	5862	limit 1 to yr="2014 -Current"
#3	639	limit 2 to (school child <7 to 12 years> or adolescent <13 to 17 years>)
#4	372	(health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*).mp.
#5	**26**	3 and 4

PsycINFO was searched using Ovid interface on 17/02/2018 for the period May 2014 to Jan 2018.

Basic Search:

Set	Results	
#1	4331	("health literacy" or (health and literacy and education)).mp.
#2	2077	limit 1 to yr="2014 -Current"
#3	754	limit 2 to (100 childhood <birth to age 12 yrs> or 180 school age <age 6 to 12 yrs> or 200 adolescence <age 13 to 17 yrs> or 320 young adulthood <age 18 to 29 yrs>)
#4	216	(health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*).mp.
#5	**40**	3 and 4

CINAHL was searched using EBSCO interface on 17/02/2018 for the period May 2014 to Jan 2018.

Basic Search:

Set	Results		
S1	health literacy OR ((health AND education AND literacy))	Limiters - Published Date: 20140501-20180131; Age Groups: Child: 6-12 years, Adolescent: 13-18 years Search modes - Boolean/Phrase	View Results (467)
S2	health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*	Limiters - Published Date: 20140501-20180131; Age Groups: Child: 6-12 years, Adolescent: 13-18 years Search modes - Boolean/Phrase	View Results (118)
S3	S1 AND S2	Search modes - Boolean/Phrase	View Results (118)

ERIC was searched using EBSCO interface on 17/02/2018 for the period May 2014 to Jan 2018.

Basic Search:

Set	Results		
S1	health literacy OR ((health AND education AND literacy))	Limiters - Date Published: 20140501-20180131 Search modes - Boolean/Phrase	View Results (292)
S2	health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool*	Limiters - Date Published: 20140501-20180131 Search modes - Boolean/Phrase	View Results (13)
S3	(S1 AND S2)	Search modes - Boolean/Phrase	View Results (13)

Cochrane Library was searched on 17/02/2018 for the period May 2014 to Jan 2018.

Set	Results	Sub-database
#1	2	Cochrane Reviews: There are 2 results from 10210 records for your search on 'health literacy in Title, Abstract, Keywords and child* OR adolescent* OR student* OR teen* OR youth OR young adult OR young people in Title, Abstract, Keywords and health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool* in Title, Abstract, Keywords , Publication Year from 2014 to 2018 in Cochrane Reviews'
#2	199	Trials: There are 199 results from 1121096 records for your search on 'health literacy in Title, Abstract, Keywords and child* OR adolescent* OR student* OR teen* OR youth OR young adult OR young people in Title, Abstract, Keywords and health literacy assess* or health literacy measur* or health literacy evaluat* or health literacy instrument* or health literacy tool* in Title, Abstract, Keywords , Publication Year from 2014 to 2018 in Trials'
#3	201	

Appendix 3. Quality criteria for measurement properties of health literacy instruments

Property	Rating	Quality criteria
Reliability		
Internal consistency	+	(Sub)scale unidimensional AND Cronbach's alpha(s) ≥ 0.70
	?	Dimensionality not known OR Cronbach's alpha not determined
	-	(Sub)scale not unidimensional OR Cronbach's alpha(s) < 0.70
Measurement error	+	MIC > SDC OR MIC outside the LOA
	?	MIC not defined
	-	MIC ≤ SDC OR MIC equals or inside LOA
Reliability	+	ICC/weighted Kappa ≥ 0.70 OR Pearson's r ≥ 0.80
	?	Neither ICC/weighted Kappa nor Pearson's r determined
	-	ICC/weighted Kappa < 0.70 OR Pearson's r < 0.80
Validity		
Content validity	+	The target population considers all items in the questionnaire to be relevant AND considers the questionnaire to be complete
	?	No target population involvement
	-	The target population considers items in the questionnaire to be irrelevant OR considers the questionnaire to be incomplete
Construct validity		
Structural validity	+	Factors should explain at least 50% of the variance
	?	Explained variance not mentioned
	-	Factors explain < 50% of the variance
Hypotheses testing	+	(Correlation with an instrument measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses) AND correlation with related constructs is higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlation with an instrument measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR correlation with related constructs is lower than with unrelated constructs
Responsiveness		
Responsiveness	+	(Correlation with an instrument measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 0.70) AND correlation with related constructs is higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlation with an instrument measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR AUC < 0.70 OR correlation with related constructs is lower than with unrelated constructs

Note: AUC, Area Under the Curve; ICC, Intra-class Correlation Coefficient; LOA, Limits of Agreement; MIC, Minimal Important Change; SDC, Smallest Detectable Change. + positive rating; ? indeterminate rating; - negative rating.

Appendix 4. Levels of evidence for the overall rating of measurement properties

Level	Rating	Criteria
Strong	+++ or ---	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
Moderate	++ or --	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality
Limited	+ or -	One study of fair methodological quality
Conflicting	±	Conflicting findings
Unknown	?	Only studies of poor methodological quality

Note: + positive result; - negative result; ±conflicting result; ? unknown result.

Appendix 5. Reliability and validity results for included instruments

Appendix Table 1. The methodological quality of each study based on reliability for each health literacy instrument

Instrument	Internal consistency		Reliability			
	Result	COSMIN score	Result	Design	Time interval	COSMIN score
NVS (Warsh *et al.*, 2014)	na	na	na	na	na	na
NVS (Driessnack *et al.*, 2014)	α=0.71 (n=47)	Poor	na	na	na	na
NVS (Hoffman *et al.*, 2013)	α=0.67 (n=229)	Poor	na	na	na	na
c-sTOFHLAd (Chang *et al.*, 2012)	α=0.85 (n=300) Item-total correlation=0.44-0.86	Fair	Correlation of test and retest was 0.95 (P<0.001)	Test-retest	1 week	Fair
TOFHLA (Chisolm and Buchanan, 2007)	na	na	na	na	na	na
s-TOFHLA (Hoffman *et al.*, 2013)	α=0.89 (n=229)	Poor	na	na	na	na
REALM-Teen (Davis *et al.*, 2006)	α=0.94 (n=388)	Poor	γ=0.98	Test-retest	1 week	Fair
REALM-Teen (Hoffman *et al.*, 2013)	α=0.92 (n=229)	Poor	na	na	na	
HLAB (Wu *et al.*, 2010)	α=0.92 (n=275) Understanding α=0.88 (n=275) Evaluating α=0.82 (n=275)	Fair	Concordance rate=95%	Inter-rater	na	Poor
MMAHL(Massey *et al.*, 2013)	α=0.83 (n=1208) Item-total correlation=0.39-0.74	Good	na	na	na	na
MHL (Levin-Zamir *et al.*, 2011)	α=0.74 (n=1316) Coefficient of reproducibility=0.84 Coefficients of scalability=0.54-0.80	Poor	na	na	na	na
DNT-39 (Mulvaney *et al.*, 2013)	α=0.93 (n=61)	Fair	na	na	na	na
DNT-14 (Mulvaney *et al.*, 2013)	α=0.82 (n=133) α=0.80 (n=61)	Fair	na	na	na	na

Instrument	Internal consistency		Reliability			
	Result	COSMIN score	Result	Design	Time interval	COSMIN score
	α=0.83 (n=72)					
eHEALS (Norman and Skinner, 2006)	α=0.88 (n=664) Item-scale correlation coefficient=0.51-0.76	Fair	The correlations between administrations ranged 0.68-0.40.	Test-retest	Immediately after the intervention; 3-month; 6-month	Fair
CHC Test (Steckelberg *et al.*, 2009)	na	na	Cohen's Kappa was excellent for 277 ratings (κ=0.9-1.0), moderate or good for 31 ratings (κ=0.7-0.89) and poor for 5 ratings (κ=<0.7)	Inter-rater	na	Poor
HKACSS (Schmidt *et al.*, 2010)	Health knowledge χ^2=6.45, *P*=0.17 (n=852) Health communication α=0.73 (n=852) Health attitudes α=0.57 (n=852)	Excellent	na	na	na	na
HLAT-51 (Harper, 2014)	Goodness of fit statistic was calculated by each domain (CFI=0.33-0.88; TLI=0.66-0.84; RMSEA=0.09-0.17). The internal consistency statistic was not calculated.	Poor	na	na	na	na
HLAT-8 (Abel *et al.*, 2014)	α=0.64 (n=7097 for male) α=0.65 (n=331 for female)	Excellent	na	na	na	na
CHLT (Liu *et al.*, 2014)	α=0.87 (the entire scale); subscales α ranged 0.59 to 0.81	Fair	na	na	na	na
VOHL (Ueno *et al.*, 2014)	na	na	The kappa value of scoring among the dentists ranged from 0.60 tooth score to 0.70 for gingiva score.	Inter-rater	na	Fair
HAS-A (Manganello *et al.*, 2015)	α=0.77 (communication) α=0.73 (confusion) α=0.76 (understanding)	Fair	na	na	na	na
MaHeLi (Naigaga *et al.* 2015)	The person separation index for the original 20-item scale was 0.91 and α=0.92. After	Fair	na	na	na	na

Instrument	Internal consistency		Reliability			
	Result	COSMIN score	Result	Design	Time interval	COSMIN score
	item reduction, the person separation index for 12-item scale was 0.90.					
QuALiSMental (de Jesus Loureiro *et al.*, 2015)	α=0.55-0.72 (component 2 and 3) α=0.44-0.59 (component 4) α=0.60-0.82 (component 5)	Fair	na	na	na	na
FCCHL-AYAC (McDonald *et al.*, 2016)	α=0.73 (FHL) α=0.63 (IHL) α=0.85 (CHL)	Fair	na	na	na	na
ICHL (Smith *et al.*, 2016)	na	na	na	na	na	na
HELMA (Ghanbari *et al.*, 2016)	α=0.93 (the entire scale); subscales α ranged 0.61 to 0.89	Good	The intraclass correlation coefficient was 0.93.	Test-retest	Two weeks	Good
HLSAC (Paakkari *et al.*, 2016)	α=0.93 (the entire scale); subscales α ranged 0.69 to 0.77	Fair	The standardised stability estimate was 0.83.	Test-retest	Two weeks	Fair
REALM-TeenS (Manganello *et al.*, 2017)	α=0.82	Good	na	na	na	na
funHLS-YA (Tsubakita *et al.*, 2017)	α=0.75	Fair	na	na	na	na
HLS-TCO (Intarakamhang *et al.*, 2017)	α=0.70-0.82 for five subscales; KR-20=0.76 for health knowledge scale	Fair	na	na	na	na
HLRS-Y (Bradley-Klug *et al.*, 2017)	α=0.88 (Knowledge) α=0.94 (Self-advocacy/ support) α=0.93 (Resiliency)	Fair	na	na	na	na
p_HLAT-8 (Quemelo *et al.*, 2017)	α=0.74 (the entire scale), subscales α ranged 0.41 to 0.71	Fair	na	na	na	na

Note: na, no information available. CFI, Comparative Fit Index; CHC Test, the Critical Health Competence Test; CHL, Critical Health Literacy; CHLT, Child Health Literacy Test; c-sTOFHLAd, the Chinese version of short-form Test of Functional Health Literacy in Adolescents; DNT, the Diabetes Numeracy Test; eHEALS, the eHealth Literacy Scale; FCCHL-AYAC, the Functional, Communicative, and Critical Health Literacy-Adolescents and Young Adults Cancer; FHL, Functional Health Literacy; funHLS-YA, Functional Health

Literacy Scale for Young Adults; HAS-A, the Health Literacy Assessment Scale for Adolescents; HELMA, the Health Literacy Measure for Adolescents; HKACSS, the Health Knowledge, Attitudes, Communication and Self-efficacy Scale; HLAB, Health Literacy Assessment Booklet; HLAT-8, the 8-item Health Literacy Assessment Tool; HLAT-51, the 51-item Health Literacy Assessment Tool; HLRS-Y, Health Literacy and Resiliency Scale: Youth Version; HLSAC, The Health Literacy for School-aged Children; HLS-TCO, Health Literacy Scale for Thai Childhood Overweight; ICHL, Interactive and Critical Health Literacy; IHL, Interactive Health Literacy; MaHeLi, the Maternal Health Literacy; MHL, the Media Health Literacy; MMAHL, the Multidimensional Measure of Adolescent Health Literacy; NVS, the Newest Vital Sign; p_HLAT-8, Portuguese version of the 8-item Health Literacy Assessment Tool; QuALiSMental, the Questionnaire for Assessment of Mental Health Literacy; REALM-Teen, the Rapid Estimate of Adolescent Literacy in Medicine; REALM-TeenS, the Rapid Estimate of Adolescent Literacy in Medicine Short Form; RMSEA, Root Mean Square Error of Approximation; s-TOFHLA, the short-form Test of Functional Health Literacy in Adults; TLI, Tucker-Lewis Index; TOFHLA, the Test of Functional Health Literacy in Adults; VOHL, the Visual Oral Health Literacy.

Appendix Table 2. The methodological quality of each study based on validity for each health literacy instrument

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	**COSMIN score**	**Results**	**COSMIN score**	**Results**	**COSMIN score**	**Results**	**COSMIN score**
NVS (Warsh *et al.*, 2014)	A panel of heath literacy experts developed the NVS according to previous experience. The NVS was then refined after feedback from patients, interviewers, and data analysts. No target population is involved in item generation.	Poor	na	na	Hypotheses regarding correlation between scores of a comparator instrument of Gray Silent Reading Test (GSRT) and NVS were formulated before data collection. The NVS and GSRT scores were highly correlated (ρ=0.71, p<0.0001). The NVS score increased with child age (ρ=0.53, p<0.0001).	Fair	na	na
NVS (Driessnack *et al.*, 2014)	A panel of heath literacy experts developed the NVS according to previous experience. The NVS was then refined after feedback from patients, interviewers, and data analysts. No target population is involved in item generation.	Poor	na	na	A moderate positive correlation was found between children's NVS scores and their age, and between children's NVS scores and their reports of books numbers (γ_s=0.43, p=0.003; γ_s=0.36, p=0.012, respectively), but not found with their parents' report of the number of children's books at home (γ_s=0.06, p=0.671).	Poor	na	na
NVS (Hoffman *et al.*, 2013)	A panel of heath literacy experts developed the NVS according to previous experience. The NVS was then refined after feedback from patients, interviewers, and data analysts. No target population is involved in	Poor	na	na	Convergent validity was measured between NVS and the TerraNova academic achievement test, with a correlation coefficient of 0.49 (p<0.01).	Fair	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
	item generation.							
c-sTOFHLAd (Chang *et al.*, 2012)	The c-sTOFHLAd was translated from the short-version of TOFHLA according to translation procedures and was tested among 30 adolescents to ensure appropriateness.	Good	Confirmatory factor analysis was conducted to determine structural validity. One-factor model indicated an acceptable fit to the data according structural equation modelling analysis.	Fair	Convergent validity was measured between c-sTOFHLAd and the rapid estimate of adult literacy in medicine (REALM), with a correlation coefficient of 0.74 (p<0.001).	Fair	Semantic equivalence was measured by the content validity index (CVI). All items were rated by the experts as having a CVI>0.85. Thirty adolescents were chosen to determine and ensure the cultural congruence of the instrument.	Fair
TOFHLA (Chisolm and Buchanan, 2007)	The TOFHLA was developed from a literacy expert after reviewing commonly used hospital texts and a pilot test. No target population is involved in item generation.	Poor	na	na	The reading comprehension component was significantly correlated with the WRAT3 and the REALM (ρ=0.59, p<0.001; ρ=0.60, p<0.001 respectively), however, no correlation were found with the numeracy component (ρ=0.11, p=0.45; ρ=0.18, p=0.22 respectively).	Fair	na	na
s-TOFHLA (Hoffman *et al.*, 2013)	The s-TOFHLA was developed based on previous data analysis, perceived importance and frequency of the task in the healthcare settings.	Poor	na	na	Convergent validity was measured between NVS and the TerraNova academic achievement test, with a correlation coefficient of 0.28 (p<0.01).	Fair	na	na
REALM-Teen (Davis *et al.*, 2006)	The REALM-Teen was developed based on a preliminary test and a structured interview among adolescents. And a	Good	na	na	Convergent validity was measured between REALM-Teen and the WRAT-3 (r=0.83) and SORT-R (r=0.93).	Fair	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
	panel of experts reviewed the word list.							
REALM-Teen (Hoffman *et al.*, 2013)	The REALM-Teen was developed based on a preliminary test and structured interview among adolescents. And a panel of experts reviewed the word list.	Poor	na	na	Convergent validity was measured between NVS and the TerraNova academic achievement test, with a correlation coefficient of 0.40 (p<0.01).	Poor	na	na
HLAB (Wu *et al.*, 2010)	Previous experience and literature review were used to develop items; 10 students were pilot-tested for appropriateness of wording, content and format of the final instrument.	Good	na	na	Correlations were assumed between socio-demographic variables and the overall scores. Socio-demographics of gender, age when came to Canada to live, speaking a language other than English were correlated with the scores of HLAB (β=-0.18, p=0.004; β=-0.22, p=0.014; β=-0.20, p=0.008 respectively). No convergent validity is assessed.	Fair	na	na
MMAHL (Massey *et al.*, 2013)	Domains were established from literature review and focus group. Items were developed either using adaptation of existing relevant items or created by the research team.	Good	Explorative principal components factor analysis was conducted and 49.8% of the variance was accounted by 6 factors.	Good	na	na	na	na
MHL (Levin-Zamir *et al.*, 2011)	The face validity was discussed in the focus group during pilot test. The content validity was	Good	na	na	As hypothesised, MHL was associated with socio-economic determinants, particularly with gender (β=1.25, p<0.001) and	Good	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
	analysed using theory and operational definitions of health literacy and media literacy, and adolescents were invited to write detailed, anonymous responses.				mother's education (β=0.16, *p*=0.04). In addition, MHL was also associated with health behaviours (β=0.03, *p*=0.05) and health empowerment (β=0.36, *p*<0.001).			
DNT-39 (Mulvaney *et al.*, 2013)	The DNT-39 was developed from the original 43-item version DNT-43 by eliminating questions specific to type 2 diabetes. An expert team developed the DNT-43 and refined it.	Poor	na	na	The DNT-39 was associated with WRAT-3 and parent education (ρ=0.40, *p*=0.001; ρ=0.29, *p*=0.028 respectively)	Fair	na	na
DNT-14 (Mulvaney *et al.*, 2013)	The DNT-14 was developed from the original 15-item version DNT-15 by eliminating 1 question specific to type 2 diabetes. An expert team developed the DNT-15 by data analysis from DNT-43.	Poor	na	na	The DNT-14 was associated with the Wide-Ranging Achievement Test (WRAT3), parent education, diabetes problem solving and HbA1c (ρ=0.36, *p*=0.005; ρ=0.31, *p*=0.019; ρ=0.27, *p*=0.023; ρ=-0.34, *p*=0.004 respectively)	Fair	na	na
eHEALS (Norman and Skinner, 2006)	The eHEALS was developed by the expert team and pilot-tested and refined by feedback from participants.	Good	Explorative principal components factor analysis was conducted and 56% of the variance was accounted by a single factor. The factor loadings ranged from	Fair	Correlations were assumed between eHEALS and other measured variables (gender, age, use of information technology overall, self-evaluations of health). However, only gender difference was found at baseline level of eHealth literacy	Fair	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
			0.60-0.84 among the 8 items.		(t=2.236, p=0.026). No convergent validity is assessed.			
CHC Test (Steckelberg *et al.*, 2009)	The CHC Test was developed by the research team and pre-tested by collecting qualitative data and quantitative field test.	Good	IRT test for determining dimensionality was performed.	Poor	na	na	na	na
HKACSS (Schmidt *et al.*, 2010)	The HKACSS items were taken from a previous health survey and selected basing on consideration of item content.	Good	na	na	As hypothesised, health communication, attitudes and self-efficacy were significantly related to each other (ρ=0.15-0.38, P<0.05). And children from higher educational background showed a better knowledge and communicated more about health topics (β=0.16, p<0.05).	Good	na	na
HLAT-51 (Harper, 2014)	The expert team evaluated the initial items using a 5-point Likert scale according to their research experience. And 144 college students were invited to complete a pilot test.	Good	Comprehension (CFI=0.80; TLI=0.78; RMSEA=0.09); health numeracy (CFI=0.57; TLI=0.48; RMSEA=0.09); media literacy (CFI=0.88; TLI=0.84; RMSEA=0.07); digital literacy (CFI=0.33; TLI=0.06; RMSEA=0.16); health information seeking (CFI=0.80; TLI=0.66; RMSEA=0.17)	Poor	na	na	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
HLAT-8 (Abel *et al.*, 2014)	The research team developed the HALT-8 drawing on literature review and their own experience. No target population is involved in item generation.	Poor	Explorative principal components factor analysis was conducted and 72.96% of the variance was accounted by four factors among male. In addition, the factor structure was validated using confirmatory factor analysis (CFI=0.99; TLI=0.97; RMSEA=0.03; SRMR=0.03).	Excellent	Hypotheses were formulated a priori regarding correlations between health literacy and gender, socio-cultural characteristics and health values. Results showed that female, higher educational status, and a stronger health valuation were associated with higher HL scores (p<0.05, respectively).	Good	na	na
CHLT (Liu *et al.*, 2014)	The research team developed the CHLT drawing on literature review, expert consultation and pilot test. 12 six graders were piloted about the instrument's readability.	Good	Confirmatory factor analysis was conducted to test the uni-dimensionality of each subscale. The factor loadings ranged from 0.20-0.58.	Fair	Hypotheses were formulated a priori regarding correlations between health literacy and gender, self-reported health and health behaviours. Results showed that female, better health status, normal BMI and healthy behaviours were positively associated with HL scores (p<0.05, respectively). Health-risky behaviours were negatively associated with health literacy scores (p<0.05).	Fair	na	na
VOHL (Ueno *et al.*, 2014)	na	na	na	na	Correlations were conducted between health literacy and gender. Results showed female students had higher gingiva scores than male students	Fair	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
					(p<0.05). However, no gender differences were found regarding tooth scores.			
HAS-A (Manganello *et al.*, 2015)	The research team developed the HAS-A drawing on literature review, expert consultation and pilot test. Scale items were piloted with undergraduates.	Good	Exploratory factor analysis was conducted and 41% of the variance was accounted by three factors.	Fair	Communication scale, confusion scale, and understanding scale were all correlated with the AURA scale (r=0.69, p<0.001; r=-0.50, p<0.001; r=-0.42, p<0.001). The correlation between communication scale, confusion scale and understanding scale and REALM-Teen and NVS were small, ranging from -0.26 to 0.08. Also health literacy scores were compared by demographics. There was no difference in scores by sex or age, but a significant difference by race/ethnicity (p<0.001).	Fair	na	na
MaHeLi (Naigaga *et al.* 2015)	The research team developed the MaHeLi based on the health belief model and integrated model of health literacy. No target population is involved in item generation.	Poor	The health-seeking behaviour (HSB) subscale brought substantial multidimensionality into the MeHeLi scale. After removing most items of the HSB subscale, the MeHeLi scale showed a uni-dimensionality	Fair	na	na	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
			construct with some but not too noticeable multi-dimensionality.					
QuALiSMental (de Jesus Loureiro *et al.*, 2015)	The questionnaire was developed based on mental health literacy framework and adapted among Portuguese adolescents and young people.	Excellent	Exploratory factor analysis was conducted for each component of the questionnaire. A five-factor solution explained 46.84% of the total variance for component 1 and 40.00% for components 2 and 3. A three-factor solution explained 47.24% of the variance for component 4 and a two-factor solution explained 55.63% for component 5.	Fair	The relationship between mental health components and mental health help-seeking intension was examined using a binary logistic regression analysis. Results showed higher levels of mental health literacy tended to associate with mental health-seeking intentions.	Fair	na	na
FCCHL-AYAC (McDonald *et al.*, 2016)	The instrument was adapted from the functional, communicative, and critical health literacy scale to be suitable for adolescents and young adults diagnosed with cancer.	Good	Exploratory factor analysis was conducted for the entire scale. The screen plot suggested the extraction of three factors (53.1% variance explained)	Fair	Health literacy scores were examined by gender, whether the measure was completed online or on paper, whether the participant was on or off treatment. Results showed no significant difference was found.	Fair	na	na
ICHL (Smith *et al.*, 2016)	The instrument was developed from formative interviews with 20 deaf/hard-of hearing high	Good	na	na	The relationship between ICHL and standard health literacy measures were examined. Result showed most ICHL items	Fair	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
	school students. Also the instrument was piloted with 18 individuals including content-expert and content-naïve deaf and hearing colleagues, teachers interpreters and students.				were related to health literacy skills instrument-short form, s-TOFHLA, and comprehensive heart disease knowledge questionnaire (p<0.05).			
HELMA (Ghanbari *et al.*, 2016)	All items were initially generated by in-depth interviews with 67 adolescents. Then items were assessed by an expert panel review and 16 adolescents.	Good	Exploratory factor analysis was conducted and 53.37% of the variance was accounted by eight factors.	Good	na	na	na	na
HLSAC (Paakkari *et al.*, 2016)	The research team developed the HLSAC drawing on literature review, expert review and pilot test. Scale items were piloted with 401 pupils (7^{th} graders and 9^{th} graders).	Good	The five-factor structure was tested using confirmatory factor analysis (RMSEA=0.08; CFI=0.96; TLI=0.92; SRMR=0.03). However, due to high correlations between factors, one-factor structure was finally determined (RMSEA=0.08; CFI=0.94; TLI=0.92; SRMR=0.04).	Fair	Correlations were assumed between the final 10-item scale and the original 15-item scale. Results showed the 10-item HLSAC predicted approximately 97% of the variance of the 15-item instrument.	Fair	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
REALM-TeenS (Manganello *et al.*, 2017)	This instrument was derived from the original 66-item REALM-Teen using the item response theory. Also, ten teenage patients were piloted.	Good	na	na	The REALM-TeenS scores were correlated with the REALM-Teen ($r=0.92$, $p<0.001$). Item fit analysis using the differential item functioning showed the REALM-TeenS functioned well for different groups of sex, race/ethnicity, and language spoken at home.	Good	na	na
funHLS-YA (Tsubakita *et al.*, 2017)	Items were generated from health materials that were frequently used in young adults and reviewed by the research team. No target population was involved in pilot test.	Poor	1-factor model was supported by the exploratory factor analysis.	Fair	The correlation between funHLS-YA and the comparator instrument of functional health literacy was 0.191 ($p<0.001$).	Fair	na	na
HLS-TCO (Intarakamhang *et al.*, 2017)	Items were developed from theories, documents and related research. Also, focus group and expert review were used to develop the instrument. 100 samples of overweight children were piloted.	Good	Confirmatory factor analysis was conducted for each subscale and results showed the model was acceptable, with factor loading ranging 0.39-0.73.	Fair	The path model of health literacy for obesity prevention behaviours was conducted using structural equation modelling. Results showed the hypothetical causal model was consistent with empirical data (chi-square=60.10, p=0.00, df=12, RMSEA=0.05, CFI=0.99; AGFI=0.99).	Fair	na	na
HLRS-Y (Bradley-Klug *et al.*, 2017)	Items were generated by focus group, expert review and a pilot test with 25 participants.	Excellent	Exploratory factor analysis was conducted, and results showed a three-factor structure of the instrument.	Fair	The relationships between health literacy scores and demographics were examined and results showed insurance type and knowledge, time since diagnosis and knowledge and	Fair	na	na

Instrument	Content validity		Structural validity		Hypotheses-testing		Cross-cultural validity	
	Results	**COSMIN score**	**Results**	**COSMIN score**	**Results**	**COSMIN score**	**Results**	**COSMIN score**
					self-advocacy.			
p_HLAT-8 (Quemelo *et al.*, 2017)	The p_HLAT-8 was translated from the HLAT-8 according to translation procedures and was tested among 10 university students to ensure appropriateness.	Good	Confirmatory factor analysis was conducted, and results showed the 4-factor model fit was fair (CFI=0.97, GFI=0.98, TLI=0.95, RMSEA=0.03).	Fair	Convergent validity was examined for each sub-scale, but the results showed that only the factor 'search for information' was adequate. Discriminant validity was only adequate for two factors ('search for information' and 'understanding information').	Fair	Three experts in the field of health forward and backward translated the scale independently. Ten university students were piloted to test and ensure the cultural congruence of the scale. Confirmatory factor analysis showed a 4-factor structure fit the model.	Fair

Note: na, no information available. AGFI, Adjusted Goodness of Fit Index; AURA, Ask, Understand, Remember and Assessment; CFI, Comparative Fit Index; CHC Test, the Critical Health Competence Test; CHLT, Child Health Literacy Test; c-sTOFHLAd, the Chinese version of short-form Test of Functional Health Literacy in Adolescents; DNT, the Diabetes Numeracy Test; eHEALS, the eHealth Literacy Scale; FCCHL-AYAC, the Functional, Communicative, and Critical Health Literacy-Adolescents and Young Adults Cancer; funHLS-YA, Functional Health Literacy Scale for Young Adults; HAS-A, the Health Literacy Assessment Scale for Adolescents; HELMA, the Health Literacy Measure for Adolescents; HKACSS, the Health Knowledge, Attitudes, Communication and Self-efficacy Scale; HLAB, Health Literacy Assessment Booklet; HLAT-8, the 8-item Health Literacy Assessment Tool; HLAT-51, the 51-item Health Literacy Assessment Tool; HLRS-Y, Health Literacy and Resiliency Scale: Youth Version; HLSAC, The Health Literacy for School-aged Children; HLS-TCO, Health Literacy Scale for Thai Childhood Overweight; ICHL, Interactive and Critical Health Literacy; MaHeLi, the Maternal Health Literacy; MHL, the Media Health Literacy; MMAHL, the Multidimensional Measure of Adolescent Health Literacy; NVS, the Newest Vital Sign; p_HLAT-8, Portuguese version of the 8-item Health Literacy Assessment Tool; QuALiSMental, the Questionnaire for Assessment of Mental Health Literacy; REALM-Teen, the Rapid Estimate of Adolescent Literacy in Medicine; REALM-TeenS, the Rapid Estimate of Adolescent Literacy in Medicine Short Form; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual; SORT-R, Slosson Oral Reading Test-Revised; s-TOFHLA, the short-form Test of Functional Health Literacy in Adults; TLI, Tucker-Lewis Index; TOFHLA, the Test of Functional Health Literacy in Adults; VOHL, the Visual Oral Health Literacy; WRAT-3, Wide-Range Achievement Test-Revised.

Appendix Table 3. The methodological quality of each study based on responsiveness for each health literacy instrument

Instrument	Responsiveness	
	Results	**COSMIN score**
VOHL (Ueno *et al.*, 2014)	Comparison of health literacy scores before and after health education showed both tooth and gingiva scores significantly increased after health education.	Fair

Note: As there was only one study examining the instrument's responsiveness, we only presented the instrument of VOHL. VOHL, the Visual Oral Health Literacy.