

**Antecedents and Consequences Associated with Health Literacy or Health Numeracy
in Adult Immigrants and Refugees:
A Systematic Review of Empirical Evidence**

by

Meng Zhu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of Master of Science
in
Public Health and Health Systems

Waterloo, Ontario, Canada, 2018

©Meng Zhu 2018

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Background: Health literacy and health numeracy refer to skills that are fundamental and essential to fulfill health activities. Limited health literacy and numeracy levels have been associated with various poor health outcomes, such as increased emergency department or hospital visits, difficulty in interpreting health messages, inability to take medications properly, and increased mortality in older people. Immigrants and refugees tend to have inadequate health literacy skills, compared to the native-born population, indicating that they may be more susceptible to suffer the negative impacts of low health literacy. Currently, many empirical studies have investigated the antecedents (factors which could influence or predict health literacy or health numeracy, such as demographics) and consequences (outcomes that result from different health literacy or numeracy levels, such as mortality) of health literacy or health numeracy in adult immigrants and refugees. However, efforts that summarize the relevant empirical evidence do not exist.

Objectives: The thesis research aimed to bring together individual empirical studies dispersed in the literature and synthesize both quantitative and qualitative evidence in regards to antecedents and consequences of health literacy or health numeracy in adult immigrants and refugees by applying the systematic review approach. My purpose was not to quantitatively synthesize the results on a specific antecedent or consequence in a specific immigrant or refugee population, but to narratively summarize the relevant evidence to provide information on what antecedents and consequences of health literacy or health numeracy have been investigated in the adult immigrants and refugees, to identify potential research gaps, and to offer insights for future research and practice.

Methods: The reporting of the systematic review mainly followed the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA) checklist. Index terms and free terms relating to health literacy, health numeracy, immigrants and refugees were searched in eight databases, including PubMed, Embase, Cumulative Index to Nursing and Allied Health Literature (CINAHL), PsycINFO, SCOPUS, the ProQuest Dissertations & Theses database, the Conference Proceedings Citation Index-Science (CPCI-S), and the Conference Proceedings Citation Index-Social Science & Humanities (CPCI-SSH). Studies that met the eligibility criteria were included. The reviewers assessed the risk of bias of each individual study and narratively synthesized the extracted evidence by thematic identification. The thematic identification mainly followed the framework approach, in which a working thematic framework was established. The working thematic framework consisted of themes and subthemes pre-existing in Sorensen's comprehensive conceptual model for health literacy and numeracy. Antecedents and consequences from the included studies were coded with these themes and charted into a matrix, the row of which represented an included study, and the column of which denoted to the themes. Additionally, we also carried out inductive thematic identification for the antecedents and consequences that did not fit the pre-existing themes.

Results: In total, 77 included studies, published between 2004 and 2018, were included. Forty-nine of them were quantitative research, out of which 47 applied the cross-sectional study design and two adopted the cohort study design. The rest 28 included studies were qualitative, all of which conducted thematic analysis with information collected by interviews or focus groups. Sixty-seven studies focused on only health literacy; 1 on health numeracy only; and 9 involved both. In terms of the research sample, 63 included studies involved immigrants, 13 investigated refugees, and one involved both. Twenty-four and 58 included studies explored antecedents and

consequences, respectively. The antecedents were coded with the following themes (subthemes): personal antecedents (personal characteristics and personal competence), societal and environmental antecedents, and personal belief, experience, and behaviour. The consequences were coded with themes including health service use or behaviour, health outcome, health cost, health experience and perception, and health knowledge and understanding.

Conclusion: The systematic review was the first study to examine the current state of the research activities on antecedents and consequences of health literacy or health numeracy in the population of adult immigrants and refugees. Our major findings suggest the following recommendations: 1) Future research needs to direct more focus on other health literacy dimensions, such as interactive critical health literacy, rather than the functional dimension of health literacy (i.e., reading and writing skills) in adult immigrants and refugees. 2) More empirical research is needed on antecedents and consequences of health numeracy in adult immigrants and refugees. 3) Inductively identified themes such as “personal belief, experience, and behavior” for antecedents and “personal experience and perception” and “knowledge and understanding” for consequences should be added for the further development of the current framework for health literacy and health numeracy in the context of immigrant and refugee health. 4) More research is required on the associations between health cost and health literacy or numeracy in adult immigrants and refugees. 5) Future quantitative studies need to apply more advanced study designs (e.g., cohort studies) and improve the sampling methods to increase their research validity. 6) Immigrant- and refugee-specific antecedents (e.g., primary language, acculturation, duration of residence in the destination country) and consequences (e.g., health needs) should draw more attention in future empirical research.

Acknowledgements

First of all, I would like to express my great appreciation to my supervisor, Dr. Frank Arocha, for your guidance and support at every stage in the past two years, from the formation of the research questions to the final presentation of this thesis.

Dr. Lorie Donelle and Dr. Mark Oremus, thank you very much for serving as my supervisory committee members. I sincerely appreciate it for your exceptional research expertise and invaluable advice.

Dr. Helen Chen and Dr. Plinio Morita, it was my great pleasure to be a member of your projects. You led me to a new research field where I had little knowledge of before. Thanks for your trust and the precious opportunities.

Dr. Yuan Zhang, thank you a lot for your contribution as an independent reviewer. Your expertise in systematic reviews was sincerely appreciated. You are a great partner.

Ms. Feng He, thank you very much for your help in the title/abstract screening process.

Ms. Dahlia Khajeei, you were always there to listen to and help me. I am forever grateful.

Dr. Hammad Qazi, thank you a lot for inspiring me with your excellent academic advice.

To every faculty, staff, and student of the School of Public Health and Health Systems, thank you for the fabulous community you gave me. It was like a home to like me.

To my parents, my partner, all my families and friends, I am so blessed to have you, your understanding, and support. I would never have achieved all these without you.

A special dedication to my Grandfather, you will be always loved in my heart.

And to God, I will always remember the most important teaching from your only Son, Jesus Christ: love God and love people. I feel so grateful that we will be saved because of Jesus's sacrifice. Please listen to our prayers and grant the path we need to go. To God be all the glory forever!

Table of Contents

Author's Declaration	ii
Abstract.....	iii
Acknowledgements	vi
List of Figures.....	xii
List of Tables	xiii
List of Abbreviations	xiv
Chapter 1 Introduction.....	1
Chapter 2 Literature Review	4
2.1 Immigrants and Refugees.....	4
2.2 Immigrant and Refugee Health	5
2.3 Health Literacy and Health Numeracy	6
2.3.1 Health Literacy.....	6
2.3.1.1 Defining Health Literacy	6
2.3.1.2 Conceptual Frameworks for Health Literacy	8
2.3.2 Health Numeracy	9
2.4 Assessment Instruments for Health Literacy and Health Numeracy.....	11
2.4.1 Assessing Health Literacy	11
2.4.2 Assessing Health Numeracy	14
2.5 Health-Related Outcomes Can Be Influenced by Different Levels of Health Literacy or Health Numeracy.....	16
2.6 Many Factors Can Impact or Predict Health Literacy and Health Numeracy.....	17
Chapter 3 Study Rationale, Research Objectives, and Research Questions	19
3.1 Study Rationale	19
3.2 Research Objectives and Research Questions.....	22
Chapter 4 Methods	23
4.1 Literature Search.....	23
4.1.1 Information Sources	23
4.1.2 Search Terms.....	23
4.2 Study Selection	25
4.2.1 Eligibility Criteria.....	25
4.2.1.1 Inclusion Criteria	25
4.2.1.2 Exclusion Criteria	26
4.2.2 Screening Process.....	27

4.3 Data Extraction	27
4.4 Assessment of Risk of Bias of Included Individual Studies.....	28
4.4.1 Assessment of Risk of Bias of Included Quantitative Studies	28
4.4.1.1 The Newcastle-Ottawa Quality Assessment Scale (NOS) Adapted for Cross-Sectional Studies	28
4.4.1.2 The Risk of Bias in Non-Randomized Studies-of Interventions (ROBINS-I) Assessment Tool (Version for Cohort-Type Studies).....	30
4.4.2 Assessment of Risk of Bias of Qualitative Studies with the Critical Appraisal Skills Program (CASP) for Qualitative Studies.....	31
4.5 Data Synthesis by Thematic Identification.....	32
Chapter 5 Results.....	36
5.1 Results of Literature Search and Selection.....	36
5.2 Characteristics of the Included Studies	37
5.2.1 Basic Information.....	37
5.2.2 Information on Health Literacy or Health Numeracy Investigated in the Included Studies....	38
5.2.3 Information on Adult Immigrants and Refugees Involved in the Included Studies	54
5.3 Assessment of Risk of Bias of Included Quantitative and Qualitative Studies.....	70
5.3.1 Risk of Bias of Quantitative Studies	70
5.3.1.1 Assessment of Risk of Bias of Included Cross-Sectional Studies	70
5.3.1.2 Assessment of Risk of Bias of Cohort Studies	71
5.3.2 Assessment of Risk of Bias of Qualitative Studies.....	71
5.4 Synthesis of the Antecedents and Consequences of Health Literacy or Numeracy Extracted from the Included Studies	74
5.4.1 Themes Identified from Antecedents of Health Literacy or Numeracy.....	74
5.4.1.1 Personal Antecedents.....	74
5.4.1.1.1 Personal Characteristics.....	74
5.4.1.1.2 Personal Competence.....	80
5.4.1.2 Societal and Environmental Antecedents	81
5.4.1.3 Personal Belief, Experience, and Behaviour.....	82
5.4.2 Themes Identified From Consequences of Health Literacy or Health Numeracy	84
5.4.2.1 Health Service Use or Behaviour	84
5.4.2.2 Health Outcome	87
5.4.2.3 Health Cost	88
5.4.2.4 Health Experience and Perception	88
5.4.2.5 Health Knowledge and Understanding	90
5.5 Summary.....	91

Chapter 6: Discussion	92
6.1 Discussion of the Study Design and Sampling Methods	92
6.2 Understanding the Concept of Health Literacy in the Included Studies	94
6.3 Understanding the Concept of Health Numeracy in the Included Studies	98
6.4 Implications of the Thematic Identification	100
6.5 Strengths and Limitations of the Systematic Review.....	103
6.6 Conclusion	105
References.....	107
Appendices.....	121
Appendix A. Combination of Search Terms and Operators in PubMed.....	121
Appendix B. Combination of Search Terms and Operators in Embase	122
Appendix C. Combination of Search Terms and Operators in CINAHL	123
Appendix D. Combination of Search Terms and Operators in PsycINFO.....	124
Appendix E. Combination of Search Terms and Operators in SCOPUS.....	125
Appendix F. Combination of Search Terms and Operators in the ProQuest Dissertations & Theses Database.....	126
Appendix G. Combination of Search Terms and Operators in the CPCI-S and CPCI-SSH	127
Appendix H. Data Items Extracted from Quantitative Studies	128
Appendix I. Data Items Extracted from Qualitative Studies.....	129
Appendix J. The Newcastle-Ottawa Quality Assessment Scale Adapted for Cross-Sectional Studies Applied in the Thesis Research.....	130
Appendix K: Risk of Bias Assessment for the Included Cross-Sectional Quantitative Studies.....	133
Appendix L. The Risk of Bias in Non-Randomized Studies-of Interventions (ROBINS-I) Assessment Tool (Version for Cohort-Type Studies).....	138
Appendix M. Risk of Bias Assessment for the Included Cohort Quantitative Studies.....	160
Appendix N. The Critical Appraisal Skills Program (CASP) for Qualitative Studies	161
Appendix O. Assessment of Risk of Bias for the Included Qualitative Studies.....	165
Appendix P. The Working Thematic Framework Containing Pre-Defined Themes Based on Sorensen’s Et Al., (2012) Conceptual Framework for Health Literacy and Numeracy.....	170
Appendix Q. Extracted Basic Information from the Included Studies.....	171
Appendix R. Descptions of the Assesment Instruments Used in the Included Studies	175
Appendix S. The Matrix Used to Summarize Antecedents of Health Literacy or Health Numeracy	182
Appendix T. The Matrix Used to Summarize Consequences of Health Literacy or Health Numeracy	187
Appendix U. Summary of Key Findings Regarding Antecedents of Health Literacy or Numeracy from the Included Studies	196

**Appendix V. Summary of Key Findings Regarding Consequences of Health Literacy or Numeracy
from the Included Studies 206**

List of Figures

Figure 1. Flow Diagram Depicting the Results of Literature Search and Selection	36
Figure 2. Distribution of the Included Studies by Year of Publication	37

List of Tables

Table 1. Index Terms Searched in the Systematic Review	24
Table 2. Free Text Terms Searched in the Systematic Review	25
Table 3. Information on Health Literacy and Health Numeracy in the Included Studies	40
Table 4. Information on Immigrants and Refugees from the Included Studies	56

List of Abbreviations

95% CI	95% Confidence Interval
AMA	American Medical Association
ANOVA	Analysis of Variance
B	Unstandardized Beta
β	Standardized Beta
BMI	Body Mass Index
CASP	Critical Appraisal Skills Program
Chew Items	Brief Health Literacy Screener
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CPCI-S	Conference Proceedings Citation Index-Science
CPCI-SSH	Conference Proceedings Citation Index-Social Science & Humanities
HALS	Health Activities Literacy Scale
HIE	Healthy Immigrant Effect
HLS-EU-Q16	Short Language Version of the European Health Literacy Questionnaire
HPV	Human Papillomavirus
IOM	Institute of Medicine
IQR	Inter-quartile Range
MSES	Mathematics Self-Efficacy Scale
NOS	Newcastle-Ottawa Quality Assessment Scale
NVS	Newest Vital Sign
OR	Odds Ratio

QoC	Quality of Care
PRISMA	Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
REALM	Rapid Estimate of Adult Literacy in Medicine
ROBINS-I	Risk of Bias in Non-Randomized Studies-of Interventions
SD	Standard Deviation
SE	Standard Error
SNS	Subjective Numeracy Scale
S-TOFHLA	Short Test of Functional Health Literacy in Adults
TOFHLA	Test of Functional Health Literacy in Adults
WHO	World Health Organization

Chapter 1 Introduction

In Canada, the immigrant population is defined as individuals “*who are, or who have been, landed immigrants or permanent residents in Canada. Such persons have been granted the right to live in Canada permanently by immigration authorities*” (Statistics Canada, 2017) (page. 7). Basically, the Canadian immigration authorities admit the immigrant population via two major categories: immigrants and refugees (Statistics Canada, 2017).

Canada is an immigration country. Statistics from the 2016 Census showed that the immigrant population has accounted for more than one-fifth (about 21.9%) of the Canadian total population (Statistics Canada, 2017). The immigrant population can bring working-age people and population growth, which may benefit countries such as Canada, to maintain the economic success and overcome the rapid population aging (United Nations Department of Economic and Social Affairs, 2017).

However, the immigrant population is often considered to be vulnerable in terms of maintaining and promoting health and at increased risk for poor health outcomes and inadequate health care due to numerous challenges and obstacles they have to face in their daily lives in a new society (Derose et al., 2007). Kreps et al., (2008) proposed that immigrants often encounter significant comprehension difficulties, cultural barriers, and economic challenges to accessing and understanding relevant health information. In terms of empirical evidence, being an immigrant or refugee has been statistically significantly associated with inadequate health literacy in several studies (Copelj et al., 2011; Lo et al., 2006; Maneze et al., 2016; J. L. Smith et al., 2003; Yin et al., 2009).

Health literacy, according to Sorensen et al., (2012), is “*linked to literacy and entails people’s knowledge, motivation and competences to access, understand, appraise, and apply*

health information in order to make judgments and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course” (page 3). Health numeracy is usually conceptually subsumed under health literacy. However, Golbeck et al., (2005, 2011) argued that health numeracy should also be viewed as an distinct construct of health literacy and defined health numeracy as *“the degree to which individuals have the capacity to access, process, interpret, communicate, and act on numerical, quantitative, graphical, biostatistical, and probabilistic health information needed to make effective health decisions”* (page 1).

According to the conceptual framework for health literacy and health numeracy developed by Sorensen’s et al., (2012), antecedents and consequences are considered as two essential components. Antecedents refer to factors that could impact or predict health literacy or health numeracy level (e.g., age (Sorensen et al., 2015) and income (van der Heide et al., 2013)), whereas consequences refer to outcomes that result from different levels of health literacy or health numeracy (e.g., hospital readmission (McEwing et al., 2017) and mortality in older adults (Baker et al., 2007)).

The research described in the thesis follows the systematic review approach to synthesize quantitative and qualitative empirical evidence to examine what antecedents and consequences of health literacy or health numeracy have been investigated so far in the adult immigrant and refugee populations. The aim of the thesis research is not to provide quantitative synthesis but to narratively synthesize the relevant evidence to inform researchers about the current state of the research activities, existing research gaps, and promising research directions.

This thesis is presented in six sections: This chapter (Chapter 1) is an introduction of the thesis. Chapter 2 includes a literature review on the topic of immigrants and refugees, health

literacy, and health numeracy. Chapter 3 contains the information on the study rationale, research objectives, and research questions. Chapter 4 provides a detailed description of the research methods. Chapter 5 presents the results, and Chapter 6 provides a discussion and lists some implications of this work.

Chapter 2 Literature Review

2.1 Immigrants and Refugees

The major difference between immigrants and refugees in the 2016 Census of Canadian population is determined by the immigration category (immigrants or refugees) via which the individuals were admitted to Canada in the immigration process. However, according to the Canadian Council of Refugees, the key difference between immigrants and refugees is the former make their own choices to permanently settle in the destination country, whereas the latter forcefully escape from their country of birth due to the fear of persecution (Canadian Council for Refugees, 2010).

By 2016, immigrants have represented about 21.9% of the Canadian total population (Statistics Canada, 2017). The Canadian government plans to accept about 310,000, 330,000 and 340,000 newcomers in 2018, 2019, and 2020, respectively (Government of Canada, 2017). The immigrant population is an important force that profoundly impact the host region or country. For instance, positive net immigration (the number of international migrants arriving at a given country is more than number of international migrants leaving the given country) is one critical contributor to slowing down the trend of population aging in the host countries because immigrants usually include a large proportion of working-age people, which contrasts with the aging native-born population (United Nations Department of Economic and Social Affairs, 2017).

2.2 Immigrant and Refugee Health

New immigrants often exceed the native-born population in terms of physical health and health behaviours. This phenomenon has been repeatedly observed and is termed as the “healthy immigrant effect” (HIE) (Chen et al., 1996; Kennedy et al., 2015; Vang et al., 2015). The HIE could be due to the selection process conducted by the destination country (e.g., medical examination), or by immigrant self-selection effects, i.e., people with sufficient financial and physical means are the ones most likely to emigrate (Halli et al., 2005; Kennedy et al., 2015).

However, the HIE diminishes and even tends to disappear among immigrants who have been residing in the host country for over 10 years (McDonald et al., 2004; Vang et al., 2015). To illustrate, the HIE for cancer is likely to deteriorate among immigrants with additional years spent in Canada (McDonald et al., 2017). Similarly, the rates of becoming overweight or obese among immigrants to Canada gradually increase over the years staying in Canada, despite substantially lower weights upon arrival to Canada (McDonald et al., 2005). In addition, the risk of developing diabetes in some subgroups of immigrants to Canada is also higher than in the Canadian-born population (Adhikari et al., 2012).

Refugees also exhibit the HIE in some aspects (Gushulak et al., 2011). For example, compared with the Canadian-born population, the risks of mortality and cancer were lower in refugees (DesMeules et al., 2005). However, the maternal and infant health in refugees were worse than the Canadian-born population (Gagnon et al., 2013). Moreover, refugees may lose their health advantage more quickly, compared to people who were admitted as immigrants (Beiser, 2005; Newbold et al., 2009).

Generally speaking, immigrants and refugees are being at increased risk for poor health outcomes (Derose et al., 2007). Their vulnerability in terms of maintaining and promoting health

may come from several sources, among which there are poor health literacy and health numeracy levels.

2.3 Health Literacy and Health Numeracy

2.3.1 Health Literacy

Literacy refers to a set of skills, including the capabilities of writing, reading, speech, speech comprehension, and basic mathematics (Kirsch, 2001). The field of health literacy, first introduced in the 1970s, can be considered as a product of the intersection of literacy and health (Green, 2007; Simonds, 1974).

2.3.1.1 Defining Health literacy

Sorensen et al., (2012), by examining 17 existing definitions of health literacy and capturing the essence of them, provided an “all-inclusive” definition, which was described in Chapter 1.

Other frequently cited definitions of health literacy included definitions provided by the World Health Organization (WHO) in 1998, the American Medical Association (AMA) in 1999, and the Institute of Medicine (IOM) in 2004 (Sorensen et al., 2012). WHO concluded that health literacy is “*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*” (Nutbeam, 1998) (page 357); the AMA defined health literacy as “*the constellation of skills, including the ability to perform basic reading and numeral tasks required to function in the health care environment*” (Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs AMA, 1999) (page 553); the IOM defined health literacy 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.” (Institute of Medicine (US) Committee on Health Literacy, 2004) (page 20).

The above definitions are similar in that they state that health literacy is a set of skills possessed by individuals. However, the definitions also differ on many aspects. First of all, the definitions reflect two distinct understandings of health literacy. Nutbeam, (2008) proposed that health literacy could be viewed as either a “clinical risk factor” or a “personal asset.” From the former perspective, limited health literacy is a potential risk factor for poor health outcomes, and it needs to be identified and dealt with in clinical contexts. From the latter perspective, oriented to health promotion, health literacy focuses on helping individuals to improve health literacy skills to maintain or promote health. Among the three definitions cited above, the WHO definition reflects the “personal asset” perspective, whereas the other two reflect the “clinical risk factor” perspective. Additionally, the three definitions state different goals that should be achieved by health literacy: the WHO definition considers the goals as health maintenance and promotion, while the AMA and the IOM definitions concern an individual’s competence in healthcare environments and health decision-making, respectively.

To define health literacy, we also need to involve health contexts under consideration. In “Health Literacy: A Prescription to End Confusion”, the IOM stressed that health literacy refers to those skills which individuals apply to fulfill the demands of various types of health contexts. Health literacy acts as a “bridge” connecting individuals and health contexts (Institute of Medicine (US) Committee on Health Literacy, 2004). Reflecting the similar points of view, Rootman et al., (2008) defined health literacy as “*the ability to access, understand, evaluate and communicate information as a way to promote, maintain and improve health in a variety of settings across the life course*” (page 11).

2.3.1.2 Conceptual Frameworks for Health literacy

Researchers have developed several conceptual frameworks for health literacy (e.g., Institute of Medicine (US) Committee on Health Literacy, 2004; Nutbeam, 2000; Squiers et al., 2012; von Wagner et al., 2009b). Sorensen et al., (2012) examined and integrated 12 existing conceptual frameworks of health literacy and presented an integrated and comprehensive framework. Due to the comprehensiveness, this framework is considered valuable in terms of serving as a guide to understand health literacy in the thesis. Therefore, in this section, I will focus on Sorensen's et al., (2012) integrated framework.

Health literacy, according to Sorensen et al., (2012), consists of four core components: a) the specific individual skills which health literacy refers to, b) the antecedents of health literacy, c) the health contexts wherein health literacy applies, and d) the consequences of health literacy.

First, the integrated framework proposed that health literacy was an individual skill set, involving the ability to access, to understand, to analyze, and to apply health information (Sorensen et al., 2012).

Second, in the integrated framework, antecedents refer to factors that could impact or predict health literacy levels. They were classified into personal factors (e.g., age, gender, educational attainment), situational factors (e.g., support from family), and societal and environmental factors (e.g., policy) (Sorensen et al., 2012).

Third, the contexts refer to the domains where individuals with diverse social, personal, and situational characteristics apply their health literacy skills. In the integrated framework (Sorensen et al., 2012), three major contexts are specified: (a) the context of health care; (b) the context of disease prevention; (c) the context of health promotion.

Finally, consequences refer to health related outcomes which result from different levels of health literacy in the integrated framework (Sorensen et al., 2012). Health literacy could lead to consequences both at the individual and population levels. The individual-level consequences, for example, can include people's use of health services, health behaviour and performance, and personal empowerment. The population-level consequences, on the other hand, can relate to equity and sustainability in public health and health costs to society.

2.3.2 Health Numeracy

There are several definitions of numeracy. Statistics Canada, (2008), on its webpage, defined numeracy as “*the knowledge and skills required to effectively manage the mathematical demands of diverse situations.*”

The concept of numeracy is often conceptually subsumed under the umbrella of literacy. In this sense, numeracy is also called “quantitative literacy” (Schwartz et al., 1997). However, some argued that numeracy can be addressed as a separate construct from literacy as research findings have suggested they may be independent. For instance, Carreiras et al., (2015) showed that processing of numerical and prose information involves different neural pathways in different regions of the brain.

Health numeracy consists of a set of numeracy skills used in health contexts (Reyna et al., 2009). Similar to the conceptual relations between numeracy and literacy, health numeracy is often subsumed under the concept of health literacy. However, Golbeck et al., (2011) showed that health literacy measured by the the Short Test of Functional Health Literacy in Adults (S-TOFHLA) reading comprehension component and health numeracy assessed by the S-TOFHLA numeracy component do not correlate in terms of understanding health information. Based on the results, they suggested researchers differentiate health numeracy from health literacy at the

conceptual level, but include a numeracy component in practice of evaluating the health literacy level of participants to recognize and address the needs of participants.

The definition of health numeracy given by Golbeck et al., (2005) (see Chapter 1) emphasizes that being numerate in health activities requires individuals not only to understand numeric concepts, but also to communicate and take actions effectively based on the concepts. Golbeck et al., (2005) also classified the degree of health numeracy into four functional categories. The first category is basic health numeracy, referring to basic skills to deal with numbers under the condition that no manipulation of numbers is required. The next level, called computational health numeracy, requires individuals to perform simple manipulations and arithmetic calculations. Analytical health numeracy, the third level, asks for higher levels of literacy skills so that individuals can make sense of numerical information. The fourth level, named statistical health numeracy, mainly refers to the capacity to understand basic statistics such as proportions, percentages, and probabilities. We should notice that these four categories are not mutually exclusive, but overlapping.

Ancker et al., (2007) tried to interpret health numeracy from the perspective of “the productive use of quantitative information for health” – beyond the levels of individual skills. They argued that the level of individual skills is only one of many variables that could affect the productive use of quantitative health information. Other variables might include the information artifact design and the communication skills of health providers.

A number of conceptual frameworks suitable for health numeracy exist in the literature. Three of them were discussed previously: Sorensen’s et al., (2012) integrated framework, A. L. Golbeck’s et al., (2005) framework, and Ancker’s et al., (2007) framework. Other than these three, Reyna et al., (2009) also proposed a framework for health numeracy from the perspective

of risk reduction and medical outcomes. They argued that numeracy played critical roles in a) the interactions between the patients, information, and environment, b) the processes of making judgements and decisions, c) the individual's health behaviours, d) and the final medical outcomes. Another framework, discussing health numeracy in the context of decision making, was based on potential effects of health numeracy in six critical domains, including computation, information seeking and processing, sense-making of numbers, assessment of value and probabilities, ability to critically think about numerical data, and behaviour changes (I. M. Lipkus et al., 2009).

2.4 Assessment Instruments for Health Literacy and Health Numeracy

To address health literacy and health numeracy in research and practice as well as understand their role in people's health, it is important for researchers and practitioners to determine the levels of health literacy and numeracy. Thus, the development, validation, and application of assessment instruments for health literacy or numeracy are necessary and essential.

2.4.1 Assessing Health Literacy

There are over 100 health literacy assessment tools (Nguyen et al., 2015)., About half of the current health literacy assessment tools include the evaluation of numeracy (Haun et al., 2014).

Altin et al., (2014) described some characteristics of the assessment tools for health literacy recently developed between 2009 to 2014. One of the important features is that these assessment tools often involve mixed measurement approaches. For example, many of the tools determine people's health literacy level by two modes of measurement: (a) requiring participants to complete tasks and (b) asking participants to self-report their perceptions of their own health

literacy competence. A second important feature is that about one third of recently developed assessment tools base themselves on the formats of existing instruments, for instance, the Rapid Estimate of Adult Literacy in Medicine (REALM) and the test of functional health literacy in adults (TOFHLA).

Referencing established instruments is a common way to develop new tools. However, such new tools may inherit the drawbacks possessed by the established tools. Some researchers emphasized the importance of developing new assessment tools without reference to established tools (Pleasant et al., 2011). Regardless of the arguments, it will be useful and insightful to review the two frequently referenced, established assessment instruments – the REALM and the TOFHLA.

The REALM (Davis et al., 1993; Davis et al., 1998) and the TOFHLA (Parker et al., 1995) are likely the most widely used tools to assess health literacy. The REALM is a word recognition test consisting of 66 common medical terms that requires the participants to read the words aloud when taking the test. The TOFHLA is a cloze-type comprehension assessment tool (i.e., a reading comprehension activity in which participants are required to fill in the blanks with the words omitted from a passage), which contains 50 items to test individuals' capacity to make sense of selected materials and fill in missing words. Notably, the TOFHLA is also able to measure numeracy skills with an additional 17 numeracy items, while the REALM does not. The TOFHLA assumes that individuals with some health literacy could understand the materials they read, while the REALM is developed based on the important role of reading skills in health settings.

Comprehension tests generally are more time-consuming, compared to word recognition tests. For instance, it usually takes 22 minutes to complete the TOFHLA, but only three minutes

to finish the REALM (Davis et al., 1993; Parker et al., 1995). This suggests that the REALM might be more practical in clinical settings to determine the patient's health literacy level. However, in the context of academic research, the TOFHLA is preferred because it measures comprehension ability, which is not part of the REALM.

Davis et al., (1993) estimated the validity of the REALM by correlating the REALM scores with the scores from three standardized reading cognition tests – the Wide Range Achievement Test-Revised, the Slosson Oral Reading Test-Revised, and the reading recognition section of the Peabody Individual Achievement Test-Revised. The Pearson's correlation coefficients were 0.88, 0.96, and 0.97, respectively. However, some researchers questioned the validity of the REALM in the determination of individual health literacy levels. Dumenci et al., (2013), using logical analysis, showed that the REALM barely covered the three dimensions of health literacy: comprehension, numeracy, and information seeking. However, the authors agreed that the REALM was still useful in terms of measuring reading and pronunciation ability. The validity coefficients of the TOFHLA, obtained through comparisons with the Wide Range Achievement Test-Revised and the REALM, were 0.74 and 0.84, respectively; the Spearman-Brown equal-length coefficient and the Cronbach's α measure of internal consistency, which represent the test-retest reliability of the TOFHLA, were 0.92 and 0.98, respectively (Parker et al., 1995).

2.4.2 Assessing Health Numeracy

Researchers have been assessing numeracy objectively (the objective measure) or subjectively (the subjective measure). The former focuses on the evaluation of one's actual numeracy abilities by mathematical tests (e.g. performing arithmetical operations, interpreting percentages and likelihoods, and making inferences or decisions according to numerical information). The latter, on the other hand, turns attention to the assessment of individuals' perceptions on their numeracy competence (e.g. evaluating individuals' feelings about how confident and comfortable they are to deal with numbers without asking them to perform any mathematical tasks) (W. Nelson et al., 2008; Reyna et al., 2009).

Most research on numeracy measurement has employed objective measures (Eklund, 2012). However, Fagerlin et al., (2007) discussed some practical drawbacks of objective assessments. The most important one is that participants may not accept mathematical tasks and get frustrated or intimidated by objective numeracy tests. The subjective measure, on the other hand, is believed to offer a less stressful way to assess numeracy levels (Reyna et al., 2009).

Investigations comparing subjective and objective measures are ongoing, and the findings are still under debate. One study showed that subjective measures increased the willingness of patients at high risk of developing breast cancer to pay for BRCA1/2 genetic testing (Miron-Shatz et al., 2014). Another study suggested that the subjective measure of numeracy, without requiring the participants to take mathematical tests, still could reach a predictive ability similar to that of an objective measure (Zikmund-Fisher et al., 2007). However, some researchers have raised the caution that a test containing a single subjective numeracy measure could not match the predictive ability of an objective numeracy test (W. L. Nelson et al., 2013). Furthermore, Dunning et al., (2004) pointed out that individuals tended to make poor judgements about their

own abilities and that self-assessment was often biased by many factors, such as over confidence and unrealistic optimism. Two examples that are commonly used to measure health numeracy are the TOFHLA and the Newest Vital Sign (NVS) (Parker et al., 1995; Weiss et al., 2005).

The 17 numeracy items in the TOFHLA evaluate numeracy skills in the context of health tasks commonly encountered, e.g., monitoring blood sugar, understanding appointment times, and using medications according to prescriptions (Parker et al., 1995). However, the application of the TOFHLA in the assessment of numeracy might be limited. One reason is that some investigations validating the TOFHLA chose comparison instruments that did not contain a recognized measure of numeracy, thus not validating the numeracy part of the TOFHLA (Reyna et al., 2009).

In the NVS, test takers need to answer six questions after viewing a label involving nutrition information (Weiss et al., 2005). This instrument evaluates people's reading comprehension skills (e.g., read and identify numbers in the nutrition label), quantitative reasoning ability (e.g., deciding what math to use and what numbers might be relevant to the question asked), and arithmetical computational capacity (e.g., performing the arithmetic operations). It only takes about three minutes to complete this test, thereby suitable for a quick screening. Weiss et al., (2005) reported that the Cronbach's α measure of internal consistency for the NVS was 0.76. As for the validity of the NVS, the coefficient (Pearson r) between the scores obtained from the NVS and the TOFHLA was 0.59.

2.5 Health-related Outcomes Can Be Influenced by Different Levels of Health Literacy or Health Numeracy

Sorensen's et al., (2012) named the outcomes that are related to health literacy or numeracy levels as consequences of health literacy or numeracy.

Generally speaking, individuals with limited health literacy or health numeracy often show worse health outcomes, compared to those with higher health literacy or numeracy. For example, an analysis based on the Canadian data from the international adult literacy and skills survey indicated that Canadian individuals with level 1 health literacy skills (the lowest level) were about 2.5 times more likely to be in poor health than the individuals with level 4 and level 5 health literacy skills even after controlling factors such as age, gender, educational level, primary language as well as immigrant status (Statistics Canada, 2005; The Canadian Council on Learning, 2008). There is also a clear trend that the prevalence of diabetes and blood pressure decrease as the average health literacy scores increase (The Canadian Council on Learning, 2008).

Health literacy and health numeracy could also play critical roles in health consequences such as health status and quality of life, evaluation of online health information, heart failure knowledge and salt knowledge, mortality of seniors, following instructions to take medications properly, understanding health messages correctly, emergency department and hospital visits, immunization of influenza, and use of mammography (Berkman et al., 2004; Berkman et al., 2011; Cajita et al., 2016; Diviani et al., 2015).

In specific populations such as immigrants and refugees, health literacy and health numeracy could also be crucial. For instance, limited health literacy was associated with poor self-reported health and having difficulties in seeking healthcare among refugees to Sweden

(Wangdahl et al., 2018). Additionally, immigrants with higher health literacy reported better quality of care than those with low health literacy skills (Calvo, 2016). We further explored this issue through this systematic review project.

2.6 Many Factors Can Impact or Predict Health Literacy and Health Numeracy

The factors which can impact or predict health literacy or health numeracy, according to Sorensen's et al., (2012) integrated framework, are also called antecedents.

Currently, various antecedents such as age, gender, and educational attainment have been reported in the body of literature. Sorensen et al., (2015) found that the proportion of people who possessed limited health literacy was substantially higher in the older age subgroup (58.2% in people between 66 and 75 years and 60.8% in people > 75 years) than that in the total sample (47.6%). The 2003 US National Assessment of Adult Literacy survey showed gender and race differences in health literacy and numeracy: females and white persons possessed a higher average level of health literacy than males and black persons (Kutner et al., 2006).

The relations between educational attainment and health literacy and numeracy levels are complex. Higher health literacy was reported to correlate with higher educational attainment (Kutner et al., 2006). However, educational level appears not to be a determinant of numeracy level. Studies have indicated that even well-educated people often lacked the necessary numeracy skills to make appropriate recommendations (Johnson et al., 2014) or accurately assess numeric risk information (I.M. Lipkus et al., 2001). As for personal competence, poor cognitive ability and low self-efficacy were often correlated with low health literacy or health numeracy levels (S. Gatobu et al., 2014; Levinthal et al., 2008; von Wagner et al., 2009a).

Poor socioeconomic status such as low income has been linked to low health literacy or numeracy levels in some empirical studies (e.g., (van der Heide et al., 2013)). Cultural factors like first language and proficiency in the official language of the host country are also two important factors that predict the health literacy and numeracy levels of immigrants. Speakers of Mandarin, which contains complex numeric concepts, showed better performance in numeric tasks than speakers of Kikuyu, a language that contains relatively simple numeric concepts (S. Gatobu et al., 2016). Several studies also showed that low English proficiency could predict a limited health literacy in immigrants (Becerra et al., 2017; Jacobson et al., 2016; H. Y. Lee et al., 2012).

In the thesis project, we explored the issues on antecedents and consequences of health literacy and health numeracy in immigrants and refugees by conducting a systematic review.

Chapter 3 Study Rationale, Research Objectives, and Research Questions

3.1 Study Rationale

By 2016, the immigrant and refugee population in Canada has accounted for about 22% of the total Canadian population (Statistics Canada, 2017). Given the large population size, the health of immigrants and refugees may have significant implications for the Canadian society and health systems.

Health literacy and health numeracy are known as important determinants of health. Limited health literacy and numeracy have been linked to various poor health outcomes, such as increased hospitalization and mortality (Berkman et al., 2004; Berkman et al., 2011).

In terms of health literacy and health numeracy in immigrants and refugees, the picture is worrying. First of all, a large proportion of immigrants and refugees do not possess sufficient health literacy and health numeracy skills. For instance, a recent Canadian cross-sectional study, which surveyed a sample of 22,818 respondents, showed that about 76% of the immigrants to Canada could not reach the adequate level of health literacy and numeracy (Ng et al., 2014). Second, compared to their native-born peers, immigrants and refugees tend to be less health literate and numerate (Copelj et al., 2011; Lo et al., 2006; Maneze et al., 2016; J. L. Smith et al., 2003; Yin et al., 2009). The above facts strongly suggest that the inadequate health literacy and numeracy issue in immigrants and refugees should be seriously considered, and empirical research on health literacy and health numeracy in the immigrant and refugee population is crucial and necessary.

Antecedents and consequences are important components of health literacy and numeracy. Currently, there are a number of empirical studies investigating them in immigrants and refugees.

For instance, in some quantitative studies, antecedents, such as economic status (Becerra et al., 2017), education (Wangdahl et al., 2014), acculturation (S. E. Choi et al., 2013), and primary language (S. Gatobu et al., 2014), have been shown to be associated with health literacy or health numeracy among immigrants or refugees. In some qualitative research, experiences prior to arrival at destination country was perceived by some refugees as a barrier to health literacy (McMichael et al., 2009).

In terms of empirical studies on consequences, limited health literacy and numeracy levels were associated with low participation in healthcare systems in immigrants (Calvasina et al., 2016). Among refugees, inadequate health literacy was shown to be associated with poor self-reported health and having difficulties in seeking healthcare (Wangdahl et al., 2018). In some qualitative studies, health literacy was considered important for health activities, such as getting involved in obesity prevention (Renzaho et al., 2017) and understanding of HIV/AIDS (Shedlin et al., 2004).

However, no efforts exist to bring together the individual empirical studies dispersed in the literature and examine the current state of the research activities on the antecedents and consequences of health literacy or health numeracy in adult immigrant and refugees. This is a critical gap to address.

There are a number of approaches that could gather research evidence from individual studies to provide an overall picture of current knowledge. In the thesis research, we will apply the systematic review approach with narrative evidence synthesis based on the following rationales: 1) Systematic reviews are well-established methods, which involve a series of transparent and reproducible processes such as the comprehensive literature search, the study selection based on pre-defined eligibility criteria, and the quality appraisal of included studies. 2)

The application of narrative synthesis is determined by our broad research focus, which aims to determine what have been done in the field. No quantitative synthesis is required because we will not attempt to determine if there is an association existing between an antecedent/consequence and health literacy or health numeracy or what the direction of the association is. 3) The narrative summary of evidence will be useful and valuable in terms of serving as a guide for future researchers to determine the feasibility of carrying out a quantitative synthesis on a specific antecedent/consequence of health literacy or health numeracy in a specific adult immigrant or refugee group. 4) The systematic review will identify the potential research gaps in current empirical studies, point out possible and promising research directions; and eventually provide insights for healthcare practitioners and policy-makers interested in health literacy and health numeracy of adult immigrants and refugees.

3.2 Research Objectives and Research Questions

The research aims to conduct a systematic review to collect, examine, summarize, and narratively synthesize both quantitative and qualitative empirical evidence on antecedents and consequences of health literacy or health numeracy in adult immigrants and refugees.

Specifically, this study will address the following two review questions:

(1) What antecedents of health literacy or health numeracy in adult immigrants and refugees have been investigated in the literature?

(2) What consequences of health literacy or health numeracy in adult immigrants and refugees have been investigated in the literature?

Chapter 4 Methods

The thesis research applied the systematic review approach. The reporting of systematic review followed the PRISMA (Shamseer et al., 2015), which is a checklist consisting of 27 items. The PRISMA has been well-recognized and commonly used as a guideline for conducting systematic reviews to increase the transparency of the processes.

4.1 Literature Search

4.1.1 Information Sources

We searched the following eight databases to identify potentially relevant studies: PubMed, Embase, CINAHL, PsycINFO, SCOPUS, the ProQuest Dissertations & Theses database, the CPCI-S, and the CPCI-SSH. These databases include various types of literature, such as journal articles, conference papers, and theses.

4.1.2 Search terms

The search strategy consisted of both index terms and free text terms relating to three major components: (1) health literacy, (2) health numeracy, and (3) immigrants and refugees. An experienced librarian at the University of Waterloo was consulted in developing the search terms.

Table 1 and Table 2 listed the index terms and the free text terms that were searched in each of the eight databases, respectively. We used Boolean logic “OR” to combine “health literacy” with “health numeracy”, and used Boolean logic “AND” to combine them with “immigrants or refugees”. Appendix A to G showed the search strategies designed for PubMed, Embase, CINAHL, PsycINFO, SCOPUS, the ProQuest Dissertations & Theses database, the CPCI-S and the CPCI-SSH, respectively.

Table 1. Index terms searched in the systematic review

Index Terms			
	Relating to health literacy	Relating to health numeracy	Relating to immigrants and refugees
PubMed	health literacy	not found	emigrants and immigrants (including undocumented immigrants), emigration and immigration, transients and migrants, refugees
Embase	health literacy	not found	immigrant, immigration, migrant, emigrant, migration, refugee (including asylum seeker), undocumented immigrant
CINAHL	health literacy	not found	immigrants (including immigrants, illegal), emigration and immigration, transients and migrants, relocation, refugees
PsycINFO	health literacy	health literacy	immigration, human migration, refugees
SCOPUS			not applicable
ProQuest Theses			not applicable
CPCI-S and CPCI-SSH			not applicable

Table 2. Free text terms searched in the systematic review

Free Text Terms		
Relating to health literacy	Relating to health numeracy	Relating to immigrants and refugees
literacy	numeracy	immigrant, immigrants, immigration, migrant, migrants, migration, emigrant, emigrants, emigration, newcomer, newcomers, foreign born, foreign-born, refugee, refugees, asylum seeker, asylum seekers, stateless person, stateless persons, illegal alien, illegal aliens, undocumented alien, undocumented aliens, irregular alien, irregular aliens, clandestine alien, clandestine aliens, unauthorized alien, unauthorized aliens

4.2 Study Selection

4.2.1 Eligibility Criteria

4.2.1.1 Inclusion Criteria

A study was considered eligible, if

(1) the study could provide quantitative or qualitative empirical evidence on antecedents or consequences of health literacy or health numeracy in its Results section, and;

(2) the research sample included adult immigrants or refugees (≥ 16 years or labelled as adults), and;

(3) the study was published in English language.

As introduced in **Chapter 2: Literature Review**, there is no universally accepted definition for health literacy or health numeracy. In the present systematic review, we considered

health literacy and health numeracy as literacy and numeracy in health contexts. Therefore, we only searched and selected empirical studies which explicitly applied the term “literacy” or “numeracy.”

There are many different terms which could describe or indicate refugees or immigrants, such as asylum seekers and stateless persons for refugees as well as migrants, foreign-born, and newcomers for immigrants. In the present systematic review, we considered the sample as refugees if they were explicitly labelled as refugees, asylum seekers, or stateless persons, or we considered the sample as immigrants if they were labelled as immigrants (excluding second or third-generation “immigrants”), migrants or newcomers from other countries or regions to the destination country (excluding migrant workers, students, and transients), or foreign-born persons (excluding destination country citizens who were born outside the destination country).

4.2.1.2 Exclusion Criteria

A study was excluded, if

- (1) the study was a review article, or;
- (2) the study sample was restricted to non-adults, or;
- (3) the study was about the development, implementation, or evaluation of interventions (including programs, curriculums, and courses) to improve health literacy or health numeracy, or;
- (4) the study aimed to develop or validate instruments, tools, or measures for assessing health literacy or health numeracy, or;
- (5) the results generated from adult immigrants or refugees in the study were mixed with results from other populations (e.g., general population, non-adults, or native-born minority groups) and could not be separated, or;

(6) the study was not able to provide adequate information due to reasons such as publications which had only abstracts (e.g., conference abstract).

4.2.2 Screening Process

The screening processes, in which we involved the software EndNote to store and manage references, consisted of three stages: de-duplication of the retrieved records, the title/abstract screening, and the full-text screening.

After excluding the duplicate records, three independent reviewers examined the titles and abstracts to determine whether the retrieved studies met the eligibility criteria. We obtained the full-texts for further screening if any one of the reviewers considered a study potentially eligible or without sufficient information to make decision based only on title and abstract. During full-text screening, two independent reviewers read through the potentially eligible studies in its entirety to reach a final decision.

We resolved any disagreement regarding the inclusion of a study between the reviewers through discussion. If necessary, a third party, my thesis supervisor, was consulted.

4.3 Data Extraction

Two data extraction tables were created to collect evidence from included quantitative and qualitative studies, respectively. Data items were extracted by two independent reviewers and could be found in Appendix H and Appendix I.

We resolved any disagreement regarding the data abstraction through discussion. If necessary, a third party, my thesis supervisor, was consulted.

4.4 Assessment of Risk of Bias of Included Individual Studies

4.4.1 Assessment of Risk of Bias of Included Quantitative Studies

Biases impact the internal validity of a study and could result in either overestimating or underestimating the true effects (The Cochrane Collaboration, 2011). It is critical to evaluate the risk of bias of individual studies in a systematic review because it can provide information to determine which studies possess relatively higher internal validity, in other words, which studies yield results that are closer to truth (The Cochrane Collaboration, 2011).

The tools for evaluation of the risk of bias in quantitative studies vary according to different study designs. The results of our literature search (see **Chapter 5, Section 5.2**) showed that the study designs of the included quantitative studies were cross-sectional and cohort studies.

4.4.1.1 The Newcastle-Ottawa Quality Assessment Scale (NOS) Adapted for Cross-sectional Studies

To date, there are no universally accepted tools for assessing cross-sectional studies. Among available tools, the NOS is frequently adapted for cross-sectional studies. The NOS was originally designed to evaluate the risk of bias of non-randomized studies (i.e., cohort and case-control studies) by the University of Newcastle and the University of Ottawa (Wells GA et al., 2011). The original NOS examines the quality of a study from three broad categories: the selection of study groups, the comparability between study groups, and the ascertainment of the exposure and outcome, and uses a star (*) system for semi-quantitative assessment of study quality (Wells GA et al., 2011). Many researchers have adapted the NOS to evaluate the risk of bias in cross-sectional studies (Eijkemans et al., 2012; Herzog et al., 2013; Mata et al., 2015; Modesti et al., 2016).

The NOS for cross-sectional studies used in the thesis research (Appendix J) was a version adapted from the original NOS for cohort studies (Wells GA et al., 2011) and a version published by Modesti et al., (2016) for cross-sectional studies. It was also modified to reflect the contexts of my research fields. For instance, in the original NOS (Wells GA et al., 2011), an item asks:

“Comparability of cohorts on the basis of the design or analysis

*a. study controls for _____ (select the most important factor) **

*b. study controls for any additional factor (This criterion could be modified to indicate specific control for a second important factor.) **”*

However, to our knowledge, in the field of health literacy and health numeracy, there has been no established “the most important factor.” Thus, in our modified version, we made the following revision:

“Comparability of subjects in different outcome groups

*a) Results adjusted for relevant confounders. ***

b) Results not adjusted for any relevant confounders, or information not provided.”

4.4.1.2 The Risk of Bias in Non-Randomized Studies-of Interventions (ROBINS-I) Assessment Tool (Version for Cohort-type Studies)

The ROBINS-I for cohort studies is an instrument for evaluation of risk of bias in cohort-type studies and recommended by the Cochrane Scientific Committee (<http://methods.cochrane.org/robins-i-tool>).

The ROBINS-I examines the risk of bias in seven domains which may occur in three stages: the stages of pre-intervention, at intervention, and post intervention. The phase of pre-intervention includes two bias domains – bias due to confounding and bias in selection of participants into the study; the stage of at intervention only involves one domain: bias in classification of interventions; the stage of post-intervention contains four bias domains: bias due to deviations from intended interventions, bias due to missing data, bias in measurement of outcomes, bias in selection of the reported result (Sterne et al., 2016). In this project, we considered the “exposure” factors of interest in the included studies as “intervention” though they were not truly interventions.

Within each domain, the users need to answer the signaling questions and provide response for each question. Generally, options for the responses include yes, probably yes, probably no, no, no information, not applicable. Considering all the signaling questions given by a bias domain, the users need to decide the level of risk of bias in this domain. Depending on the severity of the problems, there are four levels of risk of bias: low, moderate, serious, and critical risk of bias. After having the judgements of risk of bias within each domain, the users eventually have to make a judgement across the seven domains. The template of the tool is available in Appendix L.

4.4.2 Assessment of Risk of Bias of Qualitative Studies with the Critical Appraisal Skills Program (CASP) for Qualitative Studies

The CASP tool for qualitative studies is frequently applied by researchers (Domecq et al., 2014; Galdas et al., 2015; Masood et al., 2011). It consists of ten checklist-style items which try to identify information regarding several aspects of a qualitative study (e.g., the aims, design, data collection, data analysis, ethical issues, findings, and value of the research). The tool also provides hints under each item which may serve as criteria for judging whether the item is satisfied (Critical Appraisal Skills Programme, 2018) (Appendix N).

In the thesis research, we checked all the hints (criteria) under each item provided by the CASP tool. If all the criteria under an item were met, a judgement “yes” would be given to the item; if only some of the criteria were satisfied, a “limited” and the number of the unmet criteria would be provided; if none of the criteria were met, a “no” would be given.

Two reviewers independently carried out the appraisal. Any uncertainty or disagreement would be resolved by discussion. If necessary, we consulted my thesis advisor.

4.5 Data Synthesis by Thematic Identification

Thematic analysis is an umbrella term referring to several interpretive methods, putting focus on identifying ideas or patterns from the data to generate insightful understandings regarding a complex phenomenon (Braun et al., 2006). In the thesis research, synthesizing evidence by thematic identification from the included quantitative and qualitative studies could help achieve our research aim, which was to explore what antecedents and consequences of health literacy or health numeracy have been investigated in adult immigrants and refugees.

Furthermore, we carried out the thematic analysis by applying the framework approach, which was frequently adopted in health research (Heath et al., 2012; Sheard et al., 2013). We adopted the framework approach mainly for the following reasons: a) According to Ritchie et al., (2003), in order to easily manage the messy and sizable raw qualitative data, researchers often tend to jump directly from the raw data to abstract accounts. However, such behaviours ought to be avoided because the analysis process should be divided into several transparent blocks to create a structure of evidence. The framework approach aims to build the evidence structure and emphasizes transparent data management and thematic identification processes (Ritchie et al., 2003). b) Based on my literature search results, almost all the included studies provided cross-sectional descriptive data. The framework approach is particularly useful for analysis of this type of data (Ritchie et al., 2003; J. Smith et al., 2011).

A thematic framework, which is used to “*classify and organise data according to key themes, concepts and emergent categories*”, is a central component of the framework approach (Ritchie et al., 2003) (page 220). Each individual study using the framework method should establish a thematic framework containing a number of themes and sub-themes. These themes may come from initial coding of the transcripts or from pre-existing themes and need to be

discussed and agreed upon by the researchers (N. K. Gale et al., 2013; Ritchie et al., 2003). A matrix, which is one of the key features of the framework approach, is then created, in which a row and a column represent a case (e.g., an interviewee, an individual study in the thesis research) and a theme from the thematic framework, respectively. Data from each case is mapped to the appropriate cell of the matrix and eventually synthesized within the thematic framework (N. K. Gale et al., 2013; Ritchie et al., 2003).

Generally, there are several steps involved in the approach (N. K. Gale et al., 2013). We adapted the procedures to reflect the contexts of our research.

1. Recognizing and extracting the antecedents and consequences from the included studies
2. Establishing the working thematic framework and coding

In the systematic review, we decided to build our working thematic framework with themes pre-existing in Sorensen's et al., (2012) model for health literacy and numeracy. This decision was made mainly due to a) Researchers in the discipline of health literacy and numeracy have established a number of conceptual frameworks, in which components (i.e., antecedents and consequences) of health literacy and numeracy have been presented and discussed. These frameworks could be valuable and reliable theme sources for our working thematic framework. b) In order to cover as many antecedents and consequences as possible in the included studies, we chose Sorensen's et al., (2012) comprehensive framework because it integrated components from 12 existing conceptual models.

In Sorensen's et al.' (2012) framework, the antecedents were classified as personal determinants, situational determinants, and societal and environmental factors, while the

consequences as health service use, health behavior, health cost, health outcomes, participation, empowerment, equity, and sustainability.

The above themes of antecedents and consequences have the potentials to satisfy our thematic identification purpose but need to be explicitly specified or re-organized because Sorensen et al., (2012) did not explicitly clarify them or set clear boundaries between the themes.

For instance, Sorensen et al., (2012) did not provide explicit definitions for antecedents such as situational factors as well as societal and environmental factors. They only proposed that situational factors could be exemplified by “*social support, family and peer influences, media use and physical environment*” (page 10), while the societal and environment determinants might refer to factors such as “*demographic situation, culture, language, political forces, societal systems*” (page 10). Thus, for the convenience of analysis, we used the theme “societal and environmental antecedents” to include both situational factors and societal and environmental factors from Sorensen’s (2012) framework.

Moreover, Sorensen et al., (2012) were also ambiguous about the themes of consequences. For example, the categories “participation, empowerment, equity, and sustainability” were more like goals for improving health literacy and numeracy rather than consequences, compared to “health service use, health behaviour, health cost, and health outcome.”

Therefore, after several discussions with the other independent reviewer and the thesis advisor, we re-grouped and defined themes from Sorensen’s et al., (2012) framework to establish our working thematic framework, which consisted of personal antecedents (including sub-themes personal characteristics and personal competence), societal and environmental antecedents, health service use or behaviour, health cotcome, and health cost (Appendix P).

Next, we reflected on the antecedents and consequences to seeking understanding of their true meaning and applied the themes from our working thematic framework to the transcripts of each included study.

Using pre-existing themes for the thematic identification was actually a deductive process. To make sure we did not miss the important aspects of the data by the deductive thematic identification, we also used inductive coding. When we encountered a passage of data that seemed not to be fit for any of the pre-existing themes, we conducted open coding for the passage instead.

3. Charting the data to the matrix and interpretation

An Excel spreadsheet was used to generate a matrix. The row of the matrix represented an individual study, and the column denoted to the pre-defined themes and sub-themes. The data were mapped to the matrix by the pre-existing themes and sub-themes and eventually to the thematic framework. Antecedents and consequences of health literacy or numeracy were described by themes.

Chapter 5 Results

5.1 Results of Literature Search and Selection

The literature search was last updated on May 10, 2018. The literature search yielded 5692 records. After excluding the duplicates, conducting the title/abstract and full text screening, we included 76 eligible studies. One additional study from other sources (e.g., reference lists of included studies and relevant reviews, and studies done in my research group) was also identified (Figure 1). Thus, the total number of included studies was 77.

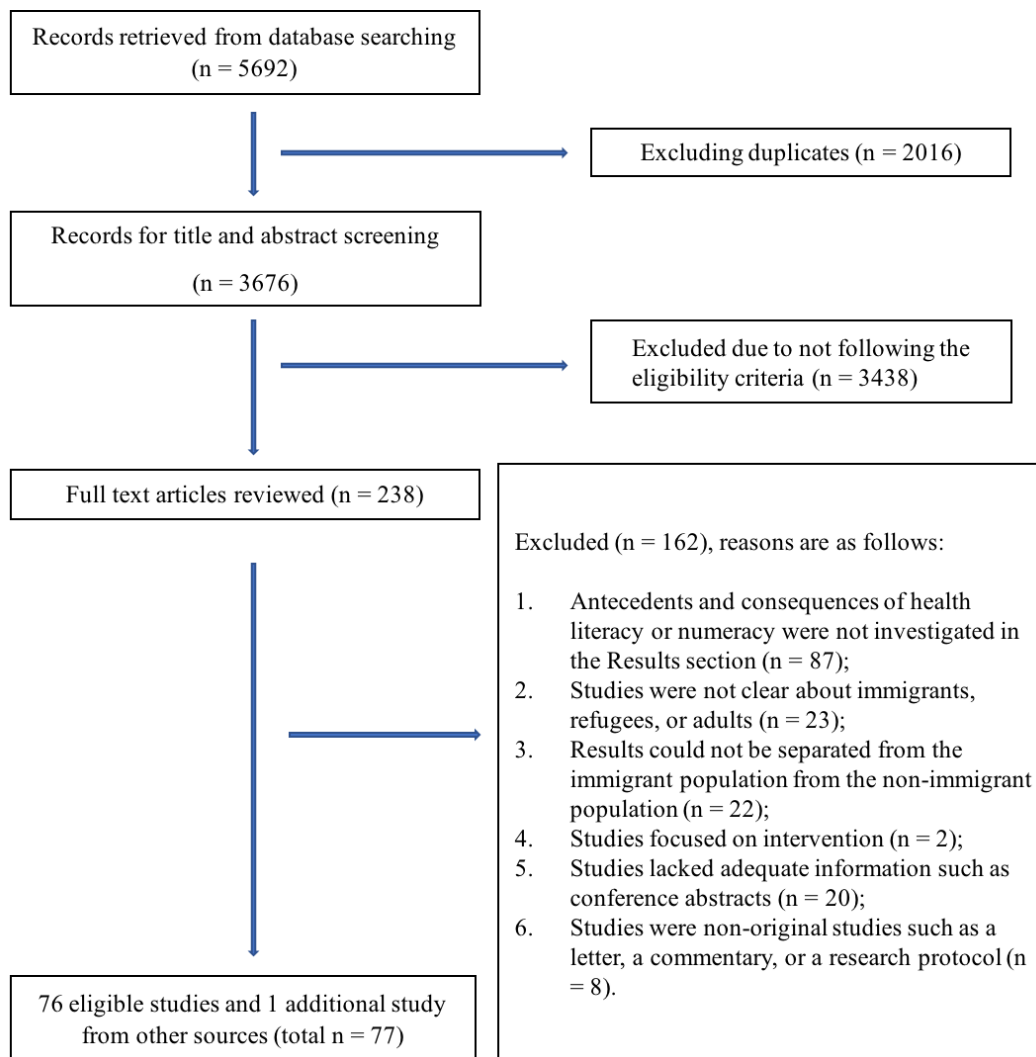


Figure 1. Flow diagram depicting the results of literature search and selection

5.2 Characteristics of the Included Studies

5.2.1 Basic Information

The included studies were published between 2004 and 2018. The average number of studies published between 2014 and 2018 was around 10, twice the average number of research published from 2009 to 2013 (Figure 2). Out of 77 included studies, 49 were quantitative and 28 were qualitative. To be noticed, two studies used a mixed quantitative and qualitative method (Jafri, 2012; Jung et al., 2017). However, we classified the two studies as qualitative because results regarding antecedents or consequences of health literacy or numeracy came from the qualitative parts.

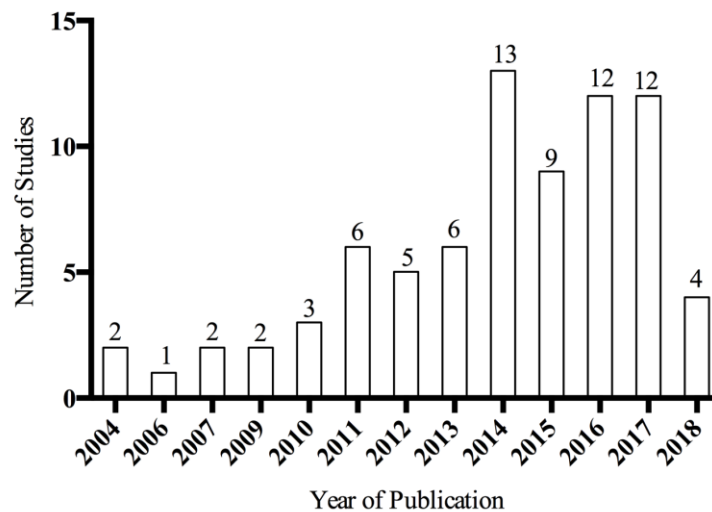


Figure 2. Distribution of the included studies by year of publication

In terms of publication type, most of the included studies (71 articles) were published as journal articles while six were PhD theses.

Lastly, among the 49 quantitative studies, 47 adopted a cross-sectional design, whereas two studies used a cohort study design (Borges et al., 2017; P. C. Smith et al., 2012). All of the 28 included qualitative studies conducted thematic analysis, mostly following an inductive way, in

which coding and theme development were directed by the content of the data collected from either interviews with individuals or focus groups, or both.

Detailed basic information regarding the first author, publication year, study type (quantitative, qualitative, or mixed), and publication type (e.g., journal article) was presented in Appendix Q.

5.2.2 Information on Health Literacy or Health Numeracy Investigated in the Included Studies

Table 3 presented the stated focus of the included studies in terms of health literacy or numeracy, the assessment instruments for health literacy or numeracy as well as the measurement methods and validation status.

Among 67 studies focusing only on health literacy, 48 explicitly used the term “health literacy”, whereas the rest applied diverse terms such as cancer literacy, dementia literacy, and obesity.

Other than a health literacy focus, nine studies also involved numeracy, four of which only mentioned numeracy when introducing the assessment tools or measures for health literacy (S. E. Choi et al., 2013; Ng et al., 2014; D. W. Omariba et al., 2011; D. W. R. Omariba et al., 2015). Among the remaining five studies, which investigated health numeracy in the Results sections, four used the term “numeracy” to represent health numeracy (Borges et al., 2017; Jacobson et al., 2016; Prins et al., 2015; Yunusa Vakkai, 2016), while one study differentiated the terms numeracy and health numeracy (S. Gatobu et al., 2016).

Only one study focused exclusively on numeracy and health numeracy (S. Gatobu et al., 2014).

There was a variety in the assessment tools used to measure the health literacy or numeracy level. Thirty-six instruments including different language versions of the same tool were adopted. The Short Test of Functional Health Literacy in Adults (S-TOFHLA) English language version was the most frequently adopted instrument (8 times), followed by the REALM English language version (3 times).

Using these assessment instruments, researchers determined the level of health literacy or numeracy in the research sample by measuring various domains, such as evaluating subjects' abilities to recognize words and comprehend information, checking people's knowledge and understanding to a given topic, and collecting self-reported outcomes. More than 70% of the included studies adopted the instruments that were validated either by themselves in the target population or by other researchers in different populations or in the same race or ethnicity.

A brief description of the assessment instruments used in the studies can be found in Appendix R.

Table 3. Information on health literacy and health numeracy in the included studies

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Becerra et al., 2017	Health literacy	Two health literacy items from the California Health Interview Survey, Asian languages (e.g., Chinese, Korean, and Vietnamese) (California Health Interview Survey, 2011)	Self-report	No
Bekker et al., 2004	Health literacy	The modified Rapid Estimate of Adult Literacy (REALM), Dutch and Standard-Arabic language versions. (Bekker et al., 2004)	Word recognition and pronunciation test	No
Beltran et al., 2016	Human papillomavirus (HPV) literacy	The 7-item Likert scale adapted from the 16-item scale of McPartland et al., (2005), English language version	Agreeing or disagreeing statements (true or false)	No
Borges et al., 2017	Parental health literacy (Both S-TOFHLA overall health literacy score and numeracy score were investigated.)	The Short Test of Functional Health Literacy in Adults (S-TOFHLA) English language version (Baker et al., 1999)	Cloze-type comprehension and quantitative skills test	Yes (Baker et al., 1999)

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Calvasina et al., 2016	Oral health literacy	The Oral Health Literacy Instrument, Portuguese language version (Calvasina et al., 2016; Sabbahi et al., 2009)	Cloze-type comprehension and quantitative skills test	Yes (Calvasina et al., 2016; Sabbahi et al., 2009)
Calvo, 2016	Health literacy	The Brief Health Literacy Screener (Chew Items), Spanish language version (Chew et al., 2004; Sarkar et al., 2011)	Self-report	Yes (Sarkar et al., 2011)
S. E. Choi et al., 2013	Health literacy (Numeracy was only mentioned once when introducing the assessment tool)	The Newest Vital Sign (NVS), Korean language version (S. E. Choi et al., 2013)	Reading comprehension and numeracy skills	Yes (S. E. Choi et al., 2013)
Y. J. Choi et al., 2016	Mental health literacy	The Mental Health Literacy Scale, Korean language version (Y. J. Choi et al., 2016)	Measuring knowledge, attitudes, and stigma	Yes (Y. J. Choi et al., 2016)
Coffman et al., 2007	Health literacy	The Short Assessment of Health Literacy for Spanish-speaking Adults, Spanish language version (S. Y. Lee et al., 2006)	Word recognition and matching test	Yes (S. Y. Lee et al., 2006)

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Coffman et al., 2010	Health literacy	The Short Assessment of Health Literacy for Spanish-speaking Adults, Spanish language version (Lee et al., S. Y. 2006)	Word recognition and matching test	Yes (S. Y. Lee et al., 2006)
Coffman et al., 2012	Health literacy	The S-TOFHLA, Spanish language version (Stonbraker et al., 2015)	Cloze-type comprehension and quantitative skills test	Yes (Aguirre et al., 2005)
Diamond et al., 2014	Dementia literacy	The dementia literacy survey, Chinese language version (Diamond et al., 2014; Woo, 2013)	Agreeing or disagreeing statements (true or false)	No
Gatobu et al., 2014	Numeracy, health numeracy	a) the addition subtraction-correction and addition tasks from the Kit of Factor-Referenced Cognition Test (French kit), English (Ekstrom et al., 1979); b) the health numeracy component of the S-TOFHLA, English language version (Baker et al., 1999); c) the NVS, English language version (Weiss et al., 2005)	a) addition subtraction-correction and addition ability; b) cloze-type comprehension and quantitative skills test; c) reading comprehension and numeracy skills	a) no; b) no; c) yes (Weiss et al., 2005)

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Gatobu et al., 2016	Health literacy, numeracy, and health numeracy	a) the addition subtraction-correction and addition tasks from the French kit, English (Ekstrom et al., 1979); b) the S-TOFHLA, English language version (Baker et al., 1999); c) the NVS, English language version (Weiss et al., 2005)	a) addition - subtraction correction and addition ability; b) cloze-type comprehension and quantitative skills test; c) reading comprehension and numeracy skills	a) no; b) yes (Baker et al., 1999); c) yes (Weiss et al., 2005)
Gele et al., 2016	Health literacy	The short language version of the European health literacy questionnaire (HLS-EU-Q16), English language version (HLS-EU Consortium, 2012)	Self-report	Yes (Gele et al., 2016)
Geltman et al., 2013	Health literacy	The S-TOFHLA, English language version (Baker et al., 1999)	Cloze-type comprehension and quantitative skills test	Yes (Baker et al., 1999)
Geltman et al., 2014	Health literacy and dental health literacy	a) the S-TOFHLA, English language version (Baker et al., 1999); b) the Rapid Estimate of Adult Literacy in Dentistry 30 Short Form, English language version (J. Y. Lee et al., 2007)	a) cloze-type comprehension and quantitative skills test; b) word recognition	a) yes (Baker et al., 1999); b) yes (J. Y. Lee et al., 2007)

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Hernandez-Mekonnen et al., 2016	Maternal health literacy	The Parental Health Literacy Activities Test, Spanish language version (Yin et al., 2012)	Reading comprehension and numeracy skills	Yes (yin et al., 2012)
Idehen et al., 2017	Literacy in Finnish/Swedish	Not reported		
Igarashi et al., 2013	Health literacy	The modified REALM, Japanese language version (Igarashi et al., 2013)	Word recognition and pronunciation test	No
Jacobson et al., 2016	English functional health literacy (Both TOFHLA overall health literacy score and numeracy score were investigated.)	The TOFHLA, English language version (Parker et al., 1995)	Cloze-type comprehension and quantitative skills test	Yes (Parker et al., 1995)
Kankou et al., 2017	French literacy	Questions about if participants can read and/or write French (Kankou et al., 2017)	Self-report	No

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Khuu et al., 2018	Health literacy	Two statements, “I am confident to understand health information given by health care professionals.” and “I have experiences of missing medication because I did not know how to take medication.” Responses were determined on a 5-point Likert-type scale (strongly disagree (1 point) to strongly agree (5 points)). (Khuu et al., 2018)	Self-report	No
Kim et al., 2013	Korean language literacy	The Korean Language Literacy Scale (Kim et al., 2013)	Self-report	Yes (Kim et al., 2013)
Kim et al., 2018	Health literacy	The Assessment of Health Literacy in Cancer Screening, English version (Han et al., 2014)	Assessing comprehension, familiarity, and recognition of words related to breast and cervical cancer screening	Yes (Han et al., 2014)
Ko, 2014	Colorectal cancer literacy	The Assessment of Colon Cancer Literacy, Korean language version (Holubar et al., 2009; Ko, 2014)	Agreeing or disagreeing statements (true or false)	Yes (Ko, 2014)

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Koch-Weser et al., 2006	Literacy	A question asking the participants 'if they read materials such as books, letters, or newspapers most weeks.' (Koch-Weser et al., 2006)	Self-report	No
H. Y. Lee et al., 2012	Health literacy	The 16-item health literacy screening scale, English and Korean versions (Chew et al., 2004)	Self-report	Only 3 items (Chew Items) in the English version were validated (Chew et al., 2004)
H. Y. Lee et al., 2014	Cancer literacy	The 12 survey items on cancer risk and prevention, English and Korean versions (H. Y. Lee et al., 2014; Stein et al., 2007)	Agreeing or disagreeing statements	No
H. Y. Lee et al., 2015a	Health literacy	The 12 items from the Chew's et al., (2004) 16-item health literacy screening scale, English and Korean versions	Self-report	Yes (H. Y. Lee et al., 2015a)
H. Y. Lee et al., 2015b	Health literacy	The two health literacy items from California Health Interview Survey, Spanish language version (California Health Interview Survey, 2011)	Self-report	No

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
H. Y. Lee et al., 2015c	Cancer literacy	Items adopted from Stein's et al., (2007) 12 survey items on cancer risk and prevention, English and Korean versions. Not clear the number of items adopted	Agreeing or disagreeing statements	No
Mantwill et al., 2017	Health literacy	a) the S-TOFHLA, Albanian, Portuguese, Serbian, German, and Italian language versions (Mantwill et al., 2017); b) the Brief Health Literacy Screeners (Chew items), Albanian, Portuguese, and Serbian language versions (Chew et al., 2004; Mantwill et al., 2017)	a) reading comprehension and quantitative skills test; b) self-report	a) yes (Mantwill et al., 2017); b) yes (Mantwill et al., 2017)
Ng et al., 2014	Health literacy (Numeracy was a part of the health literacy skills)	The Health Activities Literacy Scale (HALS), English language version (Canadian Council on Learning, 2008; Rudd et al., 2004)	Reading and using prose and document and testing quantitative skills	Yes (Rudd et al., 2004)
Njeru et al., 2016	Diabetes health literacy	The Spoken Knowledge in Low Literacy in Diabetes scale, English and Somali language versions (Njeru et al., 2016; Rothman et al., 2005)	Checking patient knowledge on diabetes self-care issues	English version: yes; Somali version: no (Njeru et al., 2016; Rothman et al., 2005)

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Omariba et al., 2011	Health literacy (Numeracy was a part of the health literacy skills)	The HALS, English language version (Canadian Council on Learning, 2008; Rudd et al., 2004)	Reading and using prose and document and testing quantitative skills	Yes (Rudd et al., 2004)
Omariba et al., 2015	Health literacy (Numeracy was a part of the health literacy skills)	The HALS, English language version (Canadian Council on Learning, 2008; Rudd et al., 2004)	Reading and using prose and document and testing quantitative skills	Yes (Rudd et al., 2004)
Prins et al., 2015	Literacy and numeracy	The literacy and numeracy tasks from the Program for the International Assessment of Adult Competencies, English language version (OECD, 2012)	Assessing literacy and numeracy skills in technology-rich environment	No
P. C. Smith et al., 2012	Functional health literacy	The TOFHLA, English and Spanish language versions (Parker et al., 1995)	Cloze-type comprehension and quantitative skills test	Yes (Parker et al., 1995)
Thabit et al., 2009	Health literacy	The REALM, English language version (Davis et al., 1993)	Word recognition and pronunciation test	Yes (Davis et al., 1993)

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Thomson et al., 2011	Health literacy	a) the S-TOFHLA, English language version (Baker et al., 1999), b) the REALM, English (Davis et al., 1993), and c) the Cloze (Thomson et al., 2011)	a) cloze-type comprehension and quantitative skills test, b) word recognition test, and c) multiple-choice Cloze test	a) yes (Baker et al., 1999), b) (Davis et al., 1993), and c) no
Todd et al., 2011a	Functional health literacy	The S-TOFHLA, English language version (Baker et al., 1999)	Cloze-type comprehension and quantitative skills test	Yes (Baker et al., 1999)
Todd et al., 2011b	Health literacy	a) the S-TOFHLA English language version (Baker et al., 1999) and b) the Cloze, Chinese and English versions (Thomson et al., 2011)	a) cloze-type comprehension and quantitative skills test, b) multiple-choice Cloze test	a) yes (Baker et al., 1999), b) no
Tsoh et al., 2016	Health literacy	The Single Item Literacy Screener, English language version (Morris et al., 2006)	Self-report	Yes (Morris et al., 2006)
Wangdahl et al., 2014	Functional health literacy and comprehensive health literacy	a) the Swedish functional health literacy scale, Arabic, Dari, Somali, and English language versions, b) the slightly modified HLS-EU-Q16, Arabic, Dari, Somali, and English language versions (Wangdahl et al., 2014, 2015)	a) self-report, b) self-report	a) no, b) no

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Wangdahl et al., 2015	Functional health literacy and comprehensive health literacy	a) the Swedish functional health literacy scale, Arabic, Dari, Somali, and English language versions (Wangdahl et al., 2014, 2015), b) the slightly modified HLS-EU-Q16, Swedish language version (Pelikan et al., 2017)	a) self-report, b) self-report	a) no, b) yes (Pelikan et al., 2017)
Wangdahl et al., 2018	Comprehensive health literacy	The slightly modified HLS-EU-Q16, Swedish language version (Pelikan et al., 2017)	Self-report	Yes (Pelikan et al., 2017)
Wister et al., 2010	Health literacy	The HALS, English language version (Canadian Council on Learning, 2008; Rudd et al., 2004)	Reading and using prose and document and testing quantitative skills	Yes (Rudd et al., 2004)
Yunusa Vakkai, 2016	Literacy and numeracy	A self-rating of one's ability to read (1=excellent, 5 = poor) and ability to work with numbers (1=excellent, 5 = poor)	Self-report	No
Alzayer et al., 2017	Health literacy	A qualitative study, no assessment tools were used		
Carroll et al., 2007	Health literacy	A qualitative study, no assessment tools were used		

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Clark et al., 2014	Literacy	A qualitative study, no assessment tools were used		
Cyril et al., 2017	Health and food literacy	A qualitative study, no assessment tools were used		
Filippi et al., 2014	Literacy	A qualitative study, no assessment tools were used		
Floyd et al., 2017	Literacy	A qualitative study, no assessment tools were used		
Gele et al., 2017	Health literacy	A qualitative study, no assessment tools were used		
Gregory, 2015	Health care literacy	A qualitative study, no assessment tools were used		
Groenenberg et al., 2015	Health literacy	A qualitative study, no assessment tools were used		
Hurley et al., 2013	Health literacy	A qualitative study, no assessment tools were used		
Jafri, 2012	Health literacy	The REALM English language version (Davis et al., 1993)	Word recognition and pronunciation test	Yes (Davis et al., 1993)
Jung et al., 2017	Health literacy	A qualitative study, no assessment tools were used		

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
F. H. Lee et al., 2014	Health literacy			A qualitative study, no assessment tools were used
Leung et al., 2014	Health literacy			A qualitative study, no assessment tools were used
Marshall et al., 2010	Health system literacy			A qualitative study, no assessment tools were used
McMichael et al., 2009	Sexual health literacy			A qualitative study, no assessment tools were used
Murray et al., 2018	Health literacy			A qualitative study, no assessment tools were used
Oliver, 2015	Health literacy			A qualitative study, no assessment tools were used
Renzaho et al., 2017	Pre-migratory obesity literacy			A qualitative study, no assessment tools were used
Schoenmakers et al., 2017	Mental health literacy			A qualitative study, no assessment tools were used
Shedlin et al., 2004	Literacy			A qualitative study, no assessment tools were used
Smaland Goth et al., 2011	Medical literacy			A qualitative study, no assessment tools were used

Included Studies	Stated Research Focus	Assessment Tools	Assessment Methods	Validation Status of the Assessment Tools
Sriphanlop et al., 2014	Health literacy	A qualitative study, no assessment tools were used		
Taiwo, 2013	Health literacy	A qualitative study, no assessment tools were used		
Thomson et al., 2012	Functional health literacy and interactive health literacy	A qualitative study, no assessment tools were used		
Todd et al., 2011c	Health literacy	The S-TOFHLA, English language version (Baker et al., 1999)	Cloze-type comprehension and quantitative skills test	Yes (Baker et al., 1999)
Watts et al., 2014	Parental sexual health literacy	A qualitative study, no assessment tools were used		
Woudstra et al., 2016	Literacy	A qualitative study, no assessment tools were used		

5.2.3 Information on Adult Immigrants and Refugees Involved in the Included Studies

Table 4 shows the information on the adult immigrants and refugees involved in the 77 quantitative and qualitative studies, including the primary research focus in terms of immigrants and refugees, country of origin of the immigrants and refugees, their destination country, gender, and age.

Sixty-one studies identified their research sample as solely immigrants, in which three studies further specified that the immigrants were all or mostly undocumented (Gregory, 2015; Hernandez-Mekonnen et al., 2016; Shedlin et al., 2004) and six studies described their research participants as migrants (Alzayer et al., 2017; Cyril et al., 2017; Filippi et al., 2014; Njeru et al., 2016; Schoenmakers et al., 2017; Smaland Goth et al., 2011).

Two additional studies primarily focused on the general population rather than immigrants, but they included terms suggesting the involvement of immigrants, such immigrant status (Borges et al., 2017; Wister et al., 2010).

Thirteen of the included studies focused only on refugees (Carroll et al., 2007; Clark et al., 2014; Floyd et al., 2017; P. Geltman et al., 2014; P. L. Geltman et al., 2013; K. Kim et al., 2018; Koch-Weser et al., 2006; McMichael et al., 2009; Oliver, 2015; Wangdahl et al., 2014, 2015, 2018; Watts et al., 2014). In addition, the research sample of one study consisted of both immigrants and refugees (Idehen et al., 2017).

The countries (regions) of origin of the immigrants and refugees in the included studies were diverse and mainly distributed through Asia, Africa, South America, and Mexico. Most of the destination countries of these immigrants and refugees were located in North America (the

USA and Canada), Australia, and Europe, in which the top three destination countries, in turn, were the USA (35 articles), Canada (15 articles), and Australia (8 articles).

There was an imbalance in the gender composition among the 49 quantitative studies. Thirteen studies investigated only female participants. Excluding six studies that were unclear on this issue or had the same gender composition, 23 studies involved more female participants than male participants, in 11 of which the number of female subjects were at least as 1.5 times as the number of male participants. The imbalance was also found in the 28 qualitative studies. Eleven of them were conducted in only women participants. None of the quantitative and qualitative studies investigated health literacy or numeracy in male participants only. The gender imbalance indicates a potential risk that immigrant and refugee men might be under-represented in comparison to women in terms of health literacy and numeracy empirical research on antecedents and consequences.

Lastly, the included studies used diverse indices to present the participants' age, such as mean age, mean age and standard deviation (SD), age range, median age, and median age and inter-quartile range (IQR). Also, some included studies only provided the percentages of participants in different age groups or did not report the age at all. As a result, we were not able to find any pattern in the age of the immigrants and refugees.

Table 4. Information on immigrants and refugees from the included studies

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Quantitative Studies					
Becerra et al., 2017	Immigrants	Not reported, foreign-born Hispanics	USA	Female: 1889, male: 1172	Mean (95% confidence interval (95% CI)): low health literacy group: 40.69 (39.62 to 41.76); adequate health literacy group: 40.32 (39.54 to 41.09)
Bekker et al., 2004	Immigrants	Morocco	Netherlands	Female: 50	Mean \pm SD: illiterate women: 47.5 \pm 6.9, literate women: 41.4 \pm 5.6
Beltran et al., 2016	Immigrants	Not reported, Hmong	USA	Female: 121, male: 71	Mean \pm SD: 29.24 \pm 8.60
Borges et al., 2017	General population (immigrant status as a variable)	Born outside Canada	Canada	Not reported (In total sample, female: 132, male: 52)	Not reported, age of total sample (mean \pm SD): 39.1 \pm 6.6
Calvasina et al., 2016	Immigrants	Brazil	Canada	Female: 74, male: 35	Mean \pm SD: 34.5 \pm 9

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Calvo, 2016	Immigrants	Not reported, Latino adults born overseas	USA	Female: 1488, male: 1508	Mean \pm SD: 41.24 \pm 0.3
S. E. Choi et al., 2013	Immigrants	Korea	USA	Among 145 immigrants, female: 50%, male: 50%	Mean \pm SD: 49.17 \pm 12.29
Y. J. Choi et al., 2016	Immigrants	Vietnamese, Korean Chinese, Chinese, Philippines, Japanese, Thai, Mongolian, Cambodian, Russian, and others	Korea	Female: 209	87 subjects (41.6%) < 30 years; 83 (39.5%) between 31 to 40 years; 35 (16.7%) between 41 to 50 years; 4 (1.9%) > 51 years
Coffman et al., 2007	Immigrants	Mexico and other central and south American countries	USA	Among 99 immigrants, female: 76.8%, male: 23.2%	Mean \pm SD: 35.7 \pm 10.79
Coffman et al., 2010	Immigrants	Mexico and other central and south American countries	USA	Among 99 immigrants, female: 76, male: 23	Mean \pm SD: 35.7 \pm 10.8
Coffman et al., 2012	Immigrants	Mexico and central and south American countries and others	USA	Female: 113, male: 31	Mean \pm SD: 43.6 \pm 12.6

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Diamond et al., 2014	Immigrants	China	USA	Female: 109, male: 42	Age range: 40 to 64
Gatobu et al., 2014	Immigrants	China and Kenya	Canada	Female: 73, male: 47	46 Kikuyu and 42 Mandarin speakers between 40 and 50 years; 14 Kikuyu and 18 Mandarin speakers > 50 years
Gatobu et al., 2016	Immigrants	China and Kenya	Canada	Female: 73, male: 47	Most were between 40 and 50 years (n = 88, 73%)
Gele et al., 2016	Immigrants	Somalia	Norway	Female: 302	Mean ± SD: 36.13 ± 8.0
Geltman et al., 2013	Refugees	Somalia	USA	Among 439 refugees, female: 58.1%, male: 41.9%	140 subjects (31.9%) between 18 and 24 years; 186 (42.4) between 25 and 44 years; 113 (25.7) ≥ 45 years
Geltman et al., 2014	Refugees	Somalia	USA	Among 439 refugees, female: 58.1%, male: 41.9%	Mean ± SD: 35.2 ± 14.7
Hernandez-Mekonnen et al., 2016	Immigrants, undocumented	Mexico	USA	Female: 87	Mean ± SD: 28.2 ± 4.1

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Idehen et al., 2017	Immigrants and refugees	Russia or the Soviet Union, Somalia, and Iraq or Iran	Finland	Female: 620	Mean \pm SD: Russian: 42.3 \pm 10.4; Somali: 39.3 \pm 9.1; Kurdish: 38.4 \pm 8.5
Igarashi et al., 2013	Immigrants	China, Brazil, Philippines, South Korea, and others	Japan	Female: 236	Maternal age (mean \pm SD): 29.2 \pm 5.3
Jacobson et al., 2016	Immigrants	Not reported, Hispanic immigrants	USA	Among 439 immigrants, female: 78%, male: 22%	14 subjects (9.7%) < 30 years; 59 (41%) between 31 to 45 years; 64 (44.4%) between 46 to 60 years; 7 (4.9%) > 60 years
Kankou et al., 2017	Immigrants	Sub-Saharan countries	France	Female: 147, male: 99	Median (IQR): 41 (35 to 48)
Khuu et al., 2018	Immigrants	Not reported, Hmong	USA	Female: 118, male: 50	Mean \pm SD: 29.24 \pm 8.60
Kim et al., 2013	Immigrants	China, Japan, Philippines, Vietnam, Mongol, others	Korea	Female: 173	Mean \pm SD: 28.6 \pm 6.2
Kim et al., 2018	Immigrants	Korea	USA	Female: 565	Mean \pm SD: 46.1 \pm 8.5

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Ko, 2014	Immigrants	Korea	USA	Female: 141, male: 113	Mean \pm SD: 60.52 \pm 8.22
Koch-Weser et al., 2006	Refugees	Southeast Asia	USA	Female: 249, male 132	Mean \pm SD: 44.3 \pm 13.2
H. Y. Lee et al., 2012	Immigrants	Korea	USA	Female: 202, male: 205	Mean \pm SD: 57.2 \pm 16.6
H. Y. Lee et al., 2014	Immigrants	Korea	USA	Female: 202, male: 205	67 subjects (16.5%) between 21 and 39 years; 58 (14.3%) between 40 and 49 years; 74 (18.2%) between 50 and 59 years; 90 (22.1%) between 60 and 69 years; 118 (29.0%) 70 and older
H. Y. Lee et al., 2015a	Immigrants	Korea	USA	Female: 202, male: 205	Mean \pm SD: 57.2 \pm 16.6
H. Y. Lee et al., 2015b	Immigrants	China, Korea, Philippines, Vietnam, South Asia	USA	Among 3053 immigrants, female: 57%, male: 43%	Age: 50.26

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
H. Y. Lee et al., 2015c	Immigrants	Hmong	USA	Female: 164	Mean: 30
Mantwill et al., 2017	Immigrants	Kosovo, Albania, Macedonia, Portugal, Serbia	Switzerland	Female and male (883 immigrants)	138 subjects (12%) between 18 and 24 years; 278 (24.3%) between 25 and 34 years; 231 (20.2%) between 35 and 44 years; 286 (25.0%) between 45 and 54 years; 157 (13.7%) between 55 and 64 years; 47 (4.11%) 65 and older; 9 (0.8%) missing
Ng et al., 2014	Immigrants	Not reported	Canada	Among 3861 immigrants, female: 51.7%, male: 48.3%	9.5% of the immigrants between 16 and 25 years; 16.1% between 26 and 35 years; 21.2% between 36 and 45 years; 19.1% between 46 and 55 years; 15% between 56 and 65 years; 19.1% > 65 and older
Njeru et al., 2016	Immigrants (migrants)	Somalia	USA	Female: 50	Mean ± SD: 52.54 ± 16.01
Omariba et al., 2011	Immigrants	Not reported	Canada	Among 3861 immigrants, female: 51.7%, male: 48.3%	9.5% of the immigrants between 16 and 25 years; 16.1% between 26 and 35 years; 21.2% between 36 and 45 years; 19.1% between 46 and 55 years; 15% between 56 and 65 years; 19.1% > 65 and older

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Omariba et al., 2015	Immigrants	Not reported	Canada	Among 3861 immigrants, female: 51.7%, male: 48.3%	9.5% of the immigrants between 16 and 25 years; 16.1% between 26 and 35 years; 21.2% between 36 and 45 years; 19.1% between 46 and 55 years; 15% between 56 and 65 years; 19.1% > 65 and older
Prins et al., 2015	Immigrants	Asians, Hispanics and other immigrants	USA	Among 613 immigrants, female: 52.1%, male: 47.9%	10.6% of the immigrants 24 years or less; 24.7% between 25 and 34 years; 27.1% between 35 and 44 years; 22.8% between 45 and 54 years; 14.8% 55 or older
P. C. Smith et al., 2012	Immigrants	Spanish-speaking immigrants	USA	Both genders (50 immigrants)	Mean: 33
Thabit et al., 2009	Immigrants	South and East Asia, Africa, Western and Eastern Europe, Middle East, South America	Ireland	Not reported (52 immigrants)	Mean \pm SD: 45.8 \pm 11.8
Thomson et al., 2011	Immigrants	Not reported	Canada	Female: 78	Mean \pm SD: 53 \pm 7.06

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Todd et al., 2011a	Immigrants	China	Canada	Female: 103	Mean \pm SD: 63.61 \pm 8.0
Todd et al., 2011b	Immigrants	China	Canada	Female: 106	Mean \pm SD: 63.2 \pm 8.2
Tsoh et al., 2016	Immigrants	China	USA	Among 705 immigrants, female: 81%, male: 19%	Mean \pm SD: 62.2 \pm 6.9
Wangdahl et al., 2014	Refugees	Somalia, Iraq, Syria, Afghanistan and others	Sweden	Among 455 immigrants, female: 46%, male: 54%	Mean \pm SD: 35.8 \pm 10.6
Wangdahl et al., 2015	Refugees	Somalia, Iraq, Syria, Afghanistan, others	Sweden	Female: 170, male: 184	Mean \pm SD: 35.4 \pm 10.5
Wangdahl et al., 2018	Refugees	Somalia, Iraq, Iran, Syria, Afghanistan, others	Sweden	Female: 112, male: 391	Mean: 38, age range: 18 to 74

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Wister et al., 2010	General population (year of immigration as a variable)	Not reported	Canada	Not reported (about 611 immigrants)	≥ 66
Yunusa Vakkai, 2016	Immigrants	66.4% were Hispanics	USA	Female: 203, male: 166	Mean ± SD: 40.41 ± 10.073
Qualitative Studies					
Alzayer et al., 2017	Immigrants (migrants)	Arabic-speaking migrants	Australia	Female: 25	5 (20%) of the participants in their 20s; 7 (28%) in their 30s; 2 (8%) in their 40s; 1 (4%) in their 50s; 5 (20) in their 60s; 5 (20) in their 70s
Carroll et al., 2007	Refugees	Somalia	USA	Female: 34	Median: 27, age range: 18 to 53
Clark et al., 2014	Refugees	Sudan, Burundi, Congo, Burma, Afghanistan and Bhutan	Australia	Female: 36	Not reported (refugee women)
Cyril et al., 2017	Immigrants (migrants)	Vietnamese, Burmese, African, Afghani, Indian	Australia	Female: 34, male: 5	Median (IQR): 35.5 (33 to 37)

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Filippi et al., 2014	Immigrants (migrants)	Somalia	USA	Female: 6, male: 5	Median: 40
Floyd et al., 2017	Refugees	Somali, Liberian, Eritrean, Ethiopian.	Canada	Female: 8	Age range: 20 to 38
Gele et al., 2017	Immigrants	Pakistani, Somali	Norway	Female: 35	27 participants between 25 and 50 years; 8 between 51 and 70 years
Gregory, 2015	Immigrants, mostly undocumented	Latino immigrants	USA	Female and male (13 immigrants)	Age range: 35 to 66
Groenenberg et al., 2015	Immigrants	Turkish, Moroccan, Hindustani, Creole	Netherlands	Not reported (60 immigrants)	Age range: 45 to 70, except for Surinamese: 35 to 70

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Hurley et al., 2013	Immigrants	Greece	Australia	Among interviewees: female: 15, male: 7; focus group participants: both female and male	Mean age and age range of interviewees: 79 (69 to 92); focus group participants: not reported (elders)
Jafri, 2012	Immigrants	Pakistan	USA	Female: 12	Mean and age range: 30.9 (27 to 35)
Jung et al., 2017	Immigrants	China and Korea	USA	Among interviewees: female: 7, male: 10; focus group participants: female: 73, male: 47	Mean \pm SD in interviewees: 52.1 \pm 12.3; mean \pm SD in focus group participants: 62.9 \pm 7.5
F. H. Lee et al., 2014	Immigrants	Vietnam	Taiwan, China	Female: 17	Mean and age range: 34.8 (32 to 56)
Leung et al., 2014	Immigrants	China	USA	Female: 11, male: 18	Mean \pm SD: 63.6 \pm 12.2

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Marshall et al., 2010	Immigrants	China, South Asia	Canada	Female: 39, male: 39	39 participants < 59 years and 39 > 50 years
McMichael et al., 2009	Refugees	Iraq, Afghanistan, Burma, Sudan, Liberia, and the Horn of Africa.	Australia	Female: 75, male: 67	Age range: 16 to 25
Murry et al., 2018	Refugees	Bhutan	Australia	Female: 3, male: 14	All > 45 years but one between 18 and 45 years
Oliver, 2015	Refugees	Bhutan	USA	Female: 32	5 subjects (15.6%) between 18 and 29 years; 8 (25.0%) between 30 and 39 years; 9 (28.1%) between 40 and 49 years; 4 (12.5%) between 50 and 59 years; 2 (6.3%) between 60 and 69 years; 3 (9.4%) 70 and 79 years; 1 (3.1%) 80 or older
Renzaho et al., 2017	Immigrants	Vietnam, India, Africa, Middle East	Australia	Female: 41, male: 7	Median: 35, inter-quartile range (31 to 39)

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Schoenmakers et al., 2017	Immigrants (migrants)	Turkish, Moroccan, Surinamese	Netherlands	Among group interviewees: female 14, male 24; individual interviewees: not reported	Group interviews: ≥ 50 , individual interviews: ≥ 60
Shedlin et al., 2004	Immigrants, mostly undocumented	Dominican, Mexican, and Central American	USA	Female: 30, male: 27	Age range: 19 to 61
Smaland Goth et al., 2011	Immigrants (migrants)	Poland, Sweden, Denmark, Iraq, Germany, Somalia, Pakistan, Bosnia-Herzegovina, Iran, Russia, Great Britain, Vietnam, Turkey,	Norway	Not reported (13 immigrants)	Not reported (migrant leaders)
Sriphanlop et al., 2014	Immigrants	Ethiopia, Gabon, Gambia, Ghana, Guinea, Mali, Nigeria, Republic of the Congo, Senegal, Sierra Leone, Swaziland, Tanzania, Uganda	USA	Female: 10, male: 12	Mean \pm SD: 40.4 \pm 15.0

Included Studies	Primary Focus in terms of Immigrants or Refugees	Country (region) of Origin	Destination Country (region)	Gender of Immigrants or Refugees (number or %)	Age of Immigrants or Refugees (years)
Taiwo, 2013	Immigrants	Nigerian population	USA	Female: 10, male: 10	≥ 40
Thomson et al., 2012	Immigrants	El Salvador, Colombia, Chile, Guatemala, Peru, Mexico, Nicaragua, Venezuela, Uruguay, Cuba, and Spain.	Canada	Female: 43	Mean \pm SD: 54 ± 7.1
Todd et al., 2011c	Immigrants	China	Canada	Female: 50	Mean \pm SD: 60.0 ± 7.3
Watts et al., 2014	Refugees	Ethiopia, Sudan, Liberia, Burundi, Sierra Leone	Australia	Female: 16	Age range: 17 to 30
Woudstra et al., 2016	Immigrants	Turkey, Morocco, and Surinam	Netherlands	Female: 18, male: 12	Age range: 48 to 74

5.3 Assessment of Risk of Bias of Included Quantitative and Qualitative Studies

Forty-seven quantitative studies adopted the cross-sectional and two studies the cohort study designs. Twenty-eight studies were qualitative and conducted thematic analysis on information collected from interviews or focus groups. The assessment tools we used in the risk of bias assessment varied according to the different study designs.

5.3.1 Risk of Bias of Quantitative Studies

5.3.1.1 Assessment of Risk of Bias of Included Cross-sectional Studies

Thirty-eight out of 47 cross-sectional studies used non-random sampling methods such as the snowball and convenience sampling, and four studies did not describe their sampling strategies. Only five studies adopted a random sampling approach. Few included studies justified the sample size (4 studies) or provided information on both the response rate and the characteristics of the non-respondents (2 studies). In terms of the comparability of subjects in different outcome groups, majority of the studies (37 out of 47) adjusted one or more confounders (or covariates) in their research.

Finally, studies scored well on the “exposure and outcome” domain. Exposures in 33 studies were measured by validated assessment tools or did not require a tool (e.g., age, educational level), whereas 11 studies adopted non-validated tools. None of the studies measured their outcomes by independent blind assessments. Instead, they either obtained outcomes from medical records or non-blind assessments using objective methods such as measuring blood glucose level (18 studies), or collected outcomes self-reported by the participants (29 studies). Most studies (41 studies) used appropriate statistical analysis and clearly described them.

A summary of the risk of bias assessment for 47 cross-sectional quantitative studies is given in Appendix J. The NOS adapted for cross sectional studies was adopted for the assessment (see Appendix K).

5.3.1.2 Assessment of Risk of Bias of Cohort Studies

Known and unknown factors might distort the associations between the health literacy or numeracy level and the outcomes. Because there was no appropriate measure to contain or adjust for the confounding in the two cohort studies, we classified them as of critical risk of bias in the confounding domain, low risk of bias in other domains. Thus, we judged the two cohort studies as critical risk of bias overall, and it should be borne in mind that the results and conclusions may be influenced by confounding factors. The assessment tool – ROBINS-I could be found in Appendix L; the summary of the risk of bias assessment in Appendix M.

5.3.2 Assessment of Risk of Bias of Qualitative Studies

The detailed CASP items as well as the criteria could be found in Appendix N. Appendix O summarized the assessment results of included qualitative studies. As previously mentioned in **Section 4.4.2**, if the individual study satisfied all or none of the criteria within an item, a “yes” or “no” would be given, respectively. If only some of the criteria were satisfied, a judgement of “limited” as well as the No. of the unmet criteria would be given.

None of the 28 qualitative studies satisfied all the 10 items. However, all of them had a clear statement of research aims (item 1) and appropriately chose the qualitative approach to interpret subjective experiences or perceptions of the research participants (item 2).

About half of the studies did not discuss how they decided which specific methods to use (item 3). In terms of the appropriateness of recruitment strategy (item 4), eight studies satisfied

all three criteria, one did not meet any criteria, and the rest were considered limited mostly because the researchers did not explain why the participants they recruited were the most proper ones to address the study's research objectives (criteria 4.2). Three studies satisfied all seven criteria within the data collection domain (item 5), while the rest were limited mainly because they did not justify the setting for data collection (criteria 5.1) or did not justify the data collection methods (criteria 5.3).

Studies did not perform well in terms of fully examining the researcher-participant relationship (item 6) and considering ethical issues (item 7). For the former item, none of the studies satisfied all the criteria, and no studies critically examined the researcher's own role, potential bias and influence in formulating research questions (criteria 6.1). Regarding the latter, most of them (25 out of 28 studies) provided information on the ethical approval (criteria 7.3). However, majority of these studies (16 out of 25 studies) were still considered limited because they did not provide adequate details about ethical issues such as how the researchers explained their research to potential participants (criteria 7.1 and 7.2)

All of the studies were limited in terms of sufficiently rigorous data analysis (item 8), most of which (18 out of 28 studies) were due to the researchers not explaining how the data presented were selected from the original sample (criteria 8.3). In terms of the clear statement of findings (item 9), ten studies met all three criteria while 18 did not mainly because the researchers did not provide adequate discussion on the evidence both for and against the researcher's arguments (criteria 9.2). Finally, 17 of the 28 studies satisfied all the criteria in the item determining the value of the research (item 10). The rest were limited because they neither identified new areas where future research was necessary (criteria 10.2) nor discussed the transferability issues (criteria 10.3).

To be noticed, none of the studies provided information on criteria 5.5 (If methods were modified during the study. If so, has the researcher explained how and why?), and only two studies mentioned criteria 6.2 (How the researcher responded to events during the study and whether they considered the implications of any changes in the research design?). This might could be due to lack of descriptions or because these two events did not occur at all.

5.4 Synthesis of the Antecedents and Consequences of Health literacy or Numeracy

Extracted from the Included Studies

As discussed in **Section 4.5**, the synthesis was mainly achieved by thematic identification via the framework approach. We built our working thematic framework based on Sorensen's et al., (2012) comprehensive framework. Antecedents (Appendix S) and consequences (Appendix T) were charted to the matrices, respectively, in which the row represented an individual study and the column denoted to the themes. Additionally, we also conducted the thematic analysis inductively when the antecedents or consequences did not fit any pre-existing themes in the working thematic framework.

5.4.1 Themes Identified from Antecedents of Health Literacy or Numeracy

In our working thematic framework, there were two major themes for antecedents of health literacy: personal antecedents, and societal and environmental antecedents.

5.4.1.1 Personal Antecedents

Under personal antecedents, there were two sub-themes – personal characteristics and personal competence. According to Sorensen et al., (2012), personal characteristics were exemplified by “*age, race, gender, cultural background, socioeconomic status, education, occupation, employment, income*”, while personal competence “*vision, hearing, verbal ability, memory and reasoning, physical abilities and social skills, and general literacy skills*” (page 10).

5.4.1.1.1 Personal Characteristics

Among 24 included studies which investigated antecedents, 21 studies involved personal characteristics, including a variety of factors such as immigrant-related factors (i.e., acculturation, duration of residence in the destination country, age at the time of immigration, country of origin,

and primary language), age, gender, marital status, health status, economic status (including income, employment, and insurance status), educational attainment.

Acculturation

As the level of acculturation increased, the level of health literacy increased (S. E. Choi et al., 2013; Gele et al., 2016; Todd et al., 2011b). However, higher level of acculturation stress was a predictor of limited health literacy (Y. J. Choi et al., 2016).

Duration of residence in the destination country

The results regarding the association between duration of residence in the destination country and health literacy were mixed. Khuu et al., (2018) showed that length of stay was positively associated with health literacy. Mantwill et al., (2017) revealed that duration of residence was positively associated with functional health literacy, when assessed with language-dependent measures (i.e., the Chew items, which asked participants to self-report their confidence in understanding and filling out medical information in the language of the host country), but not language-independent measures (i.e., the S-TOFHLA). Their results suggested that the duration of residence might impact immigrant's health literacy level via proficiency in language of the host country.

Seven additional studies, however, found the association between duration of residence and health literacy not statistically significant (Beltran et al., 2016; Diamond et al., 2014; Gele et al., 2016; H. Y. Lee et al., 2014; Thomson et al., 2011; Wangdahl et al., 2014; Wister et al., 2010).

Only one study explicitly explored the relation between duration of residence and health numeracy. Gatobu et al., (2014) showed that years of residency in Canada was positively associated with numeracy in terms of context-free numeracy, which was measured by pure

mathematical addition tasks (i.e., the French kit addition task). However, no significant associations were detected when numeracy was evaluated by the French kit addition, subtraction-correction task or the S-TOFHLA numeracy component.

Age at the time of immigration

Mantwill et al., (2017) showed that age at the time of immigration was inversely associated with health literacy as measured by the language-dependent measure (i.e., the Chew items), but not correlated with health literacy, which was assessed by the language-independent measure (i.e., the S-TOFHLA).

Country of origin

Being a Somali refugee in Sweden was more likely to have inadequate health literacy, compared to refugees from countries such as Iraq, Syria, and Afghanistan (Wangdahl et al., 2014).

Primary language

In their study investigating the role of mathematical self-efficacy in predicting numeracy and health numeracy, Gatobu et al., (2014) found that primary language was also a contributor to the differences in numeracy performance between groups: Mandarin speakers outperformed Kikuyu speakers in the French kit numeracy tasks. However, primary language was not associated with health numeracy, as measured by the S-TOFHLA numeracy component.

Gatobu et al., (2016) further explored the language difference in numeracy, health literacy, and health numeracy performance between Mandarin and Kikuyu speakers. Results showed that being a Mandarin speaker was associated with better performance in numeracy measured by the addition and addition, subtraction-correction tasks of the French kit when compared with a

Kikuyu speaker. However, Kikuyu speakers outperformed Mandarin speakers in health literacy tests, measured by the S-TOFHLA reading comprehension component. No statistical differences were found between these two languages in health numeracy levels assessed by the S-TOFHLA numeracy component or the NVS.

Age

Age was explored by one qualitative and 10 quantitative included studies. The only qualitative study indicated that age-related limitation such as vision and hearing problems was a barrier to health literacy (Leung et al., 2014). Four quantitative studies (S. E. Choi et al., 2013; H. Y. Lee et al., 2014; Thomson et al., 2011; Todd et al., 2011b) showed a statistically significant negative association between age and health literacy, and one (Jacobson et al., 2016) indicated a similar association between age and health numeracy. However, there were also some quantitative studies suggesting that age was not associated with health literacy (Gele et al., 2016; Jacobson et al., 2016; Khuu et al., 2018; H. Y. Lee et al., 2012; Ng et al., 2014; Thomson et al., 2011; Todd et al., 2011b; Wangdahl et al., 2014).

To be noticed, two of the above quantitative studies used different assessment instruments or different components of one instrument to evaluate health literacy or numeracy level in the same research sample and yielded different conclusions on the associations between age and health literacy or numeracy (Thomson et al., 2011; Todd et al., 2011b). Thomson et al., (2011) showed an inverse association between age and health literacy when health literacy was measured by the S-TOFHLA, but not by the REALM. In Todd's et al., (2011b) study, age was a predictor of health literacy measured by the S-TOFHLA, but not associated with health literacy as assessed by the colon cancer Cloze test.

Gender

There were eight quantitative studies investigating gender. Being a woman was a predictor of higher health literacy in three studies (Becerra et al., 2017; H. Y. Lee et al., 2012; H. Y. Lee et al., 2015a). However, four quantitative studies also indicated that gender was not associated with health literacy (S. E. Choi et al., 2013; H. Y. Lee et al., 2014; Ng et al., 2014; Wangdahl et al., 2014). Gatobu et al., (2014) showed that being a woman was not associated with numeracy, health literacy, or health numeracy.

Marital status

Five quantitative studies investigated marital status as an antecedent. Only one showed that being married was a predictor of low health literacy (H. Y. Lee et al., 2012); the rest showed no association between marital status and health literacy (Beltran et al., 2016; Khuu et al., 2018; H. Y. Lee et al., 2015a; H. Y. Lee et al., 2014).

Health status

Three quantitative studies involved existing health status. Khuu et al., (2018) suggested that health literacy levels could differ significantly based on self-reported health status, whereas H. Y. Lee et al., (2014) and Wangdahl et al., (2014) found that there was no significant association between health status (i.e., self-reported health status, number of chronic diseases, and long term sickness) and health literacy.

Economic status

Nine quantitative and two qualitative research involved economic status including income, employment, and insurance status. The two qualitative studies identified that not having health insurance (Leung et al., 2014) and concerns about the medical costs (Taiwo, 2013) were two

barriers to adequate health literacy. Some quantitative studies showed that living in poverty (Becerra et al., 2017) and unemployment (Gele et al., 2016) were statistically associated with low health literacy while higher income (S. E. Choi et al., 2013; Ng et al., 2014) and having health insurance (H. Y. Lee et al., 2012) were predictors of higher health literacy. There were also four quantitative studies showing no statistical associations between health insurance, income, employment and health literacy level (Beltran et al., 2016; Khuu et al., 2018; H. Y. Lee et al., 2014; Thomson et al., 2011).

Educational attainment

Ten quantitative studies presented the relationship between educational level and health literacy. Seven of them showed a positive association (Beltran et al., 2016; S. E. Choi et al., 2013; H. Y. Lee et al., 2012; H. Y. Lee et al., 2015a; H. Y. Lee et al., 2014; Ng et al., 2014; Wangdahl et al., 2014). Another two quantitative studies also suggested a positive association, which depended on the types and language versions of the assessment instruments for health literacy (Thomson et al., 2011; Todd et al., 2011b). For example, educational attainment was positively associated with health literacy measured by the S-TOFHLA and the Chinese version of the colon cancer Cloze test, but not with health literacy assessed by the English version of the Cloze test (Todd et al., 2011b). Only one quantitative study concluded that educational level did not predict health literacy (Gele et al., 2016).

One quantitative study determined the role of educational level in numeracy or health numeracy (Gatobu et al., 2014). Gatobu et al., (2014) found that university-level education was positively associated with numeracy in the model where numeracy was measured by the French kit addition subtraction-correction task. However, educational attainment was not correlated with

health numeracy when it was assessed by instruments such as the S-TOFHLA numeracy component.

5.4.1.1.2 Personal Competence

Sorensen et al., (2012) considered personal competence as individual abilities which could be exemplified by memory and reasoning, physical abilities and social skills, and cognitive skills. Out of 24 studies investigating personal antecedents, 11 explored personal competence, including proficiency in the official language of the destination country (8 studies), health literacy in different domains or different dimensions of health literacy (2 studies), and having difficulties with activities in daily life (1 study).

English proficiency

The official language of the destination country all referred to English in the eight included studies. Seven of them were quantitative studies, most of which showed that low English proficiency predicted a limited health literacy (Becerra et al., 2017; S. E. Choi et al., 2013; Jacobson et al., 2016; H. Y. Lee et al., 2012; H. Y. Lee et al., 2015a). However, there were some exceptions. H. Y. Lee et al., (2014) found that English proficiency was not a predictor of cancer literacy, measured by 12 questions about cancer risk. Results from Todd et al., (2011b) indicated that the association between English proficiency and health literacy differed depending on the assessment tools used: Higher English proficiency was associated with higher health literacy level measured by the S-TOFHLA but not by the Cloze test developed based on colon cancer prevention information from the Canadian Cancer Society. Finally, challenges with language and communication were perceived as barriers to health literacy in the qualitative study (Taiwo, 2013).

Only Jacobson et al., (2016) showed that higher English proficiency might be associated with higher numeracy sub-score of the S-TOFHLA.

Health literacy in different domains or different dimensions of health literacy

Interestingly, health literacy in some domains or dimensions were found to be antecedents of other domains or dimensions. For instance, Beltran et al., (2016) demonstrated that higher cervical cancer screening literacy was significantly associated with higher HPV literacy, whereas Wangdahl et al., (2014) presented that inadequate functional health literacy was associated with inadequate comprehensive health literacy.

Activities of daily living

Health literacy differed significantly between participants with or without difficulties with activities of daily living (Khuu et al., 2018).

5.4.1.2 Societal and Environmental Antecedents

Seven studies explored societal and environmental antecedents, including family history, family support, literacy practices at home or work, maternal education, community support, transportation issues, individual's support network, format of information.

Among the quantitative studies, family cancer history (Beltran et al., 2016; H. Y. Lee et al., 2014) and family support (H. Y. Lee et al., 2014) were found not to be significantly associated with health literacy. Immigrants who had literacy practice at home or work or whose mother had higher educational level tended to have adequate health literacy (Ng et al., 2014). Gatobu et al., (2014) found that information format was associated with health numeracy determined by the S-TOFHLA numeracy component.

Among the qualitative studies, Carroll et al., (2007) identified that community support (i.e., health education program) could exert positive effects on health literacy among Somali refugee women. Leung et al., (2014) found that transportation issues (e.g., taking public transportation in Los Angeles was inconvenient) and limited health information in Chinese-speaking communities were barriers to health literacy. Murry et al., (2018) showed that involving a person's support network and providing instruction not only in written format could be perceived as two ways to build health literacy for medication management.

5.4.1.3 Personal Belief, Experience, and Behaviour

We also conducted the thematic identification inductively for those antecedents which did not fit our pre-existing themes. As a result, the theme “personal belief, experience, and behaviour” was identified in eight quantitative and three qualitative studies.

Personal belief

Antecedents relevant to personal belief included self-efficacy, healthcare motivation, cultural modesty, unawareness of self-care responsibility, a desire to avoid being burdensome to others, high regard for authority, a desire to be together or follow a collective approach (e.g., highly valuing the experiences of peers), and religio-cultural beliefs and assumptions on health.

Mathematical self-efficacy determined by the Mathematics Self-Efficacy Scale (MSES) or the Subjective Numeracy Scale (SNS) was positively associated with numeracy measured by the French kit. Mathematical self-efficacy, when assessed by the SNS, not by the MSES, was also positively associated with health numeracy measured by the S-TOFHLA numeracy component (S. Gatobu et al., 2014).

H. Y. Lee et al., (2015a) found that having a low level of cultural modesty toward healthcare professionals were one of the predictors in having a higher health literacy level in the Korean American immigrants. On the other hand, in the same population, H. Y. Lee et al., (2014) showed that health care motivation was not associated with health literacy.

In a qualitative study done by Leung et al., (2014), unawareness of self-care responsibility, high regard for authority, and desiring to be together or following a collective approach belonged to the perceived barriers to health literacy. In another qualitative study (Taiwo, 2013), barriers to health literacy included religio-cultural beliefs and assumptions on health.

Personal experience

Experiences prior to arrival at the destination country can also affect health literacy level. The qualitative study done by McMichael et al., (2009) suggested that experiences such as disrupted schooling, long periods of time living in refugee camps, limited or no access to primary health care, experiences or threats of sexual violence and fragmented family life might have impact on sexual health literacy in refugees to Australia.

Personal behaviour

Health behaviours were explored in seven quantitative studies. Some of the health behaviours were significantly associated with health literacy. For instance, a higher number of doctor visits (Beltran et al., 2016), attending social or religious group (Khuu et al., 2018), having a primary care physician (H. Y. Lee et al., 2014 and 2015a), and participation in adult education (Ng et al., 2014) was correlated with a higher health literacy level. Health literacy was also reported to be positively associated with the Internet use (when health literacy was measured by

the S-TOFHLA) and television use (when health literacy was measured by the S-TOFHLA and REALM) (Thomson et al., 2011).

On the other hand, the associations between some behaviours and health literacy were not statistically significant. They included fruit intake, vegetable intake, and soda intake (S. E. Choi et al., 2013), having a usual source of care and annual health check-up (Beltran et al., 2016), usual place of care (Khuu et al., 2018), smoking frequency and exercise frequency (H. Y. Lee et al., 2014), and the Internet use (when health literacy was measured by the REALM and the colon cancer Cloze test) (Thomson et al., 2011).

5.4.2 Themes Identified from Consequences of Health Literacy or Health Numeracy

5.4.2.1 Health Service Use or Behaviour

Consequences from 29 included studies, including 13 quantitative and 16 qualitative studies, were categorized in the theme “health service use or behaviour.”

Three quantitative studies explored associations between health literacy and health care use. Coffman et al., (2007) showed that interaction between income and health literacy was a statistically significant predictor of health literacy in Latinos from Mexico and other Central and South American countries living in the USA. To be specific, participants with low health literacy and high income tended to use more health care, including prescription medications, screenings, laboratory and medical procedures, emergency care, and mental health services. The possible explanation to this association given by Coffman et al., (2007) was that low health literacy might cause more improper use of health services. Coffman et al., (2012) again found that a positive association between health literacy and health care use in another group of Latino immigrants to the USA. However, Coffman et al., (2012) argued that this positive association should be

cautiously interpreted because the participants, many of whom had no health insurance and disease risk factors, might overuse health care services such as emergency care.

Compared to refugees with adequate health literacy, those with inadequate health literacy tended to refrain from seeking health care services (Wangdahl et al., 2018). Additionally, four qualitative studies showed that literacy was one of the main obstacles for immigrants to access health care or health services (Clark et al., 2014; Filippi et al., 2014; Floyd et al., 2017; Hurley et al., 2013).

Four quantitative studies indicated that higher health literacy was associated with greater participation in cancer screening, such as the increased participation of the Pap test for cervical cancer (Idehen et al., 2017; Kim et al., 2018; H. Y. Lee et al., 2015c) and the sigmoidoscopy for colorectal cancer (Ko, 2014). Kim et al., (2018) further showed that the influence of health literacy on the Pap test participation was an indirect effect, mediated by decisional balance (i.e., weighing the pros and cons to adopt a target behavior) and self-efficacy (i.e., how much confidence a woman has in carrying out Pap test-related tasks).

Additionally, quantitative results from Todd et al., (2011a) suggested that health literacy levels were significantly higher in Chinese immigrants to Canada who had ever screened than those who had never screened for colon cancer, or in immigrants who were current colon cancer screeners (completed a fecal occult blood test in the past two years or a colonoscopy or sigmoidoscopy in past 5 years) than non-current screeners. However, there was no statistically significant difference in health literacy levels between Chinese immigrants who were current breast cancer screeners (completed a mammogram in the past two years) and non-current screeners (Todd et al., 2011a). Only quantitative study done by Ko, (2014) found no association

between health literacy and the participation of the fecal occult blood test or the colonoscopy for colorectal cancer.

Four qualitative studies suggested that lack of health literacy was a barrier to the participation of cervical cancer screening (Gele et al., 2017; F. H. Lee et al., 2014) and colorectal cancer screening (Jung et al., 2017; Woudstra et al., 2016). However, one qualitative study revealed that low health literacy level was not a barrier to cancer screening (Jafri, 2012).

Other than cancer screening, inadequate health literacy level was also perceived as a barrier to cardio-metabolic health check (Groenenberg et al., 2015), engagement in general preventive health activities (Oliver, 2015), and preventive screening for hepatitis B (Sriphanlop et al., 2014).

Health literacy was shown to be inversely correlated with health and disease management, such as glycemic control (Thabit et al., 2009) and HIV disclosure to a steady partner (Kankou et al., 2017). In the relevant qualitative research, Alzayer et al., (2017) found that low health literacy might affect asthma control, while Murry et al., (2018) identified lack of health literacy as a barrier to medication management.

Two quantitative studies showed an association between health literacy and dental care use. The results were mixed, though. Calvasina et al., (2016) suggested a positive association, whereas Geltman et al., (2014) found no statistical association.

Three studies (one quantitative and two qualitative) explored the impacts of parental health literacy on children's disease prevention. Hernandez-Mekonnen et al., (2016) showed that low maternal health literacy was associated with increased odds of having a child at risk of development delay, but not with the use of early intervention services. Parents considered low

health and food literacy as a major obstacles of child obesity prevention (Cyril et al., 2017) and perceived that low pre-migratory obesity literacy was negatively affecting their capacity to be involved in childhood obesity prevention programs (Renzaho et al., 2017).

5.4.2.2 Health Outcome

There were 17 studies, all of which were quantitative, investigating the associations between health literacy and health outcomes such as health status and mental health.

Four quantitative studies indicated that health literacy had a positive relationship with health status, including physical functioning and health perception, self-rated health, and self-reported health status (Bekker et al., 2004; H. Y. Lee et al., 2015b; Tsoh et al., 2016; Wangdahl et al., 2018). Moreover, Mantwill et al., (2017) suggested that general health status was associated with language-dependent health literacy (i.e., the Chew items) rather than with language-independent health literacy (i.e., the S-TOFHLA). Numeracy was also reported to be positively associated with health status in one quantitative study (Yunusa Vakkai, 2016).

However, there were also five quantitative studies concluding that health literacy or literacy (Koch-Weser et al., 2006; D. W. Omariba et al., 2011; D. W. R. Omariba et al., 2015; Prins et al., 2015; Yunusa Vakkai, 2016), numeracy (Prins et al., 2015), or the interaction of literacy and numeracy (Yunusa Vakkai, 2016) was not significantly correlated with health status including self-rated health and disability.

In terms of mental health, Y. J. Choi et al., (2016) presented that levels of mental health (measured by the General Health Questionnaire (Y. M. Lee et al., 1999)) increased as levels of mental health literacy (assessed by the Mental Health Literacy Scale (Y. J. Choi et al., 2016)) increased. Coffman et al., (2010) and H. Y. Lee et al., (2015b) showed an inverse association

between health literacy and the depression score of the Center for Epidemiologic Studies Depression Scale and between health literacy and depression symptoms, respectively. Kim et al., (2013), on the contrary, indicated a non-significant association between literacy and depression.

Other health outcomes, including the relapse rate in children with nephrotic syndrome (Borges et al., 2017), waist to hip ratio and blood glucose level (Y. J. Choi et al., 2016), decayed, missing, and filled teeth and risk of periodontal disease (P. L. Geltman et al., 2013), showed an inverse correlation with health literacy. There were also some non-significant associations between health literacy and the body mass index (BMI), systolic blood pressure, and diastolic blood pressure (S. E. Choi et al., 2013) and between health literacy and untreated teeth decay (P. L. Geltman et al., 2013) or diabetes outcomes (Njeru et al., 2016).

5.4.2.3 Health Cost

No consequences from the included studies could be coded with the theme – “health cost,” suggesting potential research gap regarding the economic burden of inadequate health literacy and numeracy in immigrants and refugees.

5.4.2.4 Health Experience and Perception

We inductively identified the theme “health experience and perception” from nine included studies (3 quantitative and 6 qualitative studies) for the consequences that did not fit the pre-existing themes.

Two quantitative studies determined the roles of health literacy in perceived quality of care. Latino immigrants to the US with higher levels of health literacy tended to report higher Quality of Care (QoC) (Calvo, 2016). Interestingly, among immigrant women to Japan, Japanese language literacy was inversely correlated with care satisfaction and with Quality of Care for

Pregnancy, Delivery and Postpartum Questionnaire sub-scores, including Respect (feeling respected) and Understanding (feeling being understood by health care providers) (Igarashi et al., 2013). However, Japanese language literacy was positively correlated with the Cold (feeling distant and unhelpful attitude) sub-score (Igarashi et al., 2013).

Another quantitative study reported that refugees with inadequate comprehensive health literacy were more likely to have the experience of receiving poor quality of communication and little health care information and the experience of not receiving any new knowledge or any help (Wangdahl et al., 2015).

The perceived low quality of health care may be explained by the low capacity to communicate with health care providers, or different needs in and expectations toward health care services, both of which are related to health literacy. Three qualitative studies suggested that immigrants or refugees perceived low health literacy as an obstacle to communicating gastrointestinal illness concerns (Gregory, 2015), meeting healthcare needs (Marshall et al., 2010), and recognizing depression symptoms (Schoenmakers et al., 2017). Thomson et al., (2012) revealed that immigrants to Canada with functional or interactive health literacy had different needs and points of views toward diet-related cancer prevention. Results from Watts et al., (2014) suggested parental health literacy as one factor that could impact the attitudes toward use of contraception in young African refugees (Watts et al., 2014).

Lastly, a qualitative study by Todd et al., (2011c), identified three themes: sources of cancer information, barriers to cancer information seeking, and strategies used during cancer information seeking. However, none of them differed between low and high health literacy groups.

5.4.2.5 Health Knowledge and Understanding

“Health knowledge and understanding” was another theme identified inductively from one quantitative and three qualitative studies.

The quantitative study conducted by P. C. Smith et al., (2012) found that Spanish-speaking immigrants to the US with inadequate functional health literacy, which was measured by the TOFHLA, attended follow-up appointments as instructed less frequently and reported less understanding of emergency department instructions, as compared to those with adequate functional health literacy,

The three qualitative studies indicated that inadequate health literacy (or literacy) was one of the barriers to understanding general practitioners and pharmacists among refugees in Australia (Clark et al., 2014), to the knowledge and understanding of HIV/AIDS in Latino immigrants to the US (Shedlin et al., 2004), or to the understanding of health system among migrants in Norway (Smaland Goth et al., 2011).

5.5 Summary

In Chapter 5, we brought together a sizeable number of empirical studies on antecedents and consequences of health literacy or health numeracy in immigrants and refugees, and examined the basic characteristics of these studies, assessed the risk of bias of each individual study, and narratively synthesized the antecedents and consequences under different themes.

We found that in the literature regarding health literacy and numeracy in the population of immigrants and refugees, a large proportion of empirical studies were quantitative, adopted a cross-sectional study designs, and investigated health literacy in immigrants. Furthermore, more studies focused on the consequences of health literacy or numeracy, compared to those focused on the antecedents. Three antecedent-related themes (sub-themes) included personal antecedents (personal characteristics and personal competence), societal and environmental antecedents, and personal belief, experience and behaviour. Also, we coded the consequences of health literacy and numeracy with five themes including health service use or behaviour, health outcome, health cost, health experience and perception, and health knowledge and understanding. Most studies centered on the consequences relating to health service use or behavior and health outcome, whereas no research investigated the potential associations between health cost and health literacy or numeracy.

Appendix U and V present detailed findings in regards to the antecedents and consequences from the included studies, respectively. They provide a summary for each individual study, containing information such as statistical methods, indicators and strengths of associations, p-values, adjusted covariates, and themes (qualitative studies).

Chapter 6: Discussion

6.1 Discussion of the Study Design and Sampling Methods

Overall, among 49 included quantitative studies, 47 adopted the cross-sectional study design, in which the researchers established potential associations at a specific point in time. Also, a majority of the quantitative studies used a non-probability approach (e.g., convenience sampling, purposive sampling, and snowball sampling) instead of a probability method (e.g., random sampling).

The almost unanimous selection of the cross-sectional study design and the wide adoption of non-probability sampling methods might be due to the “hard-to-reach” or the “hidden” characteristics of the immigrant and refugee population. A “hard-to-reach” population refers to the sub-group of a population that is difficult to get involved in research due to reasons such as physical location and social and economic situation (Shaghghi et al., 2011). The “hidden” population refers to those disadvantaged populations just wishing not to be found (Atkinson et al., 2001; United States National Institute on Drug Abuse, 1990). The immigrant and refugee population is usually culturally and linguistically diverse, but sometimes invisible from the mainstream society, and often hard to reach (Sulaiman-Hill et al., 2011).

The cross-sectional study design is a relatively quick and convenient way to develop and conduct research. Researchers can investigate several exposures or outcomes at the same time and do not need to worry about the risk of loss to follow-up of these hard-to-reach immigrants and refugees because the data collection usually happens once. However, cross-sectional research design has many limitations which also apply to our included studies. For instance, cross-sectional design could not establish the temporal relation between antecedents or consequences and health literacy or numeracy. Moreover, the validity of the results from cross-

sectional design could often be impacted by the non-response bias, if the characteristics of the participants who consent to participate in the study differ from the features of those who do not participate. Most of the included research did not provide information on the comparison of the characteristics between the respondents and non-respondents, posing a high risk of the non-response bias. To address all the above potential flaws, further investigations may need more complex study designs, such as cohort studies and case-control studies.

We need to be cautious about the statistical validity of the quantitative findings from a sample recruited with non-probability sampling methods since the samples might not be representative of the target population. However, given the hard-to-reach characteristic of immigrants and refugees, non-random sampling approach is usually more feasible and practical than the probability sampling methods. Thus, selecting the recruitment strategy that will work best for the target population should be a key. Many studies in the literature have discussed the sampling challenges in the context of hard-to-reach or hidden population (Boulos et al., 2015; Reichel et al., 2017; Shaghghi et al., 2011). For example, Shaghghi et al., (2011) reviewed the approaches which have been adopted to recruit hard-to-reach population in studies and suggested that the correct and successful selection of sampling methods depended on researchers' knowledge on the specific features of the hard-to-reach target population within the larger population.

6.2 Understanding the Concept of Health Literacy in the Included Studies

Most of the included studies centered on health literacy, rather than on health numeracy. Sorensen's et al., (2012) "all-inclusive" definition of health literacy (see **Chapter 1**) perceives health literacy as an entity built upon the concept of literacy but beyond literacy towards the goal of empowering individuals to have control over their health. In order to reflect the complex relationship between health literacy and literacy, we decided to set the boundary of our inclusion criteria to "literacy in health contexts" in the thesis research rather than limiting to the precise term "health literacy".

Our literature search results justified our decision. We found that there were some studies done with immigrants and refugees just investigating the role of literacy (e.g., literacy in Finnish/Swedish language) in health contexts. In addition, many included studies used diverse terms other than applying the exact term "health literacy" to represent health literacy, such as HPV literacy (Beltran et al., 2016), dementia literacy (Diamond et al., 2014), obesity literacy (Renzaho et al., 2017), and cancer literacy (Ko, 2014; H. Y. Lee et al., 2014). We did not know how Beltran et al., (2016), Diamond et al., (2014), Ko, (2014) and H. Y. Lee et al., (2014) defined HPV literacy, dementia literacy and cancer literacy, respectively, due to lack of descriptions. However, Renzaho et al., (2017), in their own research, defined obesity literacy as *"the individuals and communities' set of knowledge, skills, and abilities needed to understand the importance of maintaining a healthy weight, to recognize the impact of lifestyle and food choices on their weight, and to make appropriate decisions to address weight-related health issues"* (page. 2). The obesity literacy definition emphasized the context of obesity prevention and treatment as well as its related health consequences – maintaining a healthy weight and responding to weight-related health issues.

The variety in health literacy terms indicates that investigators in empirical research on health literacy among immigrants and refugees need more specific definitions of health literacy to reflect their targeted health contexts and related health consequences. However, current conceptual research on health literacy usually gives health literacy definitions in general and broad health contexts, which can be exemplified by Sorensen's et al., (2012) "all-inclusive" definition. Thus, researchers interested in conceptualizing health literacy may need to put more efforts on development of health literacy definitions for specific diseases or health issues in the future. Additionally, when applying a diverse term, empirical researchers should either cite a definition existing in the literature or clarify the term by themselves to help the readers know what their term stands for.

Health literacy is widely acknowledged as a multidimensional concept, which consists of several different dimensions of skills. For example, Nutbeam, (2000) classified health literacy into levels of functional health literacy (mainly referring to basic skills in reading and writing), interactive health literacy (mainly referring to more advanced cognitive and literacy skills including social skills), and critical health literacy (mainly referring to the most advanced skills regarding critically information analysis and use). Frisch et al., (2012) also proposed five essential dimensions of health literacy: functional literacy (referring to word recognition, comprehension and numeracy skills), factual and procedural knowledge (the former refers to context-specific knowledge, while the latter to knowledge on how to use factual knowledge), awareness (mainly referring to the ability to recognize a problem and beware of the need for help), critical dimension (referring to skills similar to Nutbeam's, (2000) critical health literacy skills), as well as affective dimension and attitudes (referring to individual's affection and attitude, such as individual's motivation).

Around 30 included studies explored the functional dimension of health literacy. They either explicitly used the term functional health literacy (e.g., P. C. Smith et al., 2012; Todd et al., 2011a) or applied relevant assessment instruments which are designed to measure functional health literacy skills, such as the TOFHLA (measuring reading comprehension skills), S-TOFHLA (measuring reading comprehension skills), and REALM (assessing word recognition abilities) (e.g., Bekker et al., 2004; Igarashi et al., 2013). This is in line with what has been noticed by some researchers that in the field of health literacy research, a considerable number of empirical studies were devoted to the functional dimension of health literacy (e.g., Diviani et al., 2014; Friedman et al., 2009).

However, health literacy is a multi-dimensional concept. The functional dimension cannot represent the whole health literacy concept; adequate functional health literacy does not necessarily mean sufficient health literacy. For example, results from Friedman et al., (2009) showed that participants only had very limited understanding of prostate cancer risk factors and preventive behaviours although they possessed adequate functional health literacy skills. In our included studies, Thomson et al., (2012), based on their qualitative findings, pointed out that functional health literacy was crucial but not sufficient for immigrants to Canada to actively engage in diet-related cancer prevention. Interactive and critical health literacy were also important and essential. Moreover, Wangdahl et al., (2014, 2015) investigated functional health literacy and comprehensive health literacy (referring to Sorensen's et al., (2012) "all-inclusive" definition) among refugees, and they found that functional health literacy was a positive predictor of comprehensive health literacy, and it was limited comprehensive health literacy, not low functional health literacy, that was associated with poor experience of health examinations for refugees. All the evidence suggests that putting too much focus or even equating health

literacy with functional health literacy is dangerous because it could lead existing health literacy gaps to go undetected. Thus, current associations established based on functional health literacy in immigrants and refugees may need to be further substantiated in other dimensions of health literacy. Additionally, over-concentrating on the functional dimension of health literacy could cause the ignorance of other health literacy dimensions, which as Diviani et al., (2014) proposed, would be very helpful for us to eventually understand the causal relation between health literacy and health outcomes.

There are also some included studies addressing health literacy beyond the functional health literacy dimension by applying assessment instruments that were designed for assessing other dimensions of health literacy skills. These instruments included knowledge tests which evaluate the knowledge of participants on a specific health topic or screening questions that require participants to self-report perceived ability on a health-related topic. Knowledge tests or screening questions are of great use because they evaluate health literacy beyond word recognition and reading comprehension abilities (i.e., functional health literacy skills). However, we should notice that these two types of measures have respective disadvantages (Frisch et al., 2012). First, the health literacy level determined based on individual's knowledge in a very specific health context may not be generalizable to a different health context. Second, screening questions ask participants about their perceived competence or confidence in completing health literacy tasks instead of their actual performance. This may affect the validity of the health literacy level assessment.

Finally, it is notable that among the included studies centering on the dimension of functional health literacy, the associations between the antecedents or consequences and health literacy in the same research sample could vary according to different assessment instruments

adopted. For instance, age was inversely correlated with health literacy level determined by the S-TOFHLA but not by the REALM (Thomson et al., 2011). The results on the relationship between age and health literacy seemed contradictory. However, the S-TOFHLA determines health literacy level by measuring reading comprehension and numeracy skills, while the REALM focused on the word recognition abilities. Thus, given the knowing finding that increasing age after adulthood is associated with cognitive declines (Deary et al., 2009), the contribution to the statistically significant association between age and health literacy, as assessed by the S-TOFHLA, might come from the impact of older age on comprehension skills. This situation suggests that when drawing conclusions about associations between antecedents or consequences and health literacy in the future, the specific skills tapped by the measurement instruments may need to be taken under consideration.

6.3 Understanding the Concept of Health Numeracy in the Included Studies

In 2005, Golbeck et al. argued that empirical evidence on health numeracy was limited and required, especially on the role of health numeracy in health outcomes (Golbeck et al., 2005). However, our systematic review showed that health numeracy research of immigrants and refugees still has not received adequate attention in the population of immigrants and refugees. Only six out of 77 dealt with numeracy or health numeracy in their Results sections (Borges et al., 2017; S. S. Gatobu et al., 2014; Gatobu et al., 2016; Jacobson et al., 2016; Prins et al., 2015; Yunusa Vakkai, 2016), among which three explored health consequences of numeracy or health numeracy (Borges et al., 2017; Prins et al., 2015; Yunusa Vakkai, 2016).

Borges et al., (2017) explored parental numeracy in outcomes of childhood nephrotic syndrome and showed that parental numeracy was not associated with children's health outcomes. However, this finding resulted from looking at the total study participants, including

both immigrant and non-immigrant parents, and could not be separated. Prins et al., (2015) and Yunusa Vakkai, (2016) both investigated the effect of numeracy in self-rated health status, but showed opposite results. The former found that numeracy measured by the numeric component of the Program for the International Assessment of Adult Competencies was not statistically associated with self-rated health status in Asians, Hispanics and other immigrants to the US, whereas the latter showed self-reported numeracy level was a predictor of self-rated health in immigrants to the US (66.4% were Hispanics). Prins et al., (2015) discovered that region of residence was the major reason for the non-significant association between numeracy and self-reported health status because immigrants to the USA tend to live in regions such as the West where self-reported health is significantly better in population than regions such as the Northeast or Midwest. Yunusa Vakkai, (2016), on the other hand, did not include the region of residence as a variable in her research. Thus, more research on this topic is required in the future, and region of residence should be involved as an important covariate.

Currently, health numeracy is usually conceptually subsumed under health literacy. However, Golbeck et al. (2011), based on their empirical evidence, emphasized the importance of recognition of health numeracy as an integral yet distinct construct of health literacy. In our systematic review, 21 studies applied assessment instruments (e.g., the S-TOFHLA and TOFHLA) which determine the overall health literacy level based on individual's performance on health literacy skills (e.g., reading comprehension) and numeracy abilities (e.g., quantitative skills). In these studies, however, many researchers did not make efforts to specify the contribution of the numeracy skills to the associations under investigation. This indicates that these researchers may not perceive health numeracy as a distinct entity from health literacy, and it raises the danger of ignoring the critical roles of health numeracy in health contexts. In the

future, researchers should try to explore the impacts of the health numeracy component in an assessment instrument evaluating both health literacy and numeracy skills.

6.4 Implications of the Thematic Identification

Due to the importance of health literacy and numeracy in individual and population health, it is critical to understand the concepts of health literacy and numeracy and the essential components involved in the concepts.

On the one hand, many researchers have made great efforts to conduct theoretical research to conceptualize health literacy and numeracy from various perspectives. For instance, Nutbeam, (2008) interpreted health literacy from both the clinical risk and the personal asset points of view; I. M. Lipkus et al., (2009) provided a numeracy theoretical framework from the perspective of medical decision making and risk communication. However, none of the interpretations specifically addressed issues of health literacy or numeracy in the context of immigrant and refugee health. On the other hand, the number of empirical studies in the discipline of health literacy or numeracy among immigrants and refugees has been gradually increasing in recent years.

The present systematic review addressed the above imbalance between the theoretical and empirical research and contributed to the improvement of current understandings of health literacy and numeracy in the context of immigrant and refugee health by synthesizing relevant quantitative and qualitative evidence via thematic identification.

From the standpoint of thematic identification process, the themes were classified into two categories: a) themes present in, and deductively identified from Sorensen's et al., (2012) comprehensive framework for health literacy and numeracy, including antecedent- (e.g., personal

characteristics, personal competence, societal and environmental antecedents), and consequence-relevant themes (e.g., health service use and behaviours, health outcome, and health cost); and b) inductively identified themes from the data, which do not exist in Sorensen's et al., (2012) framework, including antecedents, such as "personal belief, experience, and behavior" as well as consequences, such as "health experience and perception", and "health knowledge and understanding".

The value of the deductive process in understanding health literacy and numeracy framework in the context of the immigrant and refugee health was that it mapped the evidence from each empirical study to the corresponding components currently existing in Sorensen's et al., (2012) theoretical framework. The empirical evidence warranted the essence of these already existing components for researchers interested in conceptualizing health literacy and numeracy aiming at the immigrant and refugee health. However, we did not find any empirical evidence that could be charted to one of the Sorensen's et al., (2012) components – "health cost." Given the healthcare cost is major public expenditure in many immigration countries, the research of the economic burden of inadequate health literacy and numeracy of immigrants and refugees is of great public health implication. Future empirical research should be directed to the association between health cost and health literacy or numeracy in immigrants and refugees.

The inductively generated themes provided empirical evidence that did not fit in Sorensen's et al., (2012) model. For instance, the antecedents "number of doctor visits" (Beltran et al., 2016) and "experiences prior to arrival at destination country" (McMichael et al., 2009) seemed neither appropriate for personal characteristics nor for personal competence. Thus, we categorized them to the new theme "personal belief, experience, and behaviour." Also, consequences such as "having experienced poor quality of communication" (Wangdahl et al.,

2015) and “knowledge and understanding of HIV/AIDS” (Shedlin et al., 2004) did not fit Sorensen’s et al., (2012) classification of consequences. As a result, two new themes “health experience and perception” and “health knowledge and understanding” emerged from our analysis. These inductively identified themes suggest that in the context of immigrant and refugee health, new components regarding individual’s experience, belief, perception, and knowledge should be added to the antecedent and consequence domains of Sorensen’s et al. (2012) conceptual framework. Additionally, we noticed that these antecedents and consequences (especially the consequences) mainly came from qualitative studies. Therefore, the addition of antecedent- and consequence-related components will be of great use for Sorensen’s et al. (2012) framework to guide future qualitative research on health literacy and numeracy in the context of immigrant and refugee health.

Another finding from our thematic identification was that some factors were relatively unique to the population of immigrants and refugees (e.g., antecedents including acculturation, duration of residence in the destination country, age at the time of immigration, country of origin, and primary language; consequences such health needs and concerns), while others were shared by both the immigrants and refugees and the non-immigrant general population (e.g., age, gender, marital status, health status, economic status, education, health service use, and health outcomes). This finding has important implications: these unique antecedents and consequences may be the key and main contributing factors to the differences in health literacy or numeracy level or in the prediction of consequences between immigrants and the non-immigrant general population. They need more attention from future researchers who will try to explore or investigate health literacy and numeracy at both the conceptual and empirical level in the context of immigrant and refugee health.

6.5 Strengths and Limitations of the Systematic Review

The first major strength is that this systematic review is a useful information source (e.g., information on risk of bias of each individual study and key findings regarding antecedents and consequences among immigrants and refugees) for future researchers who are interested in the area to understand the progress, potential gaps, and possible directions. A second main strength is the inclusion of as many relevant empirical studies as possible by gathering both quantitative and qualitative evidence, by searching the terms “literacy” and “numeracy” instead of “health literacy” and “health numeracy”, by including immigrants and refugees aged 16 and older without any other restrictions (e.g., restrictions on country of origin, ethnicity, gender, or primary languages). Finally, this review synthesizes the relevant evidence by thematic identification guided by the working thematic framework derived from an existing health literacy and numeracy conceptual model (i.e., Sorensen’s et al. (2012) model). As a result, the systematic review connects the studies done at the empirical level and the conceptual level. Empirical evidence can substantiate the existence of variables involved in the current conceptual model and offer insights on the connections between variables. On the other hand, the conceptual model can guide future research by providing variables that have not been investigated by current empirical studies done in immigrants and refugees.

There are several limitations. First, the present systematic review narratively summarized the empirical evidence on antecedents and consequences of health literacy and numeracy among adult immigrants and refugees, but it did not provide quantitative synthesis of the evidence. However, this limitation does not lessen the potential value of the thesis research because the present systematic review focused on a wide scope which was to offer the readers a picture of current status of empirical research conducted in adult immigrants and refugees. Additionally,

the thesis research is valuable in terms of informing researchers on the feasibility of conducting a systematic review to quantitatively summarize evidence on a specific antecedent or consequences in a more specific immigrant or refugee population.

Second, during the study selection, the thesis research excluded studies that involved interventions, programs, curriculums, or courses which aimed to improve health literacy or numeracy in immigrants and refugees. In Sorensen's et al., (2012) framework, factors on interventions could be considered as antecedents of health literacy and numeracy (e.g., societal and environmental antecedents). Thus, our exclusion might result in the bias that there were relatively fewer studies focusing on societal and environmental antecedents in immigrants and refugees. However, we excluded them mainly because a) the focus of the review was on naturally occurring antecedents and consequences and not on programs designed to improve health literacy and health numeracy. We believe that it would be better to conduct a systematic review specifically devoted to interventions. In such systematic review, the researchers might also address the effectiveness of an intervention; b) There is a recent systematic review on the topic, which was published in 2017 and entitled "Health literacy interventions for immigrant populations" (Fernandez-Gutierrez et al., 2017).

Third, there are many available frameworks for health literacy and health numeracy in the literature. Among these frameworks, our systematic review mainly borrowed Sorensen's et al., (2012) conceptual framework for the purpose of thematic analysis. As a result, we might have ignored other frameworks' perspectives and their different understandings of health literacy and health numeracy. However, we believe that Sorensen's et al., (2012) framework served best for our purpose because as a comprehensive framework integrating 12 existing frameworks, Sorensen's et al., (2012) framework contained as much information as possible.

Another limitation includes the potential language bias caused by the inclusion of literature only published in English. Provided that the English language currently is the predominant and preferred language among investigators in health research, studies containing positive results are more likely to be published in English, compared to those with negative results. Hence, English language bias may lead to a biased assessment in this systematic review.

Finally, although we searched eight databases, we did not search some data sources such as the grey literature.

6.6 Conclusion

The thesis research is the first systematic review that summarizes empirical evidence about antecedents and consequences of health literacy or numeracy in the immigrant and refugee population. The review gathered individual studies and collated useful information to answer what antecedents and consequences of health literacy or health numeracy in adult immigrants and refugees have been identified in the literature. Moreover, we provided information on the risk of bias of each individual study to help readers determine the validity of each included study and inform future research on possible improvements. Most importantly, in the thesis research, we identified themes based on an existing comprehensive framework for health literacy and numeracy and applied inductive thematic identification when the extracted evidence did not fit the framework. These themes are useful and valuable in terms of substantiating and further developing current conceptual frameworks for health literacy and numeracy from the standpoint of immigrant and refugee health.

Several key findings and recommendations are highlighted as follows: 1) A considerable number of empirical studies in adult immigrants and refugees were devoted to the functional health literacy dimension (i.e., reading and writing skills). Future research need to carry out more

research on other dimensions of health literacy (e.g., interactive health literacy and critical health literacy). 2) There was a lack of empirical studies on antecedents and consequences of health numeracy in adult immigrants and refugees. More relevant research is needed. 3) No empirical evidence could be coded with the theme “health cost.” More research on the associations between health cost and health literacy or health numeracy among adult immigrants and refugees is required. 4) Inductively identified themes such as “personal belief, experience, and behavior” for antecedents and “personal experience and perception” and “knowledge and understanding” for consequences did not exist in Sorensen’s et al., (2012) integrated framework. They should be added for the further development of the current framework in the context of immigrant and refugee health. 5) Most of the included quantitative studies in adult immigrants and refugees applied the cross-sectional study design and non-random sampling methods. Future researchers could select more advanced study designs (e.g., cohort studies) and improve the sampling methods to increase the validity of their research. 6) Immigrant- and refugee-specific antecedents (e.g., primary language, acculturation, duration of residence in the destination country) and consequences (e.g., some health needs) may play key roles in the health literacy and numeracy difference between immigrants/refugees and non-immigrants/non-refugees, thus should draw more attention in future empirical research.

References

- Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs AMA. (1999). Health literacy: report of the council on scientific affairs. *J Am Med Assoc*, 281(6), 552-557.
- Adhikari, R., & Sanou, D. (2012). Risk Factors of Diabetes in Canadian Immigrants: A Synthesis of Recent Literature. *Can J Diabetes*, 36(3), 142-150. doi:http://dx.doi.org/10.1016/j.jcjd.2012.06.001
- Aguirre, A. C., Ebrahim, N., & Shea, J. A. (2005). Performance of the English and Spanish S-TOFHLA among publicly insured Medicaid and Medicare patients. *Patient Educ Couns*, 56(3), 332-339. doi:10.1016/j.pec.2004.03.007
- Altin, S. V., Finke, I., Kautz-Freimuth, S., & Stock, S. (2014). The evolution of health literacy assessment tools: a systematic review. *BMC Public Health*, 14, 1207. doi:10.1186/1471-2458-14-1207
- Alzayer, R., Chaar, B., Basheti, I., & Saini, B. (2017). Asthma management experiences of Australians who are native Arabic speakers. *Journal of Asthma*, 1-10. doi:http://dx.doi.org/10.1080/02770903.2017.1362702
- Ancker, J. S., & Kaufman, D. (2007). Rethinking health numeracy: a multidisciplinary literature review. *J Am Med Inform Assoc*, 14(6), 713-721. doi:10.1197/jamia.M2464
- Atkinson, R., & Flint, J. (2001). Accessing Hidden and Hard-to-Reach Populations: Snowball Research Strategies. *Social Research Update*(33).
- Baker, D. W., Williams, M. V., Parker, R. M., Gazmararian, J. A., & Nurss, J. (1999). Development of a brief test to measure functional health literacy. *Patient Educ Couns*, 38(1), 33-42.
- Baker, D. W., Wolf, M. S., Feinglass, J., Thompson, J. A., Gazmararian, J. A., & Huang, J. (2007). Health literacy and mortality among elderly persons. *Arch Intern Med*, 167(14), 1503-1509. doi:10.1001/archinte.167.14.1503
- Becerra, B. J., Arias, D., & Becerra, M. B. (2017). Low Health Literacy among Immigrant Hispanics. *J Racial Ethn Health Disparities*, 4(3), 480-483. doi:10.1007/s40615-016-0249-5
- Beiser, M. (2005). The health of immigrants and refugees in Canada. *Can J Public Health*, 96 Suppl 2, S30-44.
- Bekker, M. H. J., & Lhajoui, M. (2004). Health and literacy in first- and second-generation Moroccan Berber women in the Netherlands: III literacy? *International Journal for Equity in Health*, 3 (no pagination)(8). doi:http://dx.doi.org/10.1186/1475-9276-3-8
- Beltran, R., Simms, T., Lee, H. Y., & Kwon, M. (2016). HPV Literacy and Associated Factors Among Hmong American Immigrants: Implications for Reducing Cervical Cancer Disparity. *Journal of community health*, 41(3), 603-611. doi:http://dx.doi.org/10.1007/s10900-015-0135-9
- Berkman, N. D., DeWalt, D. A., Pignone, M. P., Sheridan, S. L., Lohr, K. N., Lux, L., . . . Bonito, A. T. (2004). *Literacy and Health Outcomes. Evidence Report Technology Assessment no. 87 (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center under contract 290-02-0016.)*. Retrieved from Rockville, MD: <https://www.ncbi.nlm.nih.gov/books/NBK11942/>

- Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med*, *155*(2), 97-107. doi:10.7326/0003-4819-155-2-201107190-00005
- Bonevski, B., Randell, M., Paul, C., Chapman, K., Twyman, L., Bryant, J., . . . Hughes, C. (2014). Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. *BMC Med Res Methodol*, *14*, 42. doi:10.1186/1471-2288-14-42
- Borges, K., Sibbald, C., Hussain-Shamsy, N., Vasilevska-Ristovska, J., Banh, T., Patel, V., . . . Parekh, R. S. (2017). Parental Health Literacy and Outcomes of Childhood Nephrotic Syndrome. *Pediatrics*, *139*(3), 1-8. doi:10.1542/peds.2016-1961
- Boulos, R. J., & Borrás, P. A. (2015). Why Hard to Reach Populations are Hard to Reach: Lessons Learned from Accelerometer Data Collection among New Immigrants. *Adv Obes Weight Manag Control*, *3*(2), 00050. doi:10.15406/aowmc.2015.03.00050
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101.
- Cajita, M. I., Cajita, T. R., & Han, H. R. (2016). Health Literacy and Heart Failure: A Systematic Review. *J Cardiovasc Nurs*, *31*(2), 121-130. doi:10.1097/JCN.0000000000000229
- California Health Interview Survey. (2011). CHIS 2007 Adult Questionnaire (Translated). Retrieved from [http://healthpolicy.ucla.edu/chis/design/Pages/Questionnaires\(Translated\).aspx](http://healthpolicy.ucla.edu/chis/design/Pages/Questionnaires(Translated).aspx)
- Calvasina, P., Lawrence, H. P., Hoffman-Goetz, L., & Norman, C. D. (2016). Brazilian immigrants' oral health literacy and participation in oral health care in Canada. *BMC oral health*, *16*, 18. doi:http://dx.doi.org/10.1186/s12903-016-0176-1
- Calvo, R. (2016). Health Literacy and Quality of Care among Latino Immigrants in the United States. *Health & Social Work*, *41*(1), e44-e51. doi:10.1093/hsw/hlv076
- Canadian Council for Refugees. (2010). Refugees and Immigrants: A glossary. Retrieved from http://ccrweb.ca/sites/ccrweb.ca/files/glossary_en.pdf
- Canadian Council on Learning. (2008). *Health Literacy in Canada: A Healthy Understanding*. Retrieved from Ottawa, ON: <http://www.en.copian.ca/library/research/ccl/health/health.pdf>
- Carreiras, M., Monahan, P. J., Lizarazu, M., Dunabeitia, J. A., & Molinaro, N. (2015). Numbers are not like words: Different pathways for literacy and numeracy. *Neuroimage*, *118*, 79-89. doi:10.1016/j.neuroimage.2015.06.021
- Carroll, J., Epstein, R., Fiscella, K., Gipson, T., Volpe, E., Jean-Pierre, P., . . . Jean-Pierre, P. (2007). Caring for Somali women: implications for clinician-patient communication. *Patient Education & Counseling*, *66*(3), 337-345.
- Chen, J., Ng, E., & Wilkins, R. (1996). The health of Canada's immigrants in 1994-95. *Health Rep*, *7*(4), 33-45, 37-50.
- Chew, L. D., Bradley, K. A., & Boyko, E. J. (2004). Brief questions to identify patients with inadequate health literacy. *Fam Med*, *36*(8), 588-594.
- Choi, S. E., Rush, E., & Henry, S. (2013). Health literacy in Korean immigrants at risk for type 2 diabetes. *Journal of immigrant and minority health / Center for Minority Public Health*, *15*(3), 553-559. doi:http://dx.doi.org/10.1007/s10903-012-9672-9
- Choi, Y. J., & Park, G. H. (2016). Associations among acculturation stress, mental health literacy, and mental health of married immigrant women in Korea. *International Journal of Mental Health Promotion*, *18*(4), 234-246. doi:10.1080/14623730.2016.1213654

- Clark, A., Gilbert, A., Rao, D., & Kerr, L. (2014). 'Excuse me, do any of you ladies speak English?' Perspectives of refugee women living in South Australia: Barriers to accessing primary health care and achieving the quality use of medicines. *Australian Journal of Primary Health*, 20(1), 92-97. doi:<http://dx.doi.org/10.1071/PY11118>
- Coffman, M. J., & Norton, C. K. (2010). Demands of immigration, health literacy, and depression in recent Latino immigrants. *Home Health Care Management & Practice*, 22(2), 116-122. doi:10.1177/1084822309347343
- Coffman, M. J., Norton, C. K., & Beene, L. (2012). Diabetes symptoms, health literacy, and health care use in adult latinos with diabetes risk factors. *Journal of Cultural Diversity*, 19(1), 4-9.
- Coffman, M. J., Shobe, M. A., Dmochowski, J., & Fox, S. D. (2007). Health care access and utilization among Latino immigrants. *Hispanic Health Care International*, 5(2), 73-80.
- Copelj, A., & Kiroopoulos, L. (2011). Knowledge of Depression and Depression Related Stigma in Immigrants from Former Yugoslavia. *Journal of Immigrant & Minority Health*, 13(6), 1013-1018. doi:10.1007/s10903-011-9463-8
- Critical Appraisal Skills Programme. (2018). Critical Appraisal Skills Programme (CASP) Checklist: 10 questions to help you make sense of a qualitative research. Retrieved from <https://casp-uk.net/wp-content/uploads/2018/03/CASP-Qualitative-Checklist-Download.pdf>
- Cyril, S., Nicholson, J. M., Agho, K., Polonsky, M., & Renzaho, A. M. (2017). Barriers and facilitators to childhood obesity prevention among culturally and linguistically diverse (CALD) communities in Victoria, Australia. *Australian & New Zealand Journal of Public Health*, 41(3), 287-293. doi:10.1111/1753-6405.12648
- Davis, T. C., Long, S. W., Jackson, R. H., Mayeaux, E. J., George, R. B., Murphy, P. W., & Crouch, M. A. (1993). Rapid estimate of adult literacy in medicine: a shortened screening instrument. *Fam Med*, 25(6), 391-395.
- Davis, T. C., Michielutte, R., Askov, E. N., Williams, M. V., & Weiss, B. D. (1998). Practical assessment of adult literacy in health care. *Health Educ Behav*, 25(5), 613-624. doi:10.1177/109019819802500508
- Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., . . . Starr, J. M. (2009). Age-associated cognitive decline. *Br Med Bull*, 92, 135-152. doi:10.1093/bmb/ldp033
- Derose, K. P., Escarce, J. J., & Lurie, N. (2007). Immigrants and Health Care: Sources of Vulnerability. *Health Affairs*, 26(5), 1258-1268. doi:10.1377/hlthaff.26.5.1258
- DesMeules, M., Gold, J., McDermott, S., Cao, Z., Payne, J., Lafrance, B., . . . Mao, Y. (2005). Disparities in mortality patterns among Canadian immigrants and refugees, 1980-1998: results of a national cohort study. *J Immigr Health*, 7(4), 221-232.
- Diamond, A. G., & Woo, B. K. P. (2014). Duration of residence and dementia literacy among Chinese Americans. *International Journal of Social Psychiatry*, 60(4), 406-409. doi:10.1177/0020764013491742
- Diviani, N., & Schulz, P. J. (2014). Association between Cancer Literacy and Cancer-Related Behaviour: Evidence from Ticino, Switzerland. *J Public Health Res*, 3(2), 295. doi:10.4081/jphr.2014.295
- Diviani, N., van den Putte, B., Giani, S., & van Weert, J. C. (2015). Low health literacy and evaluation of online health information: a systematic review of the literature. *J Med Internet Res*, 17(5), e112. doi:10.2196/jmir.4018

- Domecq, J. P., Prutsky, G., Elraiayah, T., Wang, Z., Nabhan, M., Shippee, N., . . . Murad, M. H. (2014). Patient engagement in research: a systematic review. *BMC Health Serv Res, 14*, 89. doi:10.1186/1472-6963-14-89
- Dumenci, L., Matsuyama, R. K., Kuhn, L., Perera, R. A., & Siminoff, L. A. (2013). On the Validity of the Rapid Estimate of Adult Literacy in Medicine (REALM) Scale as a Measure of Health Literacy. *Commun Methods Meas, 7*(2), 134-143. doi:10.1080/19312458.2013.789839
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed Self-Assessment: Implications for Health, Education, and the Workplace. *Psychol Sci Public Interest, 5*(3), 69-106. doi:10.1111/j.1529-1006.2004.00018.x
- Eijkemans, M., Mommers, M., Draaisma, J. M., Thijs, C., & Prins, M. H. (2012). Physical activity and asthma: a systematic review and meta-analysis. *PLoS ONE, 7*(12), e50775. doi:10.1371/journal.pone.0050775
- Eklund, J. (2012). *Objective and Subjective Measures of Numeracy: Is asking enough?* (Bachelor's Degree), Uppsala Universitet.
- Ekstrom, R. B., French, J. W., & Harman, H. H. (1979). Cognitive factors: Their identification and replication. *Multivariate Behavioral Research Monographs, 79*(2), 3-84.
- Ellard-Gray, A., Jeffrey, N. K., Choubak, M., & Crann, S. E. (2015). Finding the Hidden Participant - Solutions for Recruiting Hidden, Hard-to-Reach, and Vulnerable Populations. *International Journal of Qualitative Methods, 14*(5).
- Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Med Decis Making, 27*(5), 672-680. doi:10.1177/0272989X07304449
- Filippi, M. K., Faseru, B., Baird, M., Ndikum-Moffor, F., Greiner, K. A., & Daley, C. M. (2014). A pilot study of health priorities of Somalis living in Kansas City: laying the groundwork for CBPR. *Journal of immigrant and minority health / Center for Minority Public Health, 16*(2), 314-320. doi:http://dx.doi.org/10.1007/s10903-012-9732-1
- Floyd, A., & Sakellariou, D. (2017). Healthcare access for refugee women with limited literacy: Layers of disadvantage. *International Journal for Equity in Health, 16* (1) (no pagination)(195). doi:http://dx.doi.org/10.1186/s12939-017-0694-8
- Friedman, D. B., Corwin, S. J., Dominick, G. M., & Rose, I. D. (2009). African American men's understanding and perceptions about prostate cancer: why multiple dimensions of health literacy are important in cancer communication. *J Community Health, 34*(5), 449-460. doi:10.1007/s10900-009-9167-3
- Frisch, A. L., Camerini, L., Diviani, N., & Schulz, P. J. (2012). Defining and measuring health literacy: how can we profit from other literacy domains? *Health Promot Int, 27*(1), 117-126. doi:10.1093/heapro/dar043
- Gagnon, A. J., Dougherty, G., Wahoush, O., Saucier, J. F., Dennis, C. L., Stanger, E., . . . Stewart, D. E. (2013). International migration to Canada: the post-birth health of mothers and infants by immigration class. *Soc Sci Med, 76*(1), 197-207. doi:10.1016/j.socscimed.2012.11.001
- Galdas, P., Darwin, Z., Fell, J., Kidd, L., Bower, P., Blickem, C., . . . Richardson, G. (2015). A systematic review and metaethnography to identify how effective, cost-effective, accessible and acceptable self-management support interventions are for men with long-term conditions (SELF-MAN). *Health Serv Deliv Res, 3*(34). doi:10.3310/hsdr03340

- Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol*, *13*, 117. doi:10.1186/1471-2288-13-117
- Gatobu, S., Arocha, J. F., & Hoffman-Goetz, L. (2014). Numeracy and Health Numeracy Among Chinese and Kenyan Immigrants to Canada: Role of Math Self-Efficacy. *Sage Open*, *1*-10. doi:DOI: 10.1177/2158244014521437
- Gatobu, S., Arocha, J. F., & Hoffman-Goetz, L. (2016). Numeracy, health numeracy, and older immigrants' primary language: An observation-oriented exploration. *Basic and Applied Social Psychology*, *38*(4), 185-199. doi:10.1080/01973533.2016.1197129
- Gele, A. A., Pettersen, K. S., Torheim, L. E., & Kumar, B. (2016). Health literacy: the missing link in improving the health of Somali immigrant women in Oslo. *BMC Public Health*, *16*(1), 1134.
- Gele, A. A., Qureshi, S. A., Kour, P., Kumar, B., & Diaz, E. (2017). Barriers and facilitators to cervical cancer screening among Pakistani and somali immigrant women in Oslo: A qualitative study. *International Journal of Women's Health*, *9*, 487-496. doi:http://dx.doi.org/10.2147/IJWH.S139160
- Geltman, P., Hunter Adams, J., Penrose, K., Cochran, J., Rybin, D., Doros, G., . . . Paasche-Orlow, M. (2014). Health Literacy, Acculturation, and the Use of Preventive Oral Health Care by Somali Refugees Living in Massachusetts. *Journal of Immigrant & Minority Health*, *16*(4), 622-630. doi:10.1007/s10903-013-9846-0
- Geltman, P. L., Adams, J. H., Cochran, J., Doros, G., Rybin, D., Henshaw, M., . . . Paasche-Orlow, M. (2013). The impact of functional health literacy and acculturation on the oral health status of Somali refugees living in Massachusetts. *American journal of public health*, *103*(8), 1516-1523.
- Golbeck, A., Paschal, A., Jones, A., & Hsiao, T. (2011). Correlating reading comprehension and health numeracy among adults with low literacy. *Patient Educ Couns*, *84*(1), 132-134. doi:10.1016/j.pec.2010.05.030
- Golbeck, A. L., Ahlers-Schmidt, C. R., Paschal, A. M., & Dismuke, S. E. (2005). A definition and operational framework for health numeracy. *Am J Prev Med*, *29*(4), 375-376. doi:10.1016/j.amepre.2005.06.012
- Government of Canada. (2017). *2017 Annual Report to Parliament on Immigration*. Retrieved from https://www.canada.ca/en/immigration-refugees-citizenship/corporate/publications-manuals/annual-report-parliament-immigration-2017.html - sec1_2
- Green, J. (2007). Health literacy: Terminology and trends in making and communicating health-related information. *Health Issues*, *92*, 11-14.
- Gregory, J. L. (2015). *Exploring the Expression of Pain by Limited English Proficient Latino Immigrants in Language and Culturally Discordant Health Encounters*. (10027547 Ph.D.), University of California, San Francisco, Ann Arbor.
- Groenenberg, I., Crone, M. R., van Dijk, S., Gebhardt, W. A., Ben Meftah, J., Middelkoop, B. J. C., . . . Assendelft, W. J. J. (2015). 'Check it out!' Decision-making of vulnerable groups about participation in a two-stage cardiometabolic health check: A qualitative study. *Patient Education and Counseling*, *98*(2), 234-244. doi:http://dx.doi.org/10.1016/j.pec.2014.10.010
- Gushulak, B. D., Pottie, K., Hatcher Roberts, J., Torres, S., DesMeules, M., Canadian Collaboration for, I., & Refugee, H. (2011). Migration and health in Canada: health in the global village. *CMAJ*, *183*(12), E952-958. doi:10.1503/cmaj.090287

- Halli, S. S., & Anchan, J. P. (2005). Structural and Behavioural Determinants of Immigrant and Non-Immigrant Health Status: Results from the Canadian Community Health Survey *Journal of International Migration and Integration*, 6(1), 93-122.
- Han, H. R., Huh, B., Kim, M. T., Kim, J., & Nguyen, T. (2014). Development and validation of the assessment of health literacy in breast and cervical cancer screening. *J Health Commun, 19 Suppl 2*, 267-284. doi:10.1080/10810730.2014.936569
- Haun, J. N., Valerio, M. A., McCormack, L. A., Sorensen, K., & Paasche-Orlow, M. K. (2014). Health literacy measurement: an inventory and descriptive summary of 51 instruments. *J Health Commun, 19 Suppl 2*, 302-333. doi:10.1080/10810730.2014.936571
- Heath, G., Cameron, E., Cummins, C., Greenfield, S., Pattison, H., Kelly, D., & Redwood, S. (2012). Paediatric 'care closer to home': stake-holder views and barriers to implementation. *Health Place, 18*(5), 1068-1073. doi:10.1016/j.healthplace.2012.05.003
- Hernandez-Mekonnen, R., Duggan, E. K., Oliveros-Rosen, L., Gerdes, M., Wortham, S., Ludmir, J., & Bennett, I. M. (2016). Health Literacy in Unauthorized Mexican Immigrant Mothers and Risk of Developmental Delay in their Children. *Journal of Immigrant and Minority Health, 18*(5), 1228-1231. doi:http://dx.doi.org/10.1007/s10903-015-0284-z
- Herzog, R., Alvarez-Pasquin, M. J., Diaz, C., Del Barrio, J. L., Estrada, J. M., & Gil, A. (2013). Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review. *BMC Public Health, 13*, 154. doi:10.1186/1471-2458-13-154
- HLS-EU Consortium. (2012). HLS-EU-Q16. The European Health Literacy Survey Questionnaire - short version.
- Holubar, S. D., Hassinger, J. P., Dozois, E. J., Wolff, B. G., Kehoe, M., & Cima, R. R. (2009). Impact of a multimedia e-learning module on colon cancer literacy: a community-based pilot study. *J Surg Res, 156*(2), 305-311. doi:10.1016/j.jss.2009.03.099
- Hurley, C., Panagiotopoulos, G., Tsianikas, M., Newman, L., & Walker, R. (2013). Access and acceptability of community-based services for older Greek migrants in Australia: User and provider perspectives. *Health & Social Care in the Community, 21*(2), 140-149. doi:10.1111/hsc.12000
- Idehen, E. E., Korhonen, T., Castaneda, A., Juntunen, T., Kangasniemi, M., Pietilä, A. M., & Koponen, P. (2017). Factors associated with cervical cancer screening participation among immigrants of Russian, Somali and Kurdish origin: A population-based study in Finland. *BMC Women's Health, 17*(1). doi:10.1186/s12905-017-0375-1
- Igarashi, Y., Horiuchi, S., & Porter, S. E. (2013). Immigrants' experiences of maternity care in Japan. *Journal of community health, 38*(4), 781-790. doi:http://dx.doi.org/10.1007/s10900-013-9679-8
- Institute of Medicine (US) Committee on Health Literacy. (2004). *Health Literacy: A Prescription to End Confusion*. Washington (DC): National Academies Press (US).
- Jacobson, H. E., Hund, L., & Soto Mas, F. (2016). Predictors of English Health Literacy among U.S. Hispanic Immigrants: The importance of language, bilingualism and sociolinguistic environment. *Lit Numer Stud, 24*(1), 43-64. doi:10.5130/lms.v24i1.4900
- Jafri, N. F. (2012). *The role of culture and health literacy in cancer screening practices among young, middle to upper middle-class Pakistani-American women*. (73), ProQuest Information & Learning, US.

- Johnson, T. V., Abbasi, A., Schoenberg, E. D., Kellum, R., Speake, L. D., Spiker, C., . . . Master, V. A. (2014). Numeracy among trainees: are we preparing physicians for evidence-based medicine? *J Surg Educ*, *71*(2), 211-215. doi:10.1016/j.jsurg.2013.07.013
- Jung, M. Y., Holt, C. L., Ng, D., Sim, H. J., Lu, X., Le, D., . . . Lee, S. (2017). The Chinese and Korean American immigrant experience: a mixed-methods examination of facilitators and barriers of colorectal cancer screening. *Ethn Health*, 1-20. doi:10.1080/13557858.2017.1296559
- Kankou, J. M., Bouchaud, O., Lele, N., Bourgeois, D., Spire, B., Carrieri, M. P., & Abgrall, S. (2017). Factors associated with HIV status disclosure in HIV-infected sub-Saharan migrants living in France and successfully treated with antiretroviral therapy: Results from the ANRS-VIHVO Study. *Journal of Immigrant and Minority Health*, *19*(4), 843-850. doi:10.1007/s10903-016-0423-1
- Kennedy, S., Kidd, M., McDonald, J. T., & Biddle, N. (2015). The Healthy Immigrant Effect: Patterns and Evidence from Four Countries. *Journal of International Migration and Integration*, *16*(1), 317-332. doi:doi:10.1007/s12134-014-0340-x
- Khuu, B. P., Lee, H. Y., & Zhou, A. Q. (2018). Health Literacy and Associated Factors Among Hmong American Immigrants: Addressing the Health Disparities. *J Community Health*, *43*(1), 11-18. doi:10.1007/s10900-017-0381-0
- Kim, H. S., & Kim, H. S. (2013). Depression in Non-Korean Women Residing in South Korea Following Marriage to Korean Men. *Archives of Psychiatric Nursing*, *27*(3), 148-155. doi:10.1016/j.apnu.2013.02.005
- Kim, K., Xue, Q.-L., Walton-Moss, B., Nolan, M. T., & Han, H.-R. (2018). Decisional balance and self-efficacy mediate the association among provider advice, health literacy and cervical cancer screening. *European Journal of Oncology Nursing*, *32*, 55-62. doi:10.1016/j.ejon.2017.12.001
- Kirsch, I. S. (2001). The framework used in developing and interpreting the International Adult Literacy Survey (IALS). *European Journal of Psychology of Education*, *16*(3), 335-361.
- Ko, M. L. (2014). *Colorectal cancer screening behaviors among korean americans*. (75), ProQuest Information & Learning, US.
- Koch-Weser, S., Liang, S. L., & Grigg-Saito, D. C. (2006). Self-reported health among Cambodians in Lowell, Massachusetts. *Journal of Health Care for the Poor & Underserved*, *17*(2), 133-145.
- Kreps, G. L., & Sparks, L. (2008). Meeting the health literacy needs of immigrant populations. *Patient Educ Couns*, *71*(3), 328-332. doi:10.1016/j.pec.2008.03.001
- Kutner, M., Greenberg, E., Jin, Y., & Paulsen, C. (2006a). *The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy (NCES 2006-483)*. Retrieved from Washington, DC: <https://nces.ed.gov/pubs2006/2006483.pdf>
- Lee, F. H., Wang, H. H., Yang, Y. M., & Tsai, H. M. (2014). Barriers faced by Vietnamese immigrant women in Taiwan who do not regularly undergo cervical screenings: a qualitative study. *Journal of Advanced Nursing*, *70*(1), 87-96. doi:10.1111/jan.12168
- Lee, H. Y., & Choi, J. K. (2012). Pathway to Health Literacy in Korean American Immigrants: The Mediating Role of English Proficiency. *Journal of Human Behavior in the Social Environment*, *22*(3), 255-269. doi:10.1080/10911359.2012.655568
- Lee, H. Y., Choi, J. K., & Lee, M. H. (2015a). Health literacy in an underserved immigrant population: New implications toward achieving health equity. *Asian American Journal of Psychology*, *6*(1), 97-105. doi:10.1037/a0037425

- Lee, H. Y., Choi, J. K., & Park, J. H. (2014). The primary care physician and cancer literacy: Reducing health disparities in an immigrant population. *Health Education Journal*, 73(4), 435-445. doi:http://dx.doi.org/10.1177/0017896913489290
- Lee, H. Y., Rhee, T. G., Kim, N. K., & Ahluwalia, J. S. (2015b). Health Literacy as a Social Determinant of Health in Asian American Immigrants: Findings from a Population-Based Survey in California. *Journal of General Internal Medicine*, 30(8), 1118-1124. doi:http://dx.doi.org/10.1007/s11606-015-3217-6
- Lee, H. Y., Yang, P. N., Lee do, K., & Ghebre, R. (2015c). Cervical cancer screening behavior among Hmong-American immigrant women. *American Journal of Health Behavior*, 39(3), 301-307. doi:http://dx.doi.org/10.5993/AJHB.39.3.2
- Lee, J. Y., Rozier, R. G., Lee, S. Y. D., D., B., & Ruiz, R. E. (2007). Development of a word recognition instrument to test health literacy in dentistry: The REALD-30--A brief communication. *J Pub Health Dent*, 67(2), 94-98.
- Lee, S. Y., Bender, D. E., Ruiz, R. E., & Cho, Y. I. (2006). Development of an easy-to-use Spanish Health Literacy test. *Health Serv Res*, 41(4 Pt 1), 1392-1412. doi:10.1111/j.1475-6773.2006.00532.x
- Lee, Y. M., Kim, H. M., Lee, M. S., & Lee, H. Y. (1999). Measurement and determinants of mental health states for the urban poor. *Journal of the Korean Neuropsychiatric Association*, 38, 1234-1244.
- Leung, A. Y. M., Bo, A., Hsiao, H. Y., Wang, S. S., & Chi, I. (2014). Health literacy issues in the care of Chinese American immigrants with diabetes: A qualitative study. *BMJ Open*, 4(11) (no pagination)(e005294). doi:http://dx.doi.org/10.1136/bmjopen-2014-005294
- Levinthal, B. R., Morrow, D. G., Tu, W., Wu, J., & Murray, M. D. (2008). Cognition and health literacy in patients with hypertension. *J Gen Intern Med*, 23(8), 1172-1176. doi:10.1007/s11606-008-0612-2
- Lipkus, I. M., & Peters, E. (2009). Understanding the role of numeracy in health: proposed theoretical framework and practical insights. *Health Educ Behav*, 36(6), 1065-1081. doi:10.1177/1090198109341533
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Med Decis Mak*, 21(37-44).
- Lo, S., Sharif, I., & Ozuah, P. O. (2006). Health literacy among English-speaking parents in a poor urban setting. *Journal of Health Care for the Poor & Underserved*, 17(3), 504-511.
- Maneze, D., Everett, B., Astorga, C., Yogendran, D., & Salamonson, Y. (2016). The Influence of Health Literacy and Depression on Diabetes Self-Management: A Cross-Sectional Study. *Journal of Diabetes Research*, 2016 (no pagination)(3458969). doi:http://dx.doi.org/10.1155/2016/3458969
- Mantwill, S., & Schulz, P. J. (2017). Does acculturation narrow the health literacy gap between immigrants and non-immigrants-An explorative study. *Patient Education & Counseling*, 100(4), 760-767. doi:10.1016/j.pec.2016.10.021
- Marshall, E. G., Wong, S. T., Haggerty, J. L., & Levesque, J. F. (2010). Perceptions of unmet healthcare needs: what do Punjabi and Chinese-speaking immigrants think? A qualitative study. *BMC health services research*, 10, 46.
- Masood, M., Thaliath, E. T., Bower, E. J., & Newton, J. T. (2011). An appraisal of the quality of published qualitative dental research. *Community Dent Oral Epidemiol*, 39(3), 193-203. doi:10.1111/j.1600-0528.2010.00584.x

- Mata, D. A., Ramos, M. A., Bansal, N., Khan, R., Guille, C., Di Angelantonio, E., & Sen, S. (2015). Prevalence of Depression and Depressive Symptoms Among Resident Physicians: A Systematic Review and Meta-analysis. *JAMA*, *314*(22), 2373-2383. doi:10.1001/jama.2015.15845
- McDonald, J. T., Farnworth, M., & Liu, Z. (2017). Cancer and the healthy immigrant effect: a statistical analysis of cancer diagnosis using a linked Census-cancer registry administrative database. *BMC Public Health*, *17*(1), 296. doi:10.1186/s12889-017-4190-2
- McDonald, J. T., & Kennedy, S. (2004). Insights into the 'healthy immigrant effect': health status and health service use of immigrants to Canada. *Soc Sci Med*, *59*(8), 1613-1627. doi:10.1016/j.socscimed.2004.02.004
- McDonald, J. T., & Kennedy, S. (2005). Is migration to Canada associated with unhealthy weight gain? Overweight and obesity among Canada's immigrants. *Social Science & Medicine*, *61*(12), 2469-2481. doi:https://doi.org/10.1016/j.socscimed.2005.05.004
- McEwing, R., McLachlan, A., Lund, M., & Carrucan-Wood, L. (2017). The Impact of Health Literacy on the Health Outcomes of the Heart Failure Population. *Heart, Lung and Circulation*, *26*(suppl 1), S8. doi:https://doi.org/10.1016/j.hlc.2017.05.021
- McMichael, C., & Gifford, S. (2009). "It is good to know now...before it's too late": Promoting sexual health literacy amongst resettled young people with refugee backgrounds. *Sexuality & Culture: An Interdisciplinary Quarterly*, *13*(4), 218-236. doi:10.1007/s12119-009-9055-0
- McPartland, T. S., Weaver, B. A., Lee, S. K., & Koutsky, L. A. (2005). Men's perceptions and knowledge of human papillomavirus (HPV) infection and cervical cancer. *J Am Coll Health*, *53*(5), 225-230. doi:10.3200/JACH.53.5.225-230
- Miron-Shatz, T., Yaniv Hanoach, Y., Doniger, G. M., Omer, Z. B., & Ozanne, E. M. (2014). Subjective but not objective numeracy influences willingness to pay for BRCA1/2 genetic testing. *Judgment and Decision Making*, *9*(2), 152-158.
- Modesti, P. A., Reboldi, G., Cappuccio, F. P., Agyemang, C., Remuzzi, G., Rapi, S., . . . Settings, E. S. H. W. G. o. C. R. i. L. R. (2016). Panethnic Differences in Blood Pressure in Europe: A Systematic Review and Meta-Analysis. *PLoS ONE*, *11*(1), e0147601. doi:10.1371/journal.pone.0147601
- Morris, N. S., MacLean, C. D., Chew, L. D., & Littenberg, B. (2006). The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability. *BMC Fam Pract*, *7*, 21. doi:10.1186/1471-2296-7-21
- Murray, L., Elmer, S., & Elkhair, J. (2018). Perceived Barriers to Managing Medications and Solutions to Barriers Suggested by Bhutanese Former Refugees and Service Providers. *J Transcult Nurs*, 1043659618768702. doi:10.1177/1043659618768702
- Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: theory and practice. *Ann Behav Med*, *35*(3), 261-274. doi:10.1007/s12160-008-9037-8
- Nelson, W. L., Moser, R. P., & Han, P. K. (2013). Exploring objective and subjective numeracy at a population level: findings from the 2007 Health Information National Trends Survey (HINTS). *J Health Commun*, *18*(2), 192-205. doi:10.1080/10810730.2012.688450
- Newbold, K. B., & Willinsky, J. (2009). Providing family planning and reproductive healthcare to Canadian immigrants: perceptions of healthcare providers. *Cult Health Sex*, *11*(4), 369-382. doi:10.1080/13691050802710642

- Ng, E., & Omariba, D. (2014). Immigration, generational status and health literacy in Canada. *Health Education Journal*, 73(6), 668-682. doi:DOI: 10.1177/0017896913511809
- Nguyen, T. H., Paasche-Orlow, M. K., Kim, M. T., Han, H. R., & Chan, K. S. (2015). Modern Measurement Approaches to Health Literacy Scale Development and Refinement: Overview, Current Uses, and Next Steps. *J Health Commun*, 20 Suppl 2, 112-115. doi:10.1080/10810730.2015.1073408
- Njeru, J. W., Hagi-Salaad, M. F., Haji, H., Cha, S. S., & Wieland, M. L. (2016). Diabetes health literacy among Somali patients with diabetes mellitus in a US primary care setting. *Journal of Racial and Ethnic Health Disparities*, 3(2), 210-216. doi:10.1007/s40615-015-0129-4
- Nutbeam, D. (1998). Health Promotion Glossary. *Health Promotion International*, 13, 349-364.
- Nutbeam, D. (2000). Health literacy as a public goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot Int*, 15(3), 259-267.
- Nutbeam, D. (2008). The evolving concept of health literacy. *Soc Sci Med*, 67(12), 2072-2078. doi:10.1016/j.socscimed.2008.09.050
- OECD. (2012). *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD survey of Adult Skills*. Retrieved from http://www.oecd.org/skills/piaac/PIAAC Framework 2012-- Revised 28oct2013_ebook.pdf
- Oliver, B. D. (2015). *Bhutanese refugees' health behaviors and perceptions: A qualitative approach to understanding Bhutanese refugees' post-resettlement experiences with health care services in middle Tennessee*. (76), ProQuest Information & Learning, US.
- Omariba, D. W., & Ng, E. (2011). Immigration, generation and self-rated health in Canada: on the role of health literacy. *Canadian Journal of Public Health*, 102(4), 281-285.
- Omariba, D. W. R., & Ng, E. (2015). Health literacy and disability: Differences between generations of Canadian immigrants. *International Journal of Public Health*, 60(3), 389-397. doi:10.1007/s00038-014-0640-0
- Parker, R. M., Baker, D. W., Williams, M. V., & Nurss, J. R. (1995). The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. *J Gen Intern Med*, 10(10), 537-541.
- Pelikan, J., & Ganahl, K. (2017). The European Health Literacy Study: concept, instrument and selected results [German]. In D. Schaeffer & J. Pelikan (Eds.), *Health Literacy: Forschungsstand und Perspektiven* (pp. 93-125). Bern: Hogrefe.
- Pleasant, A., McKinney, J., & Rikard, R. V. (2011). Health literacy measurement: a proposed research agenda. *J Health Commun*, 16 Suppl 3, 11-21. doi:10.1080/10810730.2011.604392
- Prins, E., & Monnat, S. (2015). Examining associations between self-rated health and proficiency in literacy and numeracy among immigrants and U.S.-born adults: Evidence from the Program for the International Assessment of Adult Competencies (PIAAC). *PLoS ONE*, 10 (7) (no pagination)(e0130257). doi:http://dx.doi.org/10.1371/journal.pone.0130257
- Reichel, D., & Morales, L. (2017). Surveying immigrants without sampling frames - evaluating the success of alternative field methods. *Comp Migr Stud*, 5(1), 1. doi:10.1186/s40878-016-0044-9

- Renzaho, A. M. N., Green, J., Smith, B. J., & Polonsky, M. (2017). Exploring Factors Influencing Childhood Obesity Prevention Among Migrant Communities in Victoria, Australia: A Qualitative Study. *J Immigr Minor Health*. doi:10.1007/s10903-017-0620-6
- Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychol Bull*, 135(6), 943-973. doi:10.1037/a0017327
- Ritchie, J., & Lewis, J. (2003). *Qualitative research practice: A guide for social science students and researchers*. London: Sage.
- Rootman, I., & Gordon-El-Bihbety, D. (2008). *A vision for a health literate Canada: Report of the Expert Panel on Health Literacy*. Retrieved from Ottawa: [https://swselfmanagement.ca/uploads/ResourceDocuments/CPHA \(2008\) A Vision for a Health Literate Canada.pdf](https://swselfmanagement.ca/uploads/ResourceDocuments/CPHA%20(2008)%20A%20Vision%20for%20a%20Health%20Literate%20Canada.pdf)
- Rothman, R. L., Malone, R., Bryant, B., Wolfe, C., Padgett, P., DeWalt, D. A., . . . Pignone, M. (2005). The Spoken Knowledge in Low Literacy in Diabetes scale: a diabetes knowledge scale for vulnerable patients. *Diabetes Educ*, 31(2), 215-224. doi:10.1177/0145721705275002
- Rudd, R., Kirsch, I., & Yamamoto, K. (2004). *Literacy and Health in America. Policy Information Report*. Retrieved from <https://files.eric.ed.gov/fulltext/ED486416.pdf>
- Sabbahi, D. A., Lawrence, H. P., Limeback, H., & Rootman, I. (2009). Development and evaluation of an oral health literacy instrument for adults. *Community Dent Oral Epidemiol*, 37(5), 451-462. doi:10.1111/j.1600-0528.2009.00490.x
- Sarkar, U., Schillinger, D., Lopez, A., & Sudore, R. (2011). Validation of self-reported health literacy questions among diverse English and Spanish-speaking populations. *J Gen Intern Med*, 26(3), 265-271. doi:10.1007/s11606-010-1552-1
- Schoenmakers, D., Lamkaddem, M., & Suurmond, J. (2017). The role of the social network in access to psychosocial services for migrant elderly—a qualitative study. *International Journal of Environmental Research and Public Health*, 14(10). doi:10.3390/ijerph14101215
- Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Ann Intern Med*, 127(11), 966-972.
- Shaghghi, A., Bhopal, R. S., & Sheikh, A. (2011). Approaches to Recruiting 'Hard-To-Reach' Populations into Re-search: A Review of the Literature. *Health Promot Perspect*, 1(2), 86-94. doi:10.5681/hpp.2011.009
- Shamseer, L., Moher, D., Clarke, M., Gherzi, D., Liberati, A., Petticrew, M., . . . Group, P.-P. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*, 350, g7647. doi:10.1136/bmj.g7647
- Sheard, L., Prout, H., Dowding, D., Noble, S., Watt, I., Maraveyas, A., & Johnson, M. (2013). Barriers to the diagnosis and treatment of venous thromboembolism in advanced cancer patients: a qualitative study. *Palliat Med*, 27(4), 339-348. doi:10.1177/0269216312461678
- Shedlin, M. G., & Shulman, L. (2004). Qualitative needs assessment of HIV services among Dominican, Mexican and Central American immigrant populations living in the New York City area. *AIDS Care*, 16(4), 434-445. doi:10.1080/09540120410001683376
- Simonds, S. K. (1974). Health education as social policy. *Health Education Monograph*, 2, 1-25.

- Smaland Goth, U. G., & Berg, J. E. (2011). Migrant participation in Norwegian health care. A qualitative study using key informants. *European Journal of General Practice, 17*(1), 28-33. doi:10.3109/13814788.2010.525632
- Smith, J., & Firth, J. (2011). Qualitative data analysis: the framework approach. *Nurse Res, 18*(2), 52-62. doi:10.7748/nr2011.01.18.2.52.c8284
- Smith, J. L., & Haggerty, J. (2003). Literacy in primary care populations: is it a problem? *Canadian Journal of Public Health, 94*(6), 408-412.
- Smith, P. C., Brice, J. H., & Lee, J. (2012). The relationship between functional health literacy and adherence to emergency department discharge instructions among Spanish-speaking patients. *Journal of the National Medical Association, 104*(11-12), 521-527.
- Sorensen, K., Pelikan, J. M., Rothlin, F., Ganahl, K., Slonska, Z., Doyle, G., . . . Consortium, H.-E. (2015). Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). *Eur J Public Health, 25*(6), 1053-1058. doi:10.1093/eurpub/ckv043
- Sorensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., . . . Consortium Health Literacy Project, E. (2012). Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health, 12*, 80. doi:10.1186/1471-2458-12-80
- Squiers, L., Peinado, S., Berkman, N., Boudewyns, V., & McCormack, L. (2012). The health literacy skills framework. *J Health Commun, 17 Suppl 3*, 30-54. doi:10.1080/10810730.2012.713442
- Sriphanlop, P., Jandorf, L., Kairouz, C., Thelemaque, L., Shankar, H., & Perumalswami, P. (2014). Factors related to hepatitis B screening among Africans in New York City. *American Journal of Health Behavior, 38*(5), 745-754. doi:10.5993/AJHB.38.5.12
- Statistics Canada. (2005). *Building on our Competencies: Canadian Results of the International Adult Literacy and Skills Survey Catalogue no. 89-617-X1E*. Retrieved from Ottawa: <http://www.statcan.gc.ca/pub/89-617-x/89-617-x2005001-eng.pdf>
- Statistics Canada. (2008). Numeracy. Retrieved from <http://www.statcan.gc.ca/pub/81-004-x/def/4068737-eng.htm>
- Statistics Canada. (2017). *Immigration and ethnocultural diversity: Key results from the 2016 Census*. Retrieved from <http://www.statcan.gc.ca/daily-quotidien/171025/dq171025b-eng.pdf>
- Stein, K., Zhao, L., Crammer, C., & Gansler, T. (2007). Prevalence and sociodemographic correlates of beliefs regarding cancer risks. *Cancer, 110*(5), 1139-1148. doi:10.1002/cncr.22880
- Sterne, J. A., Hernan, M. A., Reeves, B. C., Savovic, J., Berkman, N. D., Viswanathan, M., . . . Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ, 355*, i4919. doi:10.1136/bmj.i4919
- Stonbraker, S., Schnall, R., & Larson, E. (2015). Tools to measure health literacy among Spanish speakers: An integrative review of the literature. *Patient Educ Couns.* doi:10.1016/j.pec.2015.07.012
- Sulaiman-Hill, C. M., & Thompson, S. C. (2011). Sampling challenges in a study examining refugee resettlement. *BMC Int Health Hum Rights, 11*, 2. doi:10.1186/1472-698X-11-2
- Taiwo, C. C. (2013). *Exploring health literacy among a selected Nigerian population: An ethnographic study*. (10124889 Ed.D.), University of Phoenix, Ann Arbor.
- Thabit, H., Shah, S., Nash, M., Brema, I., Nolan, J. J., & Martin, G. (2009). Globalization, immigration and diabetes self-management: An empirical study amongst immigrants with

- type 2 diabetes mellitus in Ireland. *Qjm*, 102(10), 713-720.
doi:<http://dx.doi.org/10.1093/qjmed/hcp113>
- The Cochrane Collaboration. (2011). Chapter 8: Assessing risk of bias in included studies. In Higgins P., Altman D., & Sterne J. (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions*.
- Thomson, M. D., & Hoffman-Goetz, L. (2011). Cancer information comprehension by English-as-a-second-language immigrant women. *Journal of health communication*, 16(1), 17-33.
- Thomson, M. D., & Hoffman-Goetz, L. (2012). Application of the health literacy framework to diet-related cancer prevention conversations of older immigrant women to Canada. *Health Promotion International*, 27(1), 33-44.
- Todd, L., Harvey, E., & Hoffman-Goetz, L. (2011a). Predicting breast and colon cancer screening among English-as-a-second-language older Chinese immigrant women to Canada. *Journal of Cancer Education*, 26(1), 161-169. doi:10.1007/s13187-010-0141-7
- Todd, L., & Hoffman-Goetz, L. (2011b). Predicting health literacy among English-as-a-second-Language older Chinese immigrant women to Canada: comprehension of colon cancer prevention information. *Journal of Cancer Education*, 26(2), 326-332.
doi:10.1007/s13187-010-0162-2
- Todd, L., & Hoffman-Goetz, L. (2011c). A qualitative study of cancer information seeking among English-as-a-second-Language older Chinese immigrant women to Canada: sources, barriers, and strategies. *Journal of Cancer Education*, 26(2), 333-340.
doi:10.1007/s13187-010-0174-y
- Tsoh, J. Y., Sentell, T., Gildengorin, G., Le, G. M., Chan, E., Fung, L.-C., . . . Nguyen, T. T. (2016). Healthcare communication barriers and self-rated health in older Chinese American immigrants. *Journal of Community Health: The Publication for Health Promotion and Disease Prevention*, 41(4), 741-752. doi:10.1007/s10900-015-0148-4
- United Nations Department of Economic and Social Affairs. (2017). *International Migration Report 2017 Highlights (ST/ESA/SER.A/404)*. Retrieved from http://www.un.org/en/development/desa/population/migration/publications/migrationreport/docs/MigrationReport2017_Highlights.pdf
- United States National Institute on Drug Abuse. (1990). *The Collection and Interpretation of Data from Hidden Populations*. Retrieved from Washington, DC: <https://archives.drugabuse.gov/sites/default/files/monograph98.pdf>
- van der Heide, I., Rademakers, J., Schipper, M., Droomers, M., Sorensen, K., & Uiters, E. (2013). Health literacy of Dutch adults: a cross sectional survey. *BMC Public Health*, 13, 179.
doi:10.1186/1471-2458-13-179
- Vang, Z., Sigouin, J., Flenon, A., & Gagnon, A. (2015). The Healthy Immigrant Effect in Canada: A Systematic Review. *Population Change and Lifecourse Strategic Knowledge Cluster Discussion Paper Series*, 3(1).
- von Wagner, C., Semmler, C., Good, A., & Wardle, J. (2009a). Health literacy and self-efficacy for participating in colorectal cancer screening: The role of information processing. *Patient Education and Counseling*, 75(3), 352-357.
doi:<https://doi.org/10.1016/j.pec.2009.03.015>
- von Wagner, C., Steptoe, A., Wolf, M. S., & Wardle, J. (2009b). Health literacy and health actions: a review and a framework from health psychology. *Health Educ Behav*, 36(5), 860-877. doi:10.1177/1090198108322819

- Wangdahl, J., Lytsy, P., Martensson, L., & Westerling, R. (2014). Health literacy among refugees in Sweden - a cross-sectional study. *BMC Public Health*, *14*, 1030. doi:<http://dx.doi.org/10.1186/1471-2458-14-1030>
- Wangdahl, J., Lytsy, P., Martensson, L., & Westerling, R. (2015). Health literacy and refugees' experiences of the health examination for asylum seekers - a Swedish cross-sectional study. *BMC Public Health*, *15*, 1162. doi:<http://dx.doi.org/10.1186/s12889-015-2513-8>
- Wangdahl, J., Lytsy, P., Martensson, L., & Westerling, R. (2018). Poor health and refraining from seeking healthcare are associated with comprehensive health literacy among refugees: a Swedish cross-sectional study. *Int J Public Health*. doi:10.1007/s00038-017-1074-2
- Watts, M. C. N. C., McMichael, C., & Liamputtong, P. (2014). Factors influencing contraception awareness and use: The experiences of young African Australian mothers. *Journal of Refugee Studies*, *28*(3), 368-387. doi:10.1093/jrs/feu040
- Weiss, B. D., Mays, M. Z., Martz, W., Castro, K. M., DeWalt, D. A., Pignone, M. P., . . . Hale, F. A. (2005). Quick assessment of literacy in primary care: the newest vital sign. *Ann Fam Med*, *3*(6), 514-522. doi:10.1370/afm.405
- Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, & Tugwell P. (2011). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Retrieved from http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
- Wister, A. V., Malloy-Weir, L. J., Rootman, I., & Desjardins, R. (2010). Lifelong educational practices and resources in enabling health literacy among older adults. *Journal of Aging & Health*, *22*(6), 827-854. doi:10.1177/0898264310373502
- Woo, B. K. P. (2013). Knowledge of dementia among Chinese American immigrants. *Asian Journal of Psychiatry*. doi:<http://dx.doi.org/10.1016/j.ajp.2013.03.010>.
- Woudstra, A. J., Dekker, E., Essink-Bot, M. L., & Suurmond, J. (2016). Knowledge, attitudes and beliefs regarding colorectal cancer screening among ethnic minority groups in the Netherlands—A qualitative study. *Health Expectations: An International Journal of Public Participation in Health Care & Health Policy*, *19*(6), 1312-1323. doi:10.1111/hex.12428
- Yin, H. S., Johnson, M., Mendelsohn, A. L., Abrams, M. A., Sanders, L. M., & Dreyer, B. P. (2009). The health literacy of parents in the United States: a nationally representative study. *Pediatrics*, *124*, S289-298. doi:10.1542/peds.2009-1162E
- Yin, H. S., Sanders, L. M., Rothman, R. L., Mendelsohn, A. L., Dreyer, B. P., White, R. O., . . . Perrin, E. M. (2012). Assessment of health literacy and numeracy among Spanish-Speaking parents of young children: validation of the Spanish Parental Health Literacy Activities Test (PHLAT Spanish). *Acad Pediatr*, *12*(1), 68-74. doi:10.1016/j.acap.2011.08.008
- Yunusa Vakkai, R. J. (2016). *Health status, health care access, literacy & numeracy among members of immigrant communities: The relationship of perceptions, awareness & concerns regarding the health care act.* (77), ProQuest Information & Learning, US.
- Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the Subjective Numeracy Scale: effects of low numeracy on comprehension of risk communications and utility elicitation. *Med Decis Making*, *27*(5), 663-671. doi:10.1177/0272989X07303824

Appendices

Appendix A. Combination of search terms and operators in PubMed

Search terms + Operators

1 (((("emigrants and immigrants"[Mesh]) or ("undocumented immigrants"[Mesh]) or ("emigration and immigration"[Mesh]) or ("transients and migrants"[Mesh]) or ("refugees"[Mesh]))) or (immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or foreign-born or "foreign born" or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens")))) and (("health literacy"[Mesh] or literacy) Filters: Publication date to 2018/05/10; English

2 (((("emigrants and immigrants"[Mesh]) or ("undocumented immigrants"[Mesh]) or ("emigration and immigration"[Mesh]) or ("transients and migrants"[Mesh]) or ("refugees"[Mesh]))) or (immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or foreign-born or "foreign born" or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens")))) and numeracy Filters: Publication date to 2018/05/10; English

Appendix B. Combination of search terms and operators in Embase

Search terms + Operators	
1	exp "migrant"/ or exp "migration"/ or exp "refugee"/ or "undocumented immigrant"/
2	(immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or foreign-born or "foreign born" or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens").mp.
3	1 or 2
4	exp "health literacy"/
5	literacy.mp.
6	4 or 5
7	numeracy.mp.
8	3 and 6
9	3 and 7
10	limit 8 to (english language and yr="1974 -Current")
11	limit 9 to (english language and yr="1974 -Current")

exp: Explode; .mp: multi-purpose (.mp) set of fields, usually including Title, Original Title, Abstract, and Subject Heading.

Appendix C. Combination of search terms and operators in CINAHL

Search terms + Operators

- 1 (MH "Immigrants+") or (MH "Emigration and Immigration") or (MH "Transients and Migrants") or (MH "Relocation") or (MH "Refugees")

(TX immigrant) or (TX immigrants) or (TX immigration) or (TX migrant) or (TX migrants) or (TX migration) or (TX emigrant) or (TX emigrants) or (TX emigration) or (TX newcomer) or (TX newcomers) or (TX foreign born) or (TX foreign-born) or (TX refugee) or (TX refugees) or (TX asylum seeker) or (TX asylum seekers) or (TX stateless person) or (TX stateless persons) or (TX illegal alien) or (TX illegal aliens) or (TX undocumented alien) or (TX undocumented aliens) or (TX irregular alien) or (TX irregular aliens) or (TX clandestine alien) or (TX clandestine aliens) or (TX unauthorized alien) or (TX unauthorized aliens)
- 3 1 or 2
- 4 MH "Health Literacy"
- 5 TX literacy
- 6 4 or 5
- 7 TX numeracy
- 8 3 and 6
- 9 3 and 7
- 10 8 Limiters - Published Date: 19810101-20180510; English Language
- 11 9 Limiters - Published Date: 19810101-20180510; English Language

MH: Exact subject heading; +: Explode; TX: All text

Appendix D. Combination of search terms and operators in PsycINFO

Search terms + Operators

1 (((Any Field: (literacy)))) or (((Index Terms: (health literacy)))) and (((Any Field:(immigrant*)) or (Any Field:(emigrant*)) or (Any Field:(migrant*)) or (Any Field:(immigration)) or (Any Field:(emigration)) or (Any Field: (newcomer*)) or (Any Field: (foreign born)) or (Any Field: (foreign-born)) or (Any Field:(migration)) or (Any Field:(refugee*)) or (Any Field:(asylum seeker*)) or (Any Field:(stateless person*)) or (Any Field:(illegal alien*)) or (Any Field:(undocumented alien*)) or (Any Field:(irregular alien*)) or (Any Field:(clandestine alien*)) or (Any Field:(unauthorized alien*))))) or (((Index Terms:(immigration))) or ((Index Terms:(human migration))) or ((Index Terms:(refugees)))) and Language: English

2 (((Any Field: (numeracy)))) or (((Index Terms: (health literacy)))) and (((Any Field:(immigrant*)) or (Any Field:(emigrant*)) or (Any Field:(migrant*)) or (Any Field:(immigration)) or (Any Field:(emigration)) or (Any Field: (newcomer*)) or (Any Field: (foreign born)) or (Any Field: (foreign-born)) or (Any Field:(migration)) or (Any Field:(refugee*)) or (Any Field:(asylum seeker*)) or (Any Field:(stateless person*)) or (Any Field:(illegal alien*)) or (Any Field:(undocumented alien*)) or (Any Field:(irregular alien*)) or (Any Field:(clandestine alien*)) or (Any Field:(unauthorized alien*))))) or (((Index Terms:(immigration))) or ((Index Terms:(human migration))) or ((Index Terms:(refugees)))) and Language: English

Appendix E. Combination of search terms and operators in SCOPUS

Search terms + Operators

1 TITLE-ABS-KEY (immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or "foreign born" or foreign-born or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens") and TITLE-ABS-KEY (literacy) and (LIMIT-TO (LANGUAGE, "English"))

2 TITLE-ABS-KEY (immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or "foreign born" or foreign-born or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens") and TITLE-ABS-KEY (numeracy) and (LIMIT-TO (LANGUAGE, "English"))

TITLE-ABS-KEY: title-abstract-keywords

Appendix F. Combination of search terms and operators in The ProQuest Dissertations & Theses database

Search terms + Operators

1 (SU(immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or foreign-born or "foreign born" or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens") or AB(immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or foreign-born or "foreign born" or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens")) and (SU(literacy) or AB(literacy)) and la.exact("ENG")

2 (SU(immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or foreign-born or "foreign born" or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens") or AB(immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or foreign-born or "foreign born" or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens")) and (SU(numeracy) or AB(numeracy)) and la.exact("ENG")

SU: subject; AB: abstract

Appendix G. Combination of search terms and operators in the CPCI-S and CPCI-SSH

Search terms + Operators

1 (TS=(immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or foreign-born or "foreign born" or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens") and TS=(literacy)) and LANGUAGE: (English)

2 (TS=(immigrant or immigrants or immigration or migrant or migrants or migration or emigrant or emigrants or emigration or newcomer or newcomers or foreign-born or "foreign born" or refugee or refugees or "asylum seeker" or "asylum seekers" or "stateless person" or "stateless persons" or "illegal alien" or "illegal aliens" or "undocumented alien" or "undocumented aliens" or "irregular alien" or "irregular aliens" or "clandestine alien" or "clandestine aliens" or "unauthorized alien" or "unauthorized aliens") and TS=(numeracy)) and LANGUAGE: (English)

TS: TOPIC – search the title, abstract, author keywords, and keywords plus within a record

Appendix H. Data items extracted from quantitative studies

Included studies (first author et al., year)	
Study Aim(s)	
Research Focus	Health Literacy or Health Numeracy
	Immigrants or Refugees
Literature Type	
Study Design	
Research Sample	Eligibility Criteria (if not provided, then description of sample)
	Sampling Methods
	Total Sample Size
	Number or Percentage of Immigrants or Refugees in the Total Sample
	Country (Region) of Origin of Immigrants or Refugees
	Destination Country/region
	Gender Composition of Immigrants or Refugees
	Age of Immigrants or Refugees
Assessment Instrument	For Health Literacy
	For Health Numeracy
Antecedents	Antecedent(s) Showing Statistical Significance
	Narrative Summary (quotes)
	Statistical Analysis
	Indicator of Association and its Variation
	Strength of Association
	P value
	Adjusted Covariates
	Antecedents Showing No Statistical Significance
Consequences	Consequences Showing Statistical Significance
	Narrative Summary (quotes)
	Measures for Consequences
	Statistical Analysis
	Comparison
	Indicator of Association and its Variation
	Strength of Association
	P value
	Adjusted Covariates
Consequences Showing No Statistical Significance	

Appendix I. Data items extracted from qualitative studies

Included studies (first author et al., year)	
Study Aim(s)	
Research Focus	Health Literacy or Health Numeracy
	Immigrants or Refugees
Literature Type	
Research Sample	Eligibility Criteria (if not provided, then description of sample)
	Sampling Methods
	Total Sample Size
	Number or Percentage of Immigrants or Refugees in the Total Sample
	Country (Region) of Origin of Immigrants or Refugees
	Destination Country
	Gender Composition of Immigrants or Refugees
	Age of Immigrants or Refugees
Methods	Assessment Instrument for Health Literacy
	Assessment Instrument for Health Numeracy
	Data Collection Methods
	Data Analysis Methods
Antecedents	Antecedent(s)
	Narrative summary
Consequences	Consequence(s)
	Narrative Summary

Appendix J. The Newcastle-Ottawa Quality Assessment Scale adapted for cross-sectional studies applied in the thesis research

Selection: (*Maximum 3 stars*)

1) Representativeness of the sample:

- a) Truly representative of the average in the target population. * (all subjects or random sampling)
- b) Somewhat representative of the average in the target population. * (non-random sampling such as convenience sampling, snowball sampling, purposive sampling, consecutive sampling, volunteer sampling, quota sampling, and judgemental sampling.)
- c) Selected group of participants (The researchers further selected the participants based on one or more features other than those defined in their eligibility criteria. For instance, in a study targeting immigrant women, the investigators selected immigrant women whose profession was teacher.)
- d) No description of the sampling strategy.

2) Sample size:

- a) Justified and satisfactory. *
- b) Not justified.

3) Non-respondents:

- a) Comparability between respondents and non-respondents characteristics is established, and the response rate is satisfactory. *
- b) The response rate is unsatisfactory, or the comparability between respondents and non-respondents is unsatisfactory.

c) No description of the response rate or the characteristics of the responders or the non-responders.

Comparability: (Maximum 2 stars)

1) Comparability of subjects in different outcome groups

a) Results adjusted for relevant confounders. **

b) Results not adjusted for any relevant confounders, or information not provided.

Exposures and Outcomes: (Maximum 5 stars)

1) Ascertainment of the exposure:

a) Validated measurement tool, secure record (e.g., surgical records), or exposure factors do not need assessment tools (e.g., age, educational level). **

b) Non-validated measurement tool, but the tool is available or described, or self report.*

c) No description.

2) Assessment of the outcome:

a) Independent blind assessment. **

b) Record linkage or assessment using objective methods (e.g., glycemic control measured by HbA1c). **

c) Self report. *

d) No description.

3) Statistical test:

a) The statistical test used to analyze the data is clearly described and appropriate, and the measurement of the association including p value is presented. *

b) The statistical test is not appropriate, not described or incomplete.

Appendix K: Risk of bias assessment for the included cross-sectional quantitative studies

Included Studies	Selection			Comparability	Exposure and Outcome		
	Sample representative?	Sample size justified?	Non-respondents	Comparability of subjects in different outcome groups	Ascertainment of the exposure	Assessment of the outcome	Statistical test
Becerra et al., 2017	b) *	a) *	c)	a) **	a) **	c) *	a) *
Bekker et al., 2004	c)	b)	c)	a) **	b) *	c) *	a) *
Beltran et al., 2016	b) *	a) *	c)	a) **	b) *	b) **	a) *
Calvasina et al., 2016	b) *	a) *	c)	a) **	a) **	c) *	a) *
Calvo, 2016	b) *	a) *	b)	a) **	a) **	c) *	a) *
S. E. Choi et al., 2013	b) *	b)	c)	b)	a) **	b) **	a) *
Y. J. Choi et al., 2016	b) *	a) *	a) *	b)	a) **	c) *	a) *
Coffman et al., 2007	b) *	b)	c)	a) **	a) **	c) *	a) *
Coffman et al., 2010	b) *	b)	c)	a) **	a) **	c) *	b)
Coffman et al., 2012	b) *	b)	c)	a) **	a) **	c) *	a) *

Included Studies	Selection			Comparability	Exposure and Outcome		
	Sample representative?	Sample size justified?	Non-respondents	Comparability of subjects in different outcome groups	Ascertainment of the exposure	Assessment of the outcome	Statistical test
Diamond et al., 2014	d)	b)	c)	b)	a) **	b) **	a) *
Gatobu et al., 2014	b) *	b)	c)	a) **	a) **	b) **	a) *
Gatobu et al., 2016	b) *	b)	c)	b)	a) **	b) **	a) *
Gele et al., 2016	b) *	b)	c)	a) **	c)	c) *	a) *
Geltman et al., 2013	b) *	b)	c)	a) **	a) **	b) **	a) *
Geltman et al., 2014	b) *	b)	c)	a) **	a) **	c) *	a) *
Hernandez-Mekonnen et al., 2016	d)	b)	c)	a) **	a) **	b) **	b)
Idehen et al., 2017	a) *	b)	c)	a) **	c)	c) *	a) *
Igarashi et al., 2013	b) *	b)	c)	b)	b) *	b) **	a) *
Jacobson et al., 2016	c)	b)	c)	a) **	a) **	b) **	a) *

Included Studies	Selection			Comparability	Exposure and Outcome		
	Sample representative?	Sample size justified?	Non-respondents	Comparability of subjects in different outcome groups	Ascertainment of the exposure	Assessment of the outcome	Statistical test
Kankou et al., 2017	d)	b)	a) *	a) **	b) *	c) *	b)
Khuu et al., 2018	b) *	b)	c)	a) **	a) **	c) *	a) *
Kim et al., 2013	a) *	a) *	c)	a) **	a) **	c) *	a) *
Kim et al., 2018	b) *	b)	c)	a) **	a) **	b) **	a) *
Ko, 2014	b) *	a) *	c)	b)	a) **	c) *	a) *
Koch-Weser et al., 2006	a) *	b)	a) *	a) **	c)	c) *	b)
H. Y. Lee et al., 2012	b) *	b)	a) *	a) **	a) **	c) *	a) *
H. Y. Lee et al., 2014	b) *	b)	c)	a) **	a) **	b) **	a) *
H. Y. Lee et al., 2015a	b) *	b)	c)	a) **	a) **	c) *	a) *
H. Y. Lee et al., 2015b	b) *	a) *	c)	a) **	b) *	c) *	a) *

Included Studies	Selection			Comparability	Exposure and Outcome		
	Sample representative?	Sample size justified?	Non-respondents	Comparability of subjects in different outcome groups	Ascertainment of the exposure	Assessment of the outcome	Statistical test
H. Y. Lee et al., 2015c	b) *	b)	c)	a) **	b) *	c) *	a) *
Mantwill et al., 2017	b) *	b)	c)	b)	a) **	b) **	a) *
Ng et al., 2014	b) *	a) *	a) *	a) **	a) **	b) **	b)
Njeru et al., 2016	d)	b)	c)	b)	b) *	b) **	a) *
Omariba et al., 2011	b) *	a) *	c)	a) **	a) **	c) *	a) *
Omariba et al., 2015	b) *	a) *	c)	a) **	a) **	c) *	a) *
Prins et al., 2015	a) *	b)	c)	a) **	b) *	c) *	a) *
Thabit et al., 2009	a) *	b)	c)	b)	a) **	b) **	a) *
Thomson et al., 2011	b) *	b)	c)	a) **	a) **	b) **	a) *
Todd et al., 2011a	b) *	b)	c)	b)	a) **	c) *	a) *

Included Studies	Selection			Comparability	Exposure and Outcome		
	Sample representative?	Sample size justified?	Non-respondents	Comparability of subjects in different outcome groups	Ascertainment of the exposure	Assessment of the outcome	Statistical test
Todd et al., 2011b	b) *	b)	c)	a) **	a) **	b) **	a) *
Tsoh et al., 2016	b) *	b)	c)	a) **	a) **	c) *	a) *
Wangdahl et al., 2014	b) *	b)	c)	a) **	b) *	c) *	a) *
Wangdahl et al., 2015	b) *	b)	c)	a) **	b) *	c) *	a) *
Wangdahl et al., 2018	b) *	b)	c)	a) **	a) **	c) *	a) *
Wister et al., 2010	b) *	b)	c)	a) **	a) **	b) **	b)
Yunusa Vakkai, 2016	b) *	b)	c)	a) **	b) *	c) *	a) *

Appendix L. The Risk of Bias in Non-Randomized Studies-of Interventions (ROBINS-I) Assessment Tool (version for cohort-type studies) (<https://sites.google.com/site/riskofbiastool//welcome/home>)

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool

(version for cohort-type studies)

Developed by: Jonathan AC Sterne, Miguel A Hernán, Barnaby C Reeves, Jelena Savović, Nancy D Berkman, Meera Viswanathan, David Henry, Douglas G Altman, Mohammed T Ansari, Isabelle Boutron, James Carpenter, An-Wen Chan, Rachel Churchill, Asbjørn Hróbjartsson, Jamie Kirkham, Peter Jüni, Yoon Loke, Terri Pigott, Craig Ramsay, Deborah Regidor, Hannah Rothstein, Lakhbir Sandhu, Pasqualina Santaguida, Holger J Schünemann, Beverly Shea, Ian Shrier, Peter Tugwell, Lucy Turner, Jeffrey C Valentine, Hugh Waddington, Elizabeth Waters, Penny Whiting and Julian PT Higgins

Version 1 August 2016



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants

Experimental intervention

Comparator

Outcomes

List the confounding domains relevant to all or most studies

--

List co-interventions that could be different between intervention groups and that could impact on outcomes

--

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	
Experimental intervention	
Comparator	

Is your aim for this study...?

- to assess the effect of *assignment to* intervention
- to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

--

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

--

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

“Important” confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. “Validity” refers to whether the confounding variable or variables fully measure the domain, while “reliability” refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that “no statistically significant association” is not the same as “not predictive”.

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

“Important” co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment (cohort-type studies)

Responses underlined in green are potential markers for low risk of bias, and responses in **red** are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Bias domain	Signalling questions	Elaboration	Response options
Bias due to confounding	1.1 Is there potential for confounding of the effect of intervention in this study? If <u>N/PN</u> to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:	In rare situations, such as when studying harms that are very unlikely to be related to factors that influence treatment decisions, no confounding is expected and the study can be considered to be at low risk of bias due to confounding, equivalent to a fully randomized trial. There is no NI (No information) option for this signalling question.	Y / PY / <u>PN / N</u>
	1.2. Was the analysis based on splitting participants' follow up time according to intervention received? If N/PN , answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY , proceed to question 1.3.	If participants could switch between intervention groups then associations between intervention and outcome may be biased by time-varying confounding. This occurs when prognostic factors influence switches between intended interventions.	NA / Y / PY / PN / N / NI
	1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN , answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY , answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)	If intervention switches are unrelated to the outcome, for example when the outcome is an unexpected harm, then time-varying confounding will not be present and only control for baseline confounding is required.	NA / Y / PY / PN / N / NI
	Questions relating to baseline confounding only		
	1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	Appropriate methods to control for measured confounders include stratification, regression, matching, standardization, and inverse probability weighting. They may control for individual variables or for the estimated propensity score. Inverse probability weighting is based on a function of the propensity score. Each method depends on the assumption that there is no unmeasured or residual confounding.	NA / <u>Y / PY / PN / N / NI</u>

	<p>1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?</p>	<p>Appropriate control of confounding requires that the variables adjusted for are valid and reliable measures of the confounding domains. For some topics, a list of valid and reliable measures of confounding domains will be specified in the review protocol but for others such a list may not be available. Study authors may cite references to support the use of a particular measure. If authors control for confounding variables with no indication of their validity or reliability pay attention to the subjectivity of the measure. Subjective measures (e.g. based on self-report) may have lower validity and reliability than objective measures such as lab findings.</p>	<p>NA / Y/PY / PN/N / NI</p>
	<p>1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?</p>	<p>Controlling for post-intervention variables that are affected by intervention is not appropriate. Controlling for mediating variables estimates the direct effect of intervention and may introduce bias. Controlling for common effects of intervention and outcome introduces bias.</p>	<p>NA / Y/PY / PN/N / NI</p>
<p>Questions relating to baseline and time-varying confounding</p>			
	<p>1.7. Did the authors use an appropriate analysis method that adjusted for all the important confounding domains and for time-varying confounding?</p>	<p>Adjustment for time-varying confounding is necessary to estimate the effect of starting and adhering to intervention, in both randomized trials and NRSI. Appropriate methods include those based on inverse probability weighting. Standard regression models that include time-updated confounders may be problematic if time-varying confounding is present.</p>	<p>NA / Y/PY / PN/N / NI</p>
	<p>1.8. If Y/PY to 1.7: Were confounding domains that were adjusted for measured validly and reliably by the variables available in this study?</p>	<p>See 1.5 above.</p>	<p>NA / Y/PY / PN/N / NI</p>
<p>Risk of bias judgement</p>			
	<p>Optional: What is the predicted direction of bias due to confounding?</p>	<p>Can the true effect estimate be predicted to be greater or less than the estimated effect in the study because one or more of the important confounding domains was not controlled for? Answering this question will be based on expert knowledge and results in other studies and therefore can only be completed after all of the studies in the body of evidence have been reviewed. Consider the potential effect of each of the unmeasured domains and whether all important confounding domains not controlled for in the analysis would be likely to change the estimate in the same direction, or if one important confounding domain that was not controlled for in the analysis is likely to have a dominant impact.</p>	<p>Low / Moderate / Serious / Critical / NI Favours experimental / Favours comparator / Unpredictable</p>

Bias in selection of participants into the study	<p>2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? If N/PN to 2.1: go to 2.4</p> <p>2.2. If Y/PY to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?</p> <p>2.3 If Y/PY to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?</p>	<p>This domain is concerned only with selection into the study based on participant characteristics observed <i>after</i> the start of intervention. Selection based on characteristics observed <i>before</i> the start of intervention can be addressed by controlling for imbalances between experimental intervention and comparator groups in baseline characteristics that are prognostic for the outcome (baseline confounding).</p> <p>Selection bias occurs when selection is related to an effect of either intervention or a cause of intervention and an effect of either the outcome or a cause of the outcome. Therefore, the result is at risk of selection bias if selection into the study is related to both the intervention and the outcome.</p>	<p>Y / PY / <u>PN / N</u> / NI</p> <p>NA / Y / PY / <u>PN / N</u> / NI</p> <p>NA / Y / PY / <u>PN / N</u> / NI</p>
	2.4. Do start of follow-up and start of intervention coincide for most participants?	If participants are not followed from the start of the intervention then a period of follow up has been excluded, and individuals who experienced the outcome soon after intervention will be missing from analyses. This problem may occur when prevalent, rather than new (incident), users of the intervention are included in analyses.	Y / PY / PN / N / NI
	2.5. If Y/PY to 2.2 and 2.3, or N/PN to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?	It is in principle possible to correct for selection biases, for example by using inverse probability weights to create a pseudo-population in which the selection bias has been removed, or by modelling the distributions of the missing participants or follow up times and outcome events and including them using missing data methodology. However such methods are rarely used and the answer to this question will usually be “No”.	NA / <u>Y / PY</u> / PN / N / NI
	Risk of bias judgement	See Table 1.	Low / Moderate / Serious / Critical / NI
	Optional: What is the predicted direction of bias due to selection of participants into the study?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable

Bias in classification of interventions	3.1 Were intervention groups clearly defined?	A pre-requisite for an appropriate comparison of interventions is that the interventions are well defined. Ambiguity in the definition may lead to bias in the classification of participants. For individual-level interventions, criteria for considering individuals to have received each intervention should be clear and explicit, covering issues such as type, setting, dose, frequency, intensity and/or timing of intervention. For population-level interventions (e.g. measures to control air pollution), the question relates to whether the population is clearly defined, and the answer is likely to be 'Yes'.	Y / PY / PN / N / NI
	3.2 Was the information used to define intervention groups recorded at the start of the intervention?	In general, if information about interventions received is available from sources that could not have been affected by subsequent outcomes, then differential misclassification of intervention status is unlikely. Collection of the information at the time of the intervention makes it easier to avoid such misclassification. For population-level interventions (e.g. measures to control air pollution), the answer to this question is likely to be 'Yes'.	Y / PY / PN / N / NI
	3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Collection of the information at the time of the intervention may not be sufficient to avoid bias. The way in which the data are collected for the purposes of the NRSI should also avoid misclassification.	Y / PY / PN / N / NI
	Risk of bias judgement	See Table 1.	Low / Moderate / Serious / Critical / NI
	Optional: What is the predicted direction of bias due to measurement of outcomes or interventions?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable

Bias due to deviations from intended interventions	If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
	4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	<p>Deviations that happen in usual practice following the intervention (for example, cessation of a drug intervention because of acute toxicity) are part of the intended intervention and therefore do not lead to bias in the effect of assignment to intervention.</p> <p>Deviations may arise due to expectations of a difference between intervention and comparator (for example because participants feel unlucky to have been assigned to the comparator group and therefore seek the active intervention, or components of it, or other interventions). Such deviations are not part of usual practice, so may lead to biased effect estimates. However these are not expected in observational studies of individuals in routine care.</p>	Y / PY / <u>PN</u> / N / NI
	4.2. If Y/PY to 4.1: Were these deviations from intended intervention unbalanced between groups <i>and</i> likely to have affected the outcome?	Deviations from intended interventions that do not reflect usual practice will be important if they affect the outcome, but not otherwise. Furthermore, bias will arise only if there is imbalance in the deviations across the two groups.	NA / Y / PY / <u>PN</u> / N / NI
	If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
	4.3. Were important co-interventions balanced across intervention groups?	Risk of bias will be higher if unplanned co-interventions were implemented in a way that would bias the estimated effect of intervention. Co-interventions will be important if they affect the outcome, but not otherwise. Bias will arise only if there is imbalance in such co-interventions between the intervention groups. Consider the co-interventions, including any pre-specified co-interventions, that are likely to affect the outcome and to have been administered in this study. Consider whether these co-interventions are balanced between intervention groups.	<u>Y</u> / PY / <u>PN</u> / N / NI
	4.4. Was the intervention implemented successfully for most participants?	Risk of bias will be higher if the intervention was not implemented as intended by, for example, the health care professionals delivering care during the trial. Consider whether implementation of the intervention was successful for most participants.	<u>Y</u> / PY / <u>PN</u> / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?	Risk of bias will be higher if participants did not adhere to the intervention as intended. Lack of adherence includes imperfect compliance, cessation of intervention, crossovers to the comparator intervention and switches to another active intervention. Consider available information on the proportion of study participants who continued with their assigned	<u>Y</u> / PY / <u>PN</u> / N / NI	

		<p>intervention throughout follow up, and answer ‘No’ or ‘Probably No’ if this proportion is high enough to raise concerns. Answer ‘Yes’ for studies of interventions that are administered once, so that imperfect adherence is not possible.</p> <p>We distinguish between analyses where follow-up time after interventions switches (including cessation of intervention) is assigned to (1) the new intervention or (2) the original intervention. (1) is addressed under time-varying confounding, and should not be considered further here.</p>	
	<p>4.6. If N/PN to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?</p>	<p>It is possible to conduct an analysis that corrects for some types of deviation from the intended intervention. Examples of appropriate analysis strategies include inverse probability weighting or instrumental variable estimation. It is possible that a paper reports such an analysis without reporting information on the deviations from intended intervention, but it would be hard to judge such an analysis to be appropriate in the absence of such information. Specialist advice may be needed to assess studies that used these approaches.</p> <p>If everyone in one group received a co-intervention, adjustments cannot be made to overcome this.</p>	<p>NA / <u>Y/PY</u> / PN / N / NI</p>
	<p>Risk of bias judgement</p>	<p>See Table 2</p>	
	<p>Optional: What is the predicted direction of bias due to deviations from the intended interventions?</p>	<p>If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.</p>	

Bias due to missing data	5.1 Were outcome data available for all, or nearly all, participants?	“Nearly all” should be interpreted as “enough to be confident of the findings”, and a suitable proportion depends on the context. In some situations, availability of data from 95% (or possibly 90%) of the participants may be sufficient, providing that events of interest are reasonably common in both intervention groups. One aspect of this is that review authors would ideally try and locate an analysis plan for the study.	<u>Y</u> / <u>PY</u> / <u>PN</u> / <u>N</u> / NI
	5.2 Were participants excluded due to missing data on intervention status?	Missing intervention status may be a problem. This requires that the <i>intended</i> study sample is clear, which it may not be in practice.	Y / <u>PY</u> / <u>PN</u> / <u>N</u> / NI
	5.3 Were participants excluded due to missing data on other variables needed for the analysis?	This question relates particularly to participants excluded from the analysis because of missing information on confounders that were controlled for in the analysis.	Y / <u>PY</u> / <u>PN</u> / <u>N</u> / NI
	5.4 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?	This aims to elicit whether either (i) differential proportion of missing observations or (ii) differences in reasons for missing observations could substantially impact on our ability to answer the question being addressed. “Similar” includes some minor degree of discrepancy across intervention groups as expected by chance.	NA / <u>Y</u> / <u>PY</u> / <u>PN</u> / <u>N</u> / NI
	5.5 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?	Evidence for robustness may come from how missing data were handled in the analysis and whether sensitivity analyses were performed by the investigators, or occasionally from additional analyses performed by the systematic reviewers. It is important to assess whether assumptions employed in analyses are clear and plausible. Both content knowledge and statistical expertise will often be required for this. For instance, use of a statistical method such as multiple imputation does not guarantee an appropriate answer. Review authors should seek naïve (complete-case) analyses for comparison, and clear differences between complete-case and multiple imputation-based findings should lead to careful assessment of the validity of the methods used.	NA / <u>Y</u> / <u>PY</u> / <u>PN</u> / <u>N</u> / NI
	Risk of bias judgement	See Table 2	Low / Moderate / Serious / Critical / NI
	Optional: What is the predicted direction of bias due to missing data?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable

Bias in measurement of outcomes	6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	Some outcome measures involve negligible assessor judgment, e.g. all-cause mortality or non-repeatable automated laboratory assessments. Risk of bias due to measurement of these outcomes would be expected to be low.	Y / PY / <u>PN</u> / N / NI
	6.2 Were outcome assessors aware of the intervention received by study participants?	If outcome assessors were blinded to intervention status, the answer to this question would be 'No'. In other situations, outcome assessors may be unaware of the interventions being received by participants despite there being no active blinding by the study investigators; the answer this question would then also be 'No'. In studies where participants report their outcomes themselves, for example in a questionnaire, the outcome assessor is the study participant. In an observational study, the answer to this question will usually be 'Yes' when the participants report their outcomes themselves.	Y / PY / <u>PN</u> / N / NI
	6.3 Were the methods of outcome assessment comparable across intervention groups?	Comparable assessment methods (i.e. data collection) would involve the same outcome detection methods and thresholds, same time point, same definition, and same measurements.	<u>Y</u> / PY / <u>PN</u> / N / NI
	6.4 Were any systematic errors in measurement of the outcome related to intervention received?	This question refers to differential misclassification of outcomes. Systematic errors in measuring the outcome, if present, could cause bias if they are related to intervention or to a confounder of the intervention-outcome relationship. This will usually be due either to outcome assessors being aware of the intervention received or to non-comparability of outcome assessment methods, but there are examples of differential misclassification arising despite these controls being in place.	Y / PY / <u>PN</u> / N / NI
	Risk of bias judgement	See Table 2	Low / Moderate / Serious / Critical / NI
	Optional: What is the predicted direction of bias due to measurement of outcomes?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable

Bias in selection of the reported result	Is the reported effect estimate likely to be selected, on the basis of the results, from...		
	7.1. ... multiple outcome <i>measurements</i> within the outcome domain?	For a specified outcome domain, it is possible to generate multiple effect estimates for different measurements. If multiple measurements were made, but only one or a subset is reported, there is a risk of selective reporting on the basis of results.	Y / PY / <u>PN</u> / N / NI
	7.2 ... multiple <i>analyses</i> of the intervention-outcome relationship?	Because of the limitations of using data from non-randomized studies for analyses of effectiveness (need to control confounding, substantial missing data, etc), analysts may implement different analytic methods to address these limitations. Examples include unadjusted and adjusted models; use of final value vs change from baseline vs analysis of covariance; different transformations of variables; a continuously scaled outcome converted to categorical data with different cut-points; different sets of covariates used for adjustment; and different analytic strategies for dealing with missing data. Application of such methods generates multiple estimates of the effect of the intervention versus the comparator on the outcome. If the analyst does not pre-specify the methods to be applied, and multiple estimates are generated but only one or a subset is reported, there is a risk of selective reporting on the basis of results.	Y / PY / <u>PN</u> / N / NI
	7.3 ... different <i>subgroups</i> ?	Particularly with large cohorts often available from routine data sources, it is possible to generate multiple effect estimates for different subgroups or simply to omit varying proportions of the original cohort. If multiple estimates are generated but only one or a subset is reported, there is a risk of selective reporting on the basis of results.	Y / PY / <u>PN</u> / N / NI
	Risk of bias judgement	See Table 2	Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable	

Overall bias	Risk of bias judgement	See Table 3.	Low / Moderate / Serious / Critical / NI
	Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Table 1. Reaching risk of bias judgements in ROBINS-I: pre-intervention and at-intervention domains

Judgement	Bias due to confounding	Bias in selection of participants into the study	Bias in classification of interventions
<u>Low risk of bias</u> (the study is comparable to a well-performed randomized trial with regard to this domain)	No confounding expected.	(i) All participants who would have been eligible for the target trial were included in the study; <i>and</i> (ii) For each participant, start of follow up and start of intervention coincided.	(i) Intervention status is well defined; <i>and</i> (ii) Intervention definition is based solely on information collected at the time of intervention.
<u>Moderate risk of bias</u> (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):	(i) Confounding expected, all known important confounding domains appropriately measured and controlled for; <i>and</i> (ii) Reliability and validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding.	(i) Selection into the study may have been related to intervention and outcome; <i>and</i> The authors used appropriate methods to adjust for the selection bias; <i>or</i> (ii) Start of follow up and start of intervention do not coincide for all participants; <i>and</i> (a) the proportion of participants for which this was the case was too low to induce important bias; <i>or</i> (b) the authors used appropriate methods to adjust for the selection bias; <i>or</i> (c) the review authors are confident that the rate (hazard) ratio for the effect of intervention remains constant over time.	(i) Intervention status is well defined; <i>and</i> (ii) Some aspects of the assignments of intervention status were determined retrospectively.

<u>Serious risk of bias</u> (the study has some important problems);	(i) At least one known important domain was not appropriately measured, or not controlled for; <i>or</i> (ii) Reliability or validity of measurement of an important domain was low enough that we expect serious residual confounding.	(i) Selection into the study was related (but not very strongly) to intervention and outcome; <i>and</i> This could not be adjusted for in analyses; <i>or</i> (ii) Start of follow up and start of intervention do not coincide; <i>and</i> A potentially important amount of follow-up time is missing from analyses; <i>and</i> The rate ratio is not constant over time.	(i) Intervention status is not well defined; <i>or</i> (ii) Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome.
<u>Critical risk of bias</u> (the study is too problematic to provide any useful evidence on the effects of intervention);	(i) Confounding inherently not controllable <i>or</i> (ii) The use of negative controls strongly suggests unmeasured confounding.	(i) Selection into the study was very strongly related to intervention and outcome; <i>and</i> This could not be adjusted for in analyses; <i>or</i> (ii) A substantial amount of follow-up time is likely to be missing from analyses; <i>and</i> The rate ratio is not constant over time.	(Unusual) An extremely high amount of misclassification of intervention status, e.g. because of unusually strong recall biases.
<u>No information</u> on which to base a judgement about risk of bias for this domain.	No information on whether confounding might be present.	No information is reported about selection of participants into the study or whether start of follow up and start of intervention coincide.	No definition of the intervention or no explanation of the source of information about intervention status is reported.



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Table 2. Reaching risk of bias judgements in ROBINS-I: post-intervention domains

Judgement	Bias due to deviations from intended intervention	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result
<u>Low risk of bias</u> (the study is comparable to a well-performed randomized trial with regard to this domain)	<p>Effect of assignment to intervention: (i) Any deviations from intended intervention reflected usual practice; <i>or</i> (ii) Any deviations from usual practice were unlikely to impact on the outcome.</p> <p>Effect of starting and adhering to intervention: The important co-interventions were balanced across intervention groups, and there were no deviations from the intended interventions (in terms of implementation or adherence) that were likely to impact on the outcome.</p>	<p>(i) Data were reasonably complete; <i>or</i> (ii) Proportions of and reasons for missing participants were similar across intervention groups; <i>or</i> (iii) The analysis addressed missing data and is likely to have removed any risk of bias.</p>	<p>(i) The methods of outcome assessment were comparable across intervention groups; <i>and</i> (ii) The outcome measure was unlikely to be influenced by knowledge of the intervention received by study participants (i.e. is objective) or the outcome assessors were unaware of the intervention received by study participants; <i>and</i> (iii) Any error in measuring the outcome is unrelated to intervention status.</p>	<p>There is clear evidence (usually through examination of a pre-registered protocol or statistical analysis plan) that all reported results correspond to all intended outcomes, analyses and sub-cohorts.</p>

Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):

Effect of assignment to intervention:
There were deviations from usual practice, but their impact on the outcome is expected to be slight.

Effect of starting and adhering to intervention:

(i) There were deviations from intended intervention, but their impact on the outcome is expected to be slight.

or

(ii) The important co-interventions were not balanced across intervention groups, or there were deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome;

and

The analysis was appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.

(i) Proportions of and reasons for missing participants differ slightly across intervention groups;

and

(ii) The analysis is unlikely to have removed the risk of bias arising from the missing data.

(i) The methods of outcome assessment were comparable across intervention groups;

and

(ii) The outcome measure is only minimally influenced by knowledge of the intervention received by study participants;

and

(iii) Any error in measuring the outcome is only minimally related to intervention status.

(i) The outcome measurements and analyses are consistent with an *a priori* plan; or are clearly defined and both internally and externally consistent;

and

(ii) There is no indication of selection of the reported analysis from among multiple analyses;

and

(iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.

<p><u>Serious risk of bias</u> (the study has some important problems);</p>	<p>Effect of assignment to intervention: There were deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome.</p> <p>Effect of starting and adhering to intervention: (i) The important co-interventions were not balanced across intervention groups, or there were deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome; <i>and</i> (ii) The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.</p>	<p>(i) Proportions of missing participants differ substantially across interventions; <i>or</i> Reasons for missingness differ substantially across interventions; <i>and</i> (ii) The analysis is unlikely to have removed the risk of bias arising from the missing data; <i>or</i> Missing data were addressed inappropriately in the analysis; <i>or</i> The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.</p>	<p>(i) The methods of outcome assessment were not comparable across intervention groups; <i>or</i> (ii) The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); <i>and</i> The outcome was assessed by assessors aware of the intervention received by study participants; <i>or</i> (iii) Error in measuring the outcome was related to intervention status.</p>	<p>(i) Outcomes are defined in different ways in the methods and results sections, or in different publications of the study; <i>or</i> (ii) There is a high risk of selective reporting from among multiple analyses; <i>or</i> (iii) The cohort or subgroup is selected from a larger study for analysis and appears to be reported on the basis of the results.</p>
---	--	---	--	--

Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention);

Effect of assignment to intervention:
There were substantial deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome.

Effect of starting and adhering to intervention:

(i) There were substantial imbalances in important co-interventions across intervention groups, or there were substantial deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome;

and

(ii) The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.

(i) (Unusual) There were critical differences between interventions in participants with missing data;
and
(ii) Missing data were not, or could not, be addressed through appropriate analysis.

The methods of outcome assessment were so different that they cannot reasonably be compared across intervention groups.

(i) There is evidence or strong suspicion of selective reporting of results;
and
(ii) The unreported results are likely to be substantially different from the reported results.

No information
on which to base
a judgement
about risk of bias
for this domain.

No information is reported on
whether there is deviation from the
intended intervention.

No information is reported
about missing data or the
potential for data to be
missing.

No information is reported
about the methods of
outcome assessment.

There is too little information
to make a judgement (for
example, if only an abstract is
available for the study).



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Table 3. Interpretation of domain-level and overall risk of bias judgements in ROBINS-I

Judgement	Within each domain	Across domains	Criterion
Low risk of bias	The study is comparable to a well-performed randomized trial with regard to this domain	The study is comparable to a well-performed randomized trial	The study is judged to be at low risk of bias for all domains .
Moderate risk of bias	The study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial	The study provides sound evidence for a non-randomized study but cannot be considered comparable to a well-performed randomized trial	The study is judged to be at low or moderate risk of bias for all domains .
Serious risk of bias	the study has some important problems in this domain	The study has some important problems	The study is judged to be at serious risk of bias in at least one domain, but not at critical risk of bias in any domain.
Critical risk of bias	the study is too problematic in this domain to provide any useful evidence on the effects of intervention	The study is too problematic to provide any useful evidence and should not be included in any synthesis	The study is judged to be at critical risk of bias in at least one domain .
No information	No information on which to base a judgement about risk of bias for this domain	No information on which to base a judgement about risk of bias	There is no clear indication that the study is at serious or critical risk of bias <i>and</i> there is a lack of information in one or more key domains of bias (<i>a judgement is required for this</i>).



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Appendix M. Risk of bias assessment for the included cohort quantitative studies

Included Studies	Bias due to or in							Overall Bias
	Confounding	Selection of Participants	Classification of Interventions	Deviations from Intended Interventions	Missing Data	Measurement of Outcomes	Selection of the Reported Result	
Borges et al., 2017	Critical risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Critical risk
P. C. Smith et al., 2012	Critical risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Critical risk

Appendix N. The Critical Appraisal Skills Program (CASP) for qualitative studies

Item 1: Was there a clear statement of the aims of the research? (yes or no)

Criteria for item 1:

1.1 Consider the goal of the research, its importance, and relevance

Item 2: Was a qualitative methodology appropriate? (yes or no)

Criteria for item 2:

2.1 Did the research seek to interpret or illuminate the actions and/or subjective experiences of research participants, or was qualitative research the right methodology for addressing the research goal?

Item 3: Was the research design appropriate to address the research aims? (yes or no)

Criteria for item 3:

3.1 Did the researcher justify the research design (e.g., did they discuss how they decided which methods to use)?

Item 4: Was the recruitment strategy appropriate to the aims of the research? (yes, limited, or no)

Criteria for item 4:

4.1 Did the researcher explain how the participants were selected?

4.2 Did the researcher explain why the participants they selected were the most appropriate to provide access to the type of knowledge sought by the study?

4.2 Were there any discussions around recruitment (e.g., why some people chose not to take part)?

Item 5: Were the data collected in a way that addressed the research aims? (yes, limited, or no)

Criteria for item 5:

- 5.1 Was the setting for data collection justified?
- 5.2 Was it clear how the data were collected (e.g., focus group, semistructured interview etc.)?
- 5.3 Did the researcher justify the methods chosen?
- 5.4 Did the researcher make the methods explicit (e.g. for interview method, is there an indication of how interviews were conducted, did they use a topic guide?)
- 5.5 If the methods were modified during the study, did the researcher explain how and why?
- 5.6 Was the form of data clear (e.g. tape recordings, video material)?
- 5.7 Did the researcher discuss saturation of data?

Item 6: Has the relationship between researcher and participants been adequately considered? (yes, limited, or no)

Criteria for item 6:

- 6.1 Did the researcher critically examine their own role, potential bias and influence during the formulation of the research questions?
- 6.2 Did the researcher critically examine their own role, potential bias and influence during data collection, including sample recruitment and choice of location?
- 6.3 Did the researcher discuss how they responded to events during the study, including the implications of any changes in the research design?

Item 7: Have ethical issues been taken into consideration? (yes, limited, or no)

Criteria for item 7:

- 7.1 Were there sufficient details of how the research was explained to participants for the reader to assess whether ethical standards were maintained?
- 7.2 Did the researcher discuss ethical issues raised by the study (e.g. confidentiality, informed consent, the effect of the study on the participants)
- 7.3 Was approval sought from an ethics committee?

Item 8: Was the data analysis sufficiently rigorous? (yes, limited, or no)

Criteria for item 8:

- 8.1 Was there an in-depth description of the analysis process?
- 8.2 If thematic analysis was used, was it clear how the categories/themes were derived from the data?
- 8.3 Did the researcher explain how the data presented were selected from the original sample to demonstrate the analysis process?
- 8.4 Were sufficient data presented to support the findings?
- 8.5 Were contradictory data taken into account?
- 8.6 Did the researcher critically examine their own role, potential bias and influence during analysis and selection of data for presentation?

Item 9: Is there a clear statement of findings? (yes, limited, or no)

Criteria for item 9:

- 9.1 Were the findings explicit?

9.2 Was there adequate discussion of the evidence both for and against the researcher's arguments?

9.3 Did the researcher discuss the credibility of their findings (e.g. triangulation, respondent validation, more than one analyst)?

9.4 Were the findings discussed in relation to the original research questions?

Item 10: How valuable is the research? (yes, limited, or no)

Criteria for item 10:

10.1 Did the researcher discuss the contribution the study makes to existing knowledge or understanding (e.g. did they consider the findings in relation to current practice or policy, or relevant research-based literature?)?

10.2 Did the researcher identify new areas where research is necessary?

10.3 Did the researcher discuss whether or how the findings can be transferred to other populations or consider other ways the research may be used?

Appendix O. Assessment of risk of bias for the included qualitative studies

Included Studies	1. clear statement of aims?	2. qualitative methodology appropriate?	3. research design justified?	4. recruitment strategy appropriate?	5. data collected in a way that addressed the research issue?	6. relationship between researcher and participants considered?	7. ethical issues considered?	8. data analysis sufficiently rigorous?	9. clear statement of findings?	10. How valuable is the research?
Alzayer et al., 2017	Yes	Yes	No	Limited (4.2)	Limited (5.3, 5.6)	Limited (6.1)	No	Limited (8.3, 8.5, 8.6)	Yes	Limited (10.3)
Carroll et al., 2007	Yes	Yes	No	Yes	Limited (5.1, 5.3)	Limited (6.1)	No	Limited (8.3, 8.6)	Limited (9.2)	Limited (10.3)
Clark et al., 2014	Yes	Yes	Yes	No	Limited (5.1, 5.7)	No	Yes	Limited (8.3, 8.6)	Limited (9.2, 9.3)	Limited (10.3)
Cyril et al., 2017	Yes	Yes	Yes	Limited (4.2)	Limited (5.6)	No	Limited (7.1, 7.2)	Limited (8.5, 8.6)	Yes	Limited (10.2)
Filippi et al., 2014	Yes	Yes	No	Limited (4.2)	Limited (5.1, 5.3, 5.7)	Limited (6.1)	Limited (7.1)	Limited (8.1, 8.3, 8.5)	Limited (9.2, 9.3)	Yes
Floyd et al., 2017	Yes	Yes	Yes	Yes	Limited (5.1, 5.3, 5.7)	Limited (6.1)	Yes	Limited (8.3)	Yes	Limited (10.2)

Included Studies	1. clear statement of aims?	2. qualitative methodology appropriate?	3. research design justified?	4. recruitment strategy appropriate?	5. data collected in a way that addressed the research issue?	6. relationship between researcher and participants considered?	7. ethical issues considered?	8. data analysis sufficiently rigorous?	9. clear statement of findings?	10. How valuable is the research?
Gele et al., 2017	Yes	Yes	Yes	Limited (4.2)	Limited (5.1, 5.7)	No	Yes	Limited (8.3, 8.6)	Limited (9.2)	Yes
Gregory, 2015	Yes	Yes	Yes	Limited (4.2)	Limited (5.3, 5.7)	Limited (6.1)	Yes	Limited (8.5)	Limited (9.2)	Yes
Groenberg et al., 2015	Yes	Yes	No	Yes	Limited (5.3)	Limited (6.1)	Limited (7.1, 7.2)	Limited (8.3, 8.5)	Yes	Yes
Hurley et al., 2013	Yes	Yes	No	Limited (4.2, 4.3)	Limited (5.1, 5.3, 5.7)	No	Limited (7.1, 7.2)	Limited (8.3, 8.5, 8.6)	Yes	Limited (10.2)
Jafri, 2012	Yes	Yes	Yes	Yes	Limited (5.7)	Limited (6.1)	Yes	Limited (8.3)	Yes	Yes
Jung et al., 2017	Yes	Yes	Yes	Limited (4.2)	Limited (5.1)	No	Limited (7.1, 7.2)	Limited (8.3, 8.5, 8.6)	Yes	Limited (10.2)

Included Studies	1. clear statement of aims?	2. qualitative methodology appropriate?	3. research design justified?	4. recruitment strategy appropriate?	5. data collected in a way that addressed the research issue?	6. relationship between researcher and participants considered?	7. ethical issues considered?	8. data analysis sufficiently rigorous?	9. clear statement of findings?	10. How valuable is the research?
F. H. Lee et al., 2014	Yes	Yes	No	Limited (4.2)	Limited (5.1, 5.3)	No	Limited (7.2)	Limited (8.6)	Limited (9.2)	Yes
Leung et al., 2014	Yes	Yes	Yes	Yes	Limited (5.1, 5.3, 5.7)	Limited (6.1)	Limited (7.1, 7.2)	Limited (8.3, 8.5)	Limited (9.2)	Yes
Marshall et al., 2010	Yes	Yes	Yes	Limited (4.1, 4.3)	Limited (5.1, 5.7)	No	Limited (7.1, 7.2)	Limited (8.3)	Limited (9.2)	Limited (10.3)
McMichael et al., 2009	Yes	Yes	No	Limited (4.3)	Limited (5.1, 5.7)	Limited (6.1)	Yes	Limited (8.3, 8.6)	Limited (9.2, 9.3)	Limited (10.3)
Murry et al., 2018	Yes	Yes	Yes	Limited (4.2, 4.3)	Limited (5.3, 5.7)	No	Limited (7.1, 7.2)	Limited (8.3, 8.4, 8.6)	Limited (9.2)	Limited (10.3)
Oliver, 2015	Yes	Yes	Yes	Yes	Yes	Limited (6.1)	Yes	Limited (8.3)	Limited (9.2)	Yes

Included Studies	1. clear statement of aims?	2. qualitative methodology appropriate?	3. research design justified?	4. recruitment strategy appropriate?	5. data collected in a way that addressed the research issue?	6. relationship between researcher and participants considered?	7. ethical issues considered?	8. data analysis sufficiently rigorous?	9. clear statement of findings?	10. How valuable is the research?
Renzaho et al., 2017	Yes	Yes	No	Limited (4.2)	Yes	No	Limited (7.1, 7.2)	Limited (8.6)	Limited (9.2)	Yes
Schoenmakers et al., 2017	Yes	Yes	No	Limited (4.2)	Yes	Limited (6.1)	Yes	Limited (8.3, 8.4, 8.5, 8.6)	Yes	Limited (10.2)
Shedlin et al., 2004	Yes	Yes	Yes	Limited (4.1, 4.2)	Limited (5.1, 5.7)	No	No	Limited (8.3, 8.6)	Limited (9.2, 9.3)	Yes
Smaland Goth et al., 2011	Yes	Yes	Yes	Limited (4.1)	Limited (5.1, 5.3, 5.7)	No	Limited (7.1, 7.2)	Limited (8.1, 8.2, 8.3, 8.4, 8.6)	Yes	Yes
Sriphanlop et al., 2014	Yes	Yes	No	Limited (4.2)	Limited (5.1, 5.3)	Limited (6.1)	Limited (7.1, 7.2)	Limited (8.3, 8.5, 8.6)	Limited (9.2)	Yes
Taiwo, 2013	Yes	Yes	Yes	Yes	Yes	No	Yes	Limited (8.3)	Limited (9.2)	Yes

Included Studies	1. clear statement of aims?	2. qualitative methodology appropriate?	3. research design justified?	4. recruitment strategy appropriate?	5. data collected in a way that addressed the research issue?	6. relationship between researcher and participants considered?	7. ethical issues considered?	8. data analysis sufficiently rigorous?	9. clear statement of findings?	10. How valuable is the research?
Thomson et al., 2012	Yes	Yes	No	Yes	Limited (5.1, 5.7)	Limited (6.1)	Limited (7.1, 7.2)	Limited (8.3, 8.5)	Limited (9.2)	Yes
Todd et al., 2011c	Yes	Yes	No	Limited (4.2)	Limited (5.1, 5.3, 5.6)	No	Limited (7.1, 7.2)	Limited (8.3, 8.5)	Limited (9.2)	Yes
Watts et al., 2014	Yes	Yes	Yes	Limited (4.2, 4.3)	Yes	No	Limited (7.2)	Limited (8.4)	Limited (9.2)	Yes
Woudstra et al., 2016	Yes	Yes	No	Limited (4.2)	Limited (5.3)	Limited (6.1)	Limited (7.2)	Limited (8.3, 8.4, 8.6)	Yes	Yes

Appendix P. The working thematic framework containing pre-defined themes based on Sorensen's et al., (2012) conceptual framework for health literacy and numeracy

Themes	Sub-themes	Descriptions
Antecedents		
Personal antecedents	Personal characteristics	Individual characteristics that may impact health literacy or numeracy (e.g., age, gender, race, socioeconomic status, education, occupation, employment, income, etc.)
	Personal competence	Individual abilities or skills that may impact health literacy or numeracy (e.g., memory and reasoning, physical abilities and social skills, and cognitive skills)
Societal and environmental antecedents	None	Factors from family members or friends, community, working, school, health care environment, or society (e.g., family support, policy)
Consequences		
Health service use or behaviour	None	Individuals participate or access health service (e.g., cancer screening) or individuals' belief and actions regarding their health (e.g., diabetes management)
Health outcome		Changes in individual health which result from different levels of health literacy or numeracy (e.g., mortality, health status)
Health cost		The cost of health care services (e.g., fees paid for a health care service)

Appendix Q. Extracted basic information from the included studies

No.	First Author	Publication Year	Study Type	Publication Type
1	Becerra	2017	Quantitative	Journal article
2	Bekker	2004	Quantitative	Journal article
3	Beltran	2016	Quantitative	Journal article
4	Borges	2017	Quantitative	Journal article
5	Calvasina	2016	Quantitative	Journal article
6	Calvo	2016	Quantitative	Journal article
7	S. E. Choi	2013	Quantitative	Journal article
8	Y. J. Choi	2016	Quantitative	Journal article
9	Coffman	2007	Quantitative	Journal article
10	Coffman	2010	Quantitative	Journal article
11	Coffman	2012	Quantitative	Journal article
12	Diamond	2014	Quantitative	Journal article
13	Gatobu	2014	Quantitative	Journal article
14	Gatobu	2016	Quantitative	Journal article
15	Gele	2016	Quantitative	Journal article
16	Geltman	2013	Quantitative	Journal article
17	Geltman	2014	Quantitative	Journal article
18	Hernandez-Mekonnen	2016	Quantitative	Journal article
19	Idehen	2017	Quantitative	Journal article

No.	First Author	Publication Year	Study Type	Publication Type
20	Igarashi	2013	Quantitative	Journal article
21	Jacobson	2016	Quantitative	Journal article
22	Kankou	2017	Quantitative	Journal article
23	Khuu	2018	Quantitative	Journal article
24	Kim	2013	Quantitative	Journal article
25	Kim	2018	Quantitative	Journal article
26	Ko	2014	Quantitative	PhD thesis
27	Koch-Weser	2006	Quantitative	Journal article
28	H. Y. Lee	2012	Quantitative	Journal article
29	H. Y. Lee	2014	Quantitative	Journal article
30	H. Y. Lee	2015a	Quantitative	Journal article
31	H. Y. Lee	2015b	Quantitative	Journal article
32	H. Y. Lee	2015c	Quantitative	Journal article
33	Mantwill	2017	Quantitative	Journal article
34	Ng	2014	Quantitative	Journal article
35	Njeru	2016	Quantitative	Journal article
36	Omariba	2011	Quantitative	Journal article
37	Omariba	2015	Quantitative	Journal article
38	Prins	2015	Quantitative	Journal article
39	P. C. Smith	2012	Quantitative	Journal article

No.	First Author	Publication Year	Study Type	Publication Type
40	Thabit	2009	Quantitative	Journal article
41	Thomson	2011	Quantitative	Journal article
42	Todd	2011a	Quantitative	Journal article
43	Todd	2011b	Quantitative	Journal article
44	Tsoh	2016	Quantitative	Journal article
45	Wangdahl	2014	Quantitative	Journal article
46	Wangdahl	2015	Quantitative	Journal article
47	Wangdahl	2018	Quantitative	Journal article
48	Wister	2010	Quantitative	Journal article
49	Yunusa Vakkai	2016	Quantitative	PhD thesis
50	Alzayer	2017	Qualitative	Journal article
51	Carroll	2007	Qualitative	Journal article
52	Clark	2014	Qualitative	Journal article
53	Cyril	2017	Qualitative	Journal article
54	Filippi	2014	Qualitative	Journal article
55	Floyd	2017	Qualitative	Journal article
56	Gele	2017	Qualitative	Journal article
57	Gregory	2015	Qualitative	PhD thesis
58	Groenenberg	2015	Qualitative	Journal article
59	Hurley	2013	Qualitative	Journal article

No.	First Author	Publication Year	Study Type	Publication Type
60	Jafri	2012	Mixed	PhD thesis
61	Jung	2017	Mixed	Journal article
62	F. H. Lee	2014	Qualitative	Journal article
63	Leung	2014	Qualitative	Journal article
64	Marshall	2010	Qualitative	Journal article
65	McMichael	2009	Qualitative	Journal article
66	Murry	2018	Qualitative	Journal article
67	Oliver	2015	Qualitative	PhD thesis
68	Renzaho	2017	Qualitative	Journal article
69	Schoenmakers	2017	Qualitative	Journal article
70	Shedlin	2004	Qualitative	Journal article
71	Smaland Goth	2011	Qualitative	Journal article
72	Sriphanlop	2014	Qualitative	Journal article
73	Taiwo	2013	Qualitative	PhD thesis
74	Thomson	2012	Qualitative	Journal article
75	Todd	2011c	Qualitative	Journal article
76	Watts	2014	Qualitative	Journal article
77	Woudstra	2016	Qualitative	Journal article

Appendix R. Descriptions of the assessment instruments used in the included studies

No.	Assessment tools	Brief Description
1	The Rapid Estimate of Adult Literacy (REALM), English language version (Davis et al., 1993)	The REALM consisting of 66 common medical terms that requires the participants to read the words aloud when taking the test. It does not assess the ability to comprehension.
2	The modified REALM, Japanese language version (Igarashi et al., 2013)	The REALM Japanese version used 30 items translated from the REALM English version.
3	The modified REALM, Dutch and Standard-Arabic language versions. (Bekker et al., 2004)	These two versions, translated from the original English version, removed and added some words or terms with higher relevance for the women under study.
4	The Test of Functional Health Literacy (TOFHLA) English and Spanish language versions (Parker et al., 1995)	The English version is a cloze-type comprehension assessment tools, which contain 50 reading comprehension items in four passages to test individuals' capacity to make sense of selected materials and fill in missing words and 17 numeracy items to measure numeracy skills. The Spanish version was developed by translating the TOFHLA English version into Spanish, and then translating it back into English.
5	The Short TOFHLA (S-TOFHLA) English language version (Baker et al., 1999)	A short version of TOFHLA English version, which contains 36 reading comprehension items in two passages and 4 numeracy items.
6	The S-TOFHLA Spanish language version (The study of initial development was not found. (Stonbraker et al., 2015))	A short version of TOFHLA Spanish version, which contains 36 reading comprehension items in two passages and 4 numeracy items.
7	The S-TOFHLA, Albanian, Portuguese, Serbian, German, and Italian language versions (Mantwill et al., 2017)	These tools were translated from S-TOFHLA English version.

No.	Assessment tools	Brief Description
8	The short language version of the European health literacy questionnaire (HLS-EU-Q16), English language version (HLS-EU Consortium, 2012)	A questionnaire, developed by the HLS-EU Consortium for assessing health literacy level of general populations, contains 16 statements. Participants are asked to provide their selection on a 5-point Likert scale (very easy, fairly easy, fairly difficult, very difficult, I don't know).
9	The modified HLS-EU-Q16, Arabic, Dari, Somali, English, and Swedish language versions (Pelikan et al., 2017; Wangdahl et al., 2014, 2015)	The modification of original HLS-EU-Q16 include providing full sentences in each item. For example, in the original Swedish version, the Q1 is "find information on treatments of illnesses that concern you? (very easy, fairly easy, fairly difficult, very difficult, I don't know)", while in the modified version, it is "How easy/difficult is it for you to find information on treatments of illnesses that concern you? (very easy, fairly easy, fairly difficult, very difficult, I don't know)."
10	The Newest Vital Sign (NVS), English language version (Weiss et al., 2005)	An instrument assessing participant's reading comprehension and numeracy skills based on a nutrition label from an ice cream container.
11	The NVS Korean language version (S. E. Choi et al., 2013)	A Korean version translated from the NVS English version.
12	The Oral Health Literacy Instrument Portuguese language version (Calvasina et al., 2016; Sabbahi et al., 2009)	A dentistry specific health literacy test which contains 38 cloze-type comprehension questions and 19 numeric calculations. The Portuguese version was translated from the English version.
13	The Parental Health Literacy Activities Test, Spanish language version (Yin et al., 2012)	A test for assessing Spanish-speaking caregiver or parental health literacy. The test asks questions such as "Demonstrates how to make a 4-ounce bottle of formula using powder-based formula."
14	The Rapid Estimate of Adult Literacy in Dentistry 30 Short Form, English language version (J. Y. Lee et al., 2007)	A word recognition test consisting of 30 common dental words with various degrees of difficulty.

No.	Assessment tools	Brief Description
15	The Short Assessment of Health Literacy for Spanish-speaking Adults, Spanish language version (Lee et al., 2006)	A test containing 30 understanding items and 17 evaluation items and assessing a Spanish-speaking adult's capability to read and understand common medical terms.
16	The Single Item Literacy Screener, English language version (Morris et al., 2006)	One question asking "How often do you need to have someone help you when you read instructions, pamphlets, or other written material from your doctor or pharmacy?" Participants select one of the following responses: 1-Never, 2-Rarely, 3-Sometimes, 4-Often, and 5-Always. This tool assesses health literacy by identifying adults in need of help with printed health material.
17	The Swedish functional health literacy scale, Arabic, Dari, Somali, and English language versions (Wangdahl et al., 2014, 2015)	Self-assessment tools each of which consists of 5 items measuring different domains of functional health literacy. They were translated from the original Swedish version.
18	The TOFHLA English and Spanish language versions (Parker et al., 1995)	The English version is a cloze-type comprehension assessment tools, which contain 50 reading comprehension items in four passages to test individuals' capacity to make sense of selected materials and fill in missing words and 17 numeracy items to measure numeracy skills. The Spanish version was developed by translating the TOFHLA English version into Spanish, and then translating it back into English.
19	The 12 survey items on cancer risk and prevention, English and Korean (H. Y. Lee et al., 2014, 2015c; Stein et al., 2007)	The original version (Stein et al., 2007) is a true or false questionnaire consisting of 12 questions regarding cancer risk and prevention. H. Y. Lee et al., (2014) adopted it to assess cancer health literacy. Participants' responses were scored (correct answer (disagree) = 1, incorrect (agree) = 0, and don't know was not calculated). High combined score indicates high cancer literacy. H. Y. Lee et al., (2015c) categorized the responses into a 5-point Likert scale ranging from "strongly disagree" (1 point) to "strongly agree" (5 points).

No.	Assessment tools	Brief Description
20	The Chew et al., (2004) 16-item health literacy screening scale, English and Korean	The English tool contains 16 health literacy screening questions on a 5-point Likert scale. H. Y. Lee et al., (2012) conducted this survey by four bilingual interviewers (Korean and English) for participants age 60 and older.
21	The 12 items from the Chew et al., (2004) 16-item health literacy screening scale, English and Korean	This tools adopted 12 items from the Chew et al., (2004) 16 items. Four items were removed due to poor reliability (H. Y. Lee et al., 2015a). H. Y. Lee et al., (2015a) conducted this survey by four bilingual interviewers (Korean and English) for participants age 60 and older.
22	The Brief Health Literacy Screeners (Chew Items), English language version (Chew et al., 2004)	This tool asks participants three questions “How often do you have someone help you read hospital materials?”, “How confident are you in filling out medical forms by yourself?” and “How often do you have problems learning about your medical condition because of difficulty understanding written information?” The 3 questions were validated from the Chew et al., (2004) 16 items as effective items to identify inadequate health literacy. Each item was rated with a 5-point Likert scale.
23	The Brief Health Literacy Screeners (Chew items), Albanian, Portuguese, and Serbian language versions ((Chew et al., 2004; Mantwill et al., 2017)	These items were translated from the Chew Items English version.
24	The Brief Health Literacy Screener (Chew Items), Spanish language version (Chew et al., 2004; Sarkar et al., 2011)	This tool was translated from the original English version.

No.	Assessment tools	Brief Description
25	The Korean Language Literacy Scale (Kim et al., 2013)	A scale for assessing the self-reported Korean language (Hangul) fluency level. Four items were included: “How well do you speak Korean?”; “How well do you write Korean?”; “How well do you read Korean?”; and “How well do you understand Korean?” Each item was rated with a 5-point Likert scale: poor; fair; good; very good; and excellent. A higher score indicates a greater perception of Korean language literacy.
26	The Assessment of Colon Cancer Literacy, Korean language version (Holubar et al., 2009; Ko, 2014)	An instrument consisting of 10 items requiring the participants to provide their responses. The possible responses include true, false, or unsure. An example of the items is "Men get cancer of the bowel more often than women. (true, false, unsure)"
27	The Assessment of Health Literacy in Cancer Screening, English version (Han et al., 2014)	An instrument modeled from the REALM and the TOFHLA and consisting of four subscales: print literacy, numeracy, comprehension, and familiarity relating to breast and cervical cancer screening.
28	The Spoken Knowledge in Low Literacy in Diabetes scale, English and Somali language versions (Njeru et al., 2016; Rothman et al., 2005)	A scale consisting 10 questions regarding patient's knowledge on diabetes self-care issues. The test administrator checks if participant's response was consistent with acceptable responses. For example: "How often should a person with diabetes check his or her feet? Once a day, once a week, or once a month? Accept: Daily" A score (0-100%) will be given, with 0 representing lowest literacy, and 100% as the highest.
29	The Cloze (Thomson et al., 2011)	The Cloze, based on an online colon cancer prevention information page (one page) developed by the Canadian Cancer Society for the general public, consists of 22 questions requiring participants to fill in the blanks.

No.	Assessment tools	Brief Description
30	The 7-item scale adapted from the 16-item scale of McPartland et al., (2005), English language version	A scale reducing McPartland's et al., (2005) 16 items to 7 items, including "Most people with genital HPV have no visible signs or symptoms. HPV infection can cause cervical cancer. A vaccine exists to prevent HPV infection. I can transmit HPV to my partner(s) even if I have no HPV symptoms. If a woman's Pap test is normal, she does not have HPV. A negative test for HPV means that you do not have HPV. Pap tests will almost always detect HPV." The participants need to provide true or false responses. The more correct answers the participants give, the greater their HPV literacy.
31	The Mental Health Literacy Scale, Korean language version (Y. J. Choi et al., 2016)	A scale consisting 25 items using a 4-point Likert scale. The scale examines three domains: knowledge (concept, epidemiology), attitudes (symptoms, treatment), and stigma (shame, fear) about mental health problem.
32	The dementia literacy survey, Chinese language version (Diamond et al., 2014; Woo, 2013)	A questionnaire assessing comprehension of dementia symptoms, cause, treatment and prognosis consisting of 11 statements requiring the participants to provide true or false responses. A higher score indicates a more complete understanding of dementia.
33	The two health literacy items California Health Interview Survey (California Health Interview Survey, 2011)	Two questions requiring participants to provide responses with a 4-point Likert scale: very easy, somewhat easy, somewhat difficult, or very difficult. Originally, a higher score indicates a lower level of health literacy. Becerra et al., (2017) considered participants as low health literacy if they reported somewhat difficult or very difficult to either of the questions; H. Y. Lee et al., (2015b) reversed the scores of the responses so that a higher score indicated a higher health literacy.
34	The Health Activities Literacy Scale (HALS), English language version (Canadian Council on Learning, 2008; Rudd et al., 2004)	A scale consisting 191 items and assessing health-related abilities in health promotion, health protection, disease prevention, health care maintenance, and system navigation. Scores range from 0 to 500. Participants with a score greater than 275 is considered to have adequate health literacy.

No.	Assessment tools	Brief Description
35	The literacy and numeracy tasks from the Program for the International Assessment of Adult Competencies, English language version (OECD, 2012)	The literacy tasks address three cognitive strategies: access and identify, integrate and interpret, and evaluate and reflect, while the numeracy items assess "the ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life." The test is computer-based.
36	The addition-subtraction correction and the addition tasks from the Kit of Factor-Referenced Cognition Test (the French kit) (Ekstrom et al., 1979)	The addition-subtraction correction task includes 60 already solved problems, which require the test takers to mark the answers as "correct" or "incorrect." The score was the difference between the number of correct answers and incorrect answers. The addition task involves 60 unsolved problems of one or two digits. Participants need to provide an answer for each problem, with the score indicating the number of correct answers.

Appendix S. The matrix used to summarize antecedents of health literacy or health numeracy

Included Studies	Personal Antecedents		Societal and Environmental Antecedents	Inductively Identified Theme
	Personal Characteristics	Personal Competence		Personal Belief, Experience, and Behaviour
Becerra et al., 2017	Sex; living in poverty; lacking consistent health insurance	Limited english language proficiency		
Beltran et al., 2016	Marital status; years in USA; health insurance; participants' level of education	Cervical cancer screening literacy	Family cancer history	Number of doctor visits; annual health check-up; having a usual source of care
S. E. Choi et al., 2013	Age; gender; income; years of education; acculturation	English proficiency		Health behaviours such as fruit intake, vegetable intake, soda intake, activity
Y. J. Choi et al., 2016	Levels of acculturation stress			
Diamond et al., 2014	Duration of residence			
Gatobu et al., 2014	Gender, primary language, residency in Canada, education- > high school < university, education-university, education-Graduate		Format of information	Self-efficacy

Included Studies	Personal Antecedents		Societal and Environmental Antecedents	Inductively Identified Theme
	Personal Characteristics	Personal Competence		Personal Belief, Experience, and Behaviour
Gatobu et al., 2016	Primary language difference (Mandarin and Kikuyu)			
Gele et al., 2016	Acculturation; age; years in Norway; employment; education			
Jacobson et al., 2016	Age	English proficiency		
Khuu et al., 2018	The number of years participants have lived in the U.S.; health status; age; marital status; income	Whether they had difficulties with activities of daily living		Social or religious group attendance; usual place of care
H. Y. Lee et al., 2012	Education; health insurance; gender (women); marital status (married); age	English proficiency		

Included Studies	Personal Antecedents		Societal and Environmental Antecedents	Inductively Identified Theme
	Personal Characteristics	Personal Competence		Personal Belief, Experience, and Behaviour
H. Y. Lee et al., 2014	Age; educational attainment; sex; marital status; years in the USA; monthly income; subjective economic status; healthcare insurance; subjective self-rated health status; number of chronic diseases	English proficiency	Family support; family cancer history	Having a primary care physician; smoking frequency, exercise frequency; healthcare motivation;
H. Y. Lee et al., 2015a	Gender (female); educational attainment	english proficiency		Cultural modesty, having a primary care physician
Mantwill et al., 2017	Length of stay; age at immigration			
Ng et al., 2014	Age; own education; household income		Maternal education; literacy practices at work; literacy practice at home	Participation in adult education
Thomson et al., 2011	Age; acculturation; years of Spanish education; employment; duration of residency			Media-internet use; and media-television use

Included Studies	Personal Antecedents		Societal and Environmental Antecedents	Inductively Identified Theme
	Personal Characteristics	Personal Competence		Personal Belief, Experience, and Behaviour
Todd et al., 2011b	Acculturation; age; education	Self-reported proficiency reading english		
Wangdahl et al., 2014	Low education; sex; age; years of resident permit; long term sickness; country of origin; being born in Somalia	Having inadequate functional health literacy		
Wister et al., 2010	Years of immigration			
Carroll et al., 2007			Community-based health education programs	
Leung et al., 2014	Insurance; age related limitations		Transportation issues; limited information in chinese-speaking communities	Namely unawareness of self-care responsibility; a desire to avoid being burdensome to others; high regard for authority; a desire to be together or follow a collective approach
McMichael et al., 2009				Experiences prior to arrival at destination country

Included Studies	Personal Antecedents		Societal and Environmental Antecedents	Inductively Identified Theme
	Personal Characteristics	Personal Competence		Personal Belief, Experience, and Behaviour
Murry et al., 2018			Involve the person's support network; less reliance on written instruction	
Taiwo, 2013	Affordability concerns to healthcare literacy and access care	Challenges with language and communication issues		Religio-cultural beliefs and assumptions on health

Appendix T. The matrix used to summarize consequences of health literacy or health numeracy

Included Studies	Health Service Use or Behaviour	Health Outcome	Health Cost	Inductively Identified Theme	
				Health Experience and Perception	Health Knowledge and Understanding
Bekker et al., 2004		Physical functioning, health perception			
Borges et al., 2017		Relapse rate in children with nephrotic syndrome			
Calvasina et al., 2016	Not visiting a dentist in the preceding year; not having a dentist as the primary source of dental information; not participating in shared dental treatment decision making				
Calvo, 2016				Reporting high qoc	
S. E. Choi et al., 2013		Waist to hip ratio (WHR); blood glucose level; BMI; systolic blood pressure; diastolic blood pressure			

Included Studies	Health Service Use or Behaviour	Health Outcome	Health Cost	Inductively Identified Theme	
				Health Experience and Perception	Health Knowledge and Understanding
Y. J. Choi et al., 2016		Mental health			
Coffman et al., 2007	Health care service use				
Coffman et al., 2010		Depression score of the 20-item Spanish Language Center for Epidemiologic Studies Depression Scale			
Coffman et al., 2012	Health care use				
Geltman et al., 2013		Decayed, missing, and filled teeth; risk of periodontal disease; untreated decay			
Geltman et al., 2014	Preventive dental care				
Hernandez-Mekonnen et al., 2016	Use of early intervention services	Child risk for developmental delay			

Included Studies	Health Service Use or Behaviour	Health Outcome	Health Cost	Inductively Identified Theme	
				Health Experience and Perception	Health Knowledge and Understanding
Idehen et al., 2017	Likelihood of Pap test (cervical cancer screening) participation				
Igarashi et al., 2013				Quality of Care for Pregnancy, Delivery and Postpartum Questionnaire (QCQ) scores, including Respect (feeling respected), Understanding (feeling being understood by health care providers) and Cold (feeling distant and unhelpful attitude) subscores; (b) care satisfaction	
Kankou et al., 2017	HIV disclosure to a steady partner, relatives, or others				
Kim et al., 2013		Depression			

Included Studies	Health Service Use or Behaviour	Health Outcome	Health Cost	Inductively Identified Theme	
				Health Experience and Perception	Health Knowledge and Understanding
Kim et al., 2018	Pap test use				
Ko, 2014	Having a sigmoidoscopy; having a colonoscopy; having a fecal occult blood test (colorectal cancer screening)				
Koch-Weser et al., 2006		Self-rated health			
H. Y. Lee et al., 2015b		Self-reported health status; depression symptoms			
H. Y. Lee et al., 2015c	Cervical cancer screening				
Mantwill et al., 2017		General health status			
Murry et al., 2018	Medication management				
Njeru et al., 2016		Diabetes outcomes			
Omariba et al., 2011		Self-rated health			
Omariba et al., 2015		Disability			
Prins et al., 2015		Self-rated health			

Included Studies	Health Service Use or Behaviour	Health Outcome	Health Cost	Inductively Identified Theme	
				Health Experience and Perception	Health Knowledge and Understanding
P. C. Smith et al., 2012					Follow up with a physician as instructed (frequency); understanding discharge instructions; filling prescriptions as instructed
Thabit et al., 2009	Glycemic control				
Todd et al., 2011a	Colon and breast cancer screening				
Tsoh et al., 2016		Self-rated health			
Wangdahl et al., 2015				(a) having experienced poor quality of communication, (b) having received little health care information, (c) not having received any new knowledge, (d) not having received any help	

Included Studies	Health Service Use or Behaviour	Health Outcome	Health Cost	Inductively Identified Theme	
				Health Experience and Perception	Health Knowledge and Understanding
Wangdahl et al., 2018	Having refrained from seeking healthcare	Poor health status; impaired psychological well-being			
Yunusa Vakkai, 2016		General health status			
Alzayer et al., 2017	Asthma control				
Clark et al., 2014	Accessing primary health care				Understanding gps and pharmacists
Cyril et al., 2017	Childhood obesity prevention				
Filippi et al., 2014	Access to health care, information, and treatment				
Floyd et al., 2017	Accessing health care				
Gele et al., 2017	Participation in cervical cancer screening				
Gregory, 2015				Communicating gastrointestinal illness concerns	

Included Studies	Health Service Use or Behaviour	Health Outcome	Health Cost	Inductively Identified Theme	
				Health Experience and Perception	Health Knowledge and Understanding
Groenenberg et al., 2015	Likelihood of participation in the Dutch two-stage cardiometabolic health check (health risk assessment and prevention consultations)				
Hurley et al., 2013	Community-based services use				
Jafri, 2012	Participation in breast cancer screening				
Jung et al., 2017	Colorectal cancer screening				
F. H. Lee et al., 2014	Receiving cervical screening				
Marshall et al., 2010				Unmet healthcare needs	
Murry et al., 2018	Medication management				

Included Studies	Health Service Use or Behaviour	Health Outcome	Health Cost	Inductively Identified Theme	
				Health Experience and Perception	Health Knowledge and Understanding
Oliver, 2015	Limited engagement in preventive health behaviors; challenges with post-resettlement health care system				
Renzaho et al., 2017	Ability to get involved in childhood obesity prevention				
Schoenmakers et al., 2017				Immigrant social network not supporting immigrants to recognize depression symptoms	
Shedlin et al., 2004					Knowledge and understanding of HIV/AIDS
Smaland Goth et al., 2011					Understanding the health system
Sriphanlop et al., 2014	Attaining preventative services and screenings				

Included Studies	Health Service Use or Behaviour	Health Outcome	Health Cost	Inductively Identified Theme	
				Health Experience and Perception	Health Knowledge and Understanding
Thomson et al., 2012				Needs and views in diet-related cancer prevention	
Todd et al., 2011c				Sources of cancer information; barriers to cancer information seeking; strategies used during cancer information seeking	
Watts et al., 2014				Attitudes towards and use of contraception	
Woudstra et al., 2016	Informed participation in CRC screening				

Appendix U. Summary of key findings regarding antecedents of health literacy or numeracy from the included studies

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
Quantitative Studies			
Becerra et al., 2017	(a) living in poverty; (b) lacking consistent health insurance; (c) limited English language proficiency; (d) gender (women)	None	Multivariable logistic regression analysis showed that in immigrant Hispanics to the US, low health literacy was associated with living in poverty (Odds Ratio (OR) = 1.63, 95% CI (1.18, 2.26), p = 0.003), lacking consistent health insurance (OR = 1.40, 95% CI (1.08, 1.82), p = 0.012), and limited English language proficiency (OR = 3.22, 95% CI (2.50, 4.14), p < 0.001), while women were less likely than men (OR = 0.59, 95% CI (0.47, 0.75), p < 0.001) to report low health literacy. Adjusted covariates included age and education.
Beltran et al., 2016	(a) participants' level of education; (b) cervical cancer screening literacy; (c) number of doctor visits	Marital status; years in US; health insurance; having a usual source of care; annual health check-up; family cancer history	Multiple regression analysis showed that three variables significantly contributing to the prediction of HPV literacy among Hmong American immigrants were participants' level of education (the unstandardized beta (B) = 0.229, the standardized beta (β) = 0.185, p < 0.05), cervical cancer screening literacy (B = 0.210, β = 0.173; p < 0.05), and number of doctor visits (B = 0.225, β = 0.188, p < 0.05). Adjusted covariates included marital status, number of years in the U.S., having health insurance, having a usual source of care, having an annual health check-up, and family cancer history.

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
S. E. Choi et al., 2013	(a) income; (b) years of education; (c) English proficiency; (d) acculturation; (e) age	Gender; health behaviours including fruit intake, vegetable intake, soda intake, activity	Bivariate correlation analysis showed that Korean immigrants to the US with higher income ($r = 0.23$, $p < 0.01$), more years of education ($r = 0.28$, $p < 0.01$), more English proficiency ($r = 0.41$, $p < 0.001$), higher acculturation ($r = 0.41$, $p < 0.001$), and younger age ($r = -0.46$, $p < 0.001$) had higher dietary health literacy.
Y. J. Choi et al., 2016	Levels of acculturation stress	None	Bivariate correlation analysis showed that among immigrant women to South Korea, as levels of acculturation stress increased, levels of mental health literacy (Pearson $r = 0.137$, $p = 0.048$) decreased.
Diamond et al., 2014	None	Duration of residence	Duration of residence did not greatly impact dementia literacy (the understanding of dementia) between Chinese immigrants to the US with a < 20 -year versus a ≥ 20 -year duration of residence.

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
<p>Gatobu et al., 2014</p>	<p>1. Predictors of numeracy (tools used: the French kit addition task and the Mathematics Self-Efficacy Scale (MSES)): (a) self-efficacy, (b) primary language, (c) residency in Canada; 2. Predictors of numeracy (tools used: the French kit addition task and the Subjective Numeracy Scale (SNS)): (d) self-efficacy, (e) primary language, (f) residency in Canada; 3. Predictors of numeracy (the French Kit addition, subtraction-correction task and MSES): (g) self-efficacy, (h) primary language, (i) education-University; 4. Predictor of numeracy (the French kit addition, subtraction-correction task and SNS): (j) self-efficacy, (k) primary language, (l) education-University; 5. Predictors of numeracy (the S-TOFHLA numeracy and MSES): (m) format (numbers vs. numbers with explanation); 6 Predictors of numeracy (the S-TOFHLA numeracy and SNS): (n) self-efficacy, (o) format.</p>	<p>1. gender, education; 2. gender, education; 3. gender, residency in Canada, education - >High School <University, education-Graduate; 4. gender, residency in Canada, education- >High School <University, Education-Graduate; 5. gender, residency in Canada, primary language, education, self-efficacy; 6. gender, residency in Canada, primary language, education</p>	<p>1. In the multiple regression model where mathematical self-efficacy was measured by the MSES and numeracy was measured by the French Kit Addition Task in the immigrants from China and Kenya to Canada: (a) self-efficacy (β (standard error (SE)) = 0.67 (0.25), $p = 0.008$), (b) primary language (β (SE) = 2.85 (0.90), $p = 0.002$), (c) residency in Canada (β (SE) = 2.83 (1.11), $p = 0.010$); 2. In the model where mathematical self-efficacy was measured by the SNS and numeracy was measured by the French Kit Addition Task: (d) self-efficacy (β (SE) = 1.34 (0.50), $p = 0.008$), (e) primary language (β (SE) = 2.40 (0.95), $p = 0.013$), (f) residency in Canada (β (SE) = 2.73 (1.11), $p = 0.014$); 3. In the model where mathematical self-efficacy was measured by the MSES and numeracy was measured by the French Kit addition, subtraction-correction task: (g) self-efficacy (β (SE) = 1.07 (0.44), $p = 0.015$), (h) primary language (β (SE) = 10.76(1.57), $p < 0.001$), (i) education-University (β (SE) = 9.20 (3.02), $p = 0.003$); 4. In the model where mathematical self-efficacy was measured by the SNS and numeracy was measured by the French Kit addition, subtraction-correction task: (j) self-efficacy (β (SE) = 3.16 (0.84), $p < 0.001$), (k) primary language (β (SE) = 9.31 (1.60), $p < 0.001$), (l) education-University (β (SE) = 7.71 (2.99), $p = 0.011$); 5. In the model where mathematical self-efficacy was measured by the MSES and health numeracy was measured by the S-TOFHLA numeracy component: (m) format (numbers vs. numbers with explanation) (β (SE) = 3.48 (1.50), $p = 0.022$); 6. In the model where mathematical self-efficacy was measured by the SNS and health numeracy was measured by the S-TOFHLA numeracy component: (n) self-efficacy (β (SE) = 1.08 (0.51), $p = 0.038$), (o) format (β (SE) = 3.63 (1.48), $p = 0.015$). Covariates included gender, residency in Canada, primary language, Education: >High School and <University, Education: University, Education: Graduate, and format.</p>

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
Gatobu et al., 2016	a) Predictor of numeracy measured by the French kit: language difference (Mandarin and Kikuyu) b) Predictor of health literacy measured by the S-TOFHLA's reading comprehension component: language difference (Mandarin and Kikuyu)	Mandarin and Kikuyu speakers showed similar performance in the S-TOFHLA's numeracy component and the NVS.	Analyses showed that a) Mandarin speakers outperformed Kikuyu speakers in the addition task (observation-oriented analyses: percent of correct classification (PCC) = 68.91%, $c = 0.24$) and the addition–subtraction task of the French Kit (observation-oriented analyses: PCC = 87.39%, $c = 0.001$); b) the superiority of the Kikuyu speakers in S-TOFHLA's reading comprehension component than the mandarin speakers (observation-oriented analyses: PCC = 73.11%, $c = 0.001$; Mann-Whitney test: $U = 2640$, $p < 3.488e-06$).
Gele et al., 2016	(a) employment; (b) acculturation	Age; education; years in Norway	Multivariate logistic regression analysis showed that unemployment (adjusted OR = 3.66, 95% CI (1.08, 12.3), $p < 0.05$) and poor acculturation (adjusted OR = 8.17, 95% CI (1.21, 54.8), $p < 0.05$) were independent predictors for a limited health literacy among Somali women in Norway. Adjusted covariates included age, education, and years in Norway.
Jacobson et al., 2016	(a) English proficiency for the TOFHLA total health literacy score and numeracy and reading comprehension subscores; (b) age for the TOFHLA numeracy subscores	Associations between age and the TOFHLA total health literacy score and reading comprehension subscores	Multiple linear regression analysis showed that English proficiency is the strongest predictor of TOFHLA scores among Hispanic immigrants to the US, including numeracy and reading comprehension scores. The relationship between English proficiency and the three continuous outcomes, total TOFHLA (changes in the total score for a 1 standard deviation increase in the predictor = 9.49, 95% CI (7.24, 11.73)), weighted numeracy (changes in the numeracy score for a 1 standard deviation increase in the predictor = 3.42, 95% CI (1.83, 5.01)), and reading comprehension (changes in the reading comprehension score for a 1 standard deviation increase in the predictor = 6.04, 95% CI (4.93, 7.15)), remained statistically significant ($p < 0.001$). This study also found that younger participants obtained higher numeracy scores, after adjusting for English proficiency (changes in the numeracy score for a 1 standard deviation increase in the predictor = -1.86, 95% CI (-3.46, -0.26), $p = 0.023$).

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
Khuu et al., 2018	(a) the number of years participants have lived in the US; (b) social or religious group attendance; (c) health status; (d) whether they had difficulties with activities of daily living	Age; marital status; income; usual place of care	Hierarchical regression analysis showed that among Hmong immigrants to the US, health literacy levels were found to differ significantly in the number of years participants have lived in the U.S. ($B = 0.05, \beta = 0.22, p < 0.05$), their social or religious group attendance ($B = 0.58, \beta = 0.18, p < 0.05$), health status ($B = 0.80, \beta = 0.26, p < 0.001$), and whether they had difficulties with activities of daily living ($B = -0.80, \beta = 0.16, p < 0.05$). Adjusted covariates included age, marital status, income, usual place of care.
H. Y. Lee et al., 2012	(a) education; (b) language proficiency; (c) health insurance; (d) gender (women); (e) marital status (married)	Age	Structural equation modeling showed that in Korean American immigrants, education (total effect = 0.20, $p < 0.001$) and language proficiency (total effect = 0.17, $p < 0.001$) were estimated to be the most influential predictors of health literacy. Health insurance (total effect = 0.13, $p < 0.01$), gender (total effect = 0.13, $p < 0.01$), and being married (total effect = -0.03, $p < 0.05$) variables were also found to have modest but significant effects on health literacy. Adjusted covariates included age, education, language proficiency, health insurance, gender (women), marital status (married).
H. Y. Lee et al., 2014	(a) age; (b) educational attainment; (c) having a primary care physician	Sex; marital status; years in the US; English proficiency; monthly income; subjective economic status; healthcare insurance; healthcare motivation; family support; exercise frequency; subjective self-rated health status; number of chronic diseases; family cancer history; smoking frequency	Multiple regression analysis showed that among Korean American immigrants, age (beta = -0.23, $p < 0.05$), educational attainment (beta = 0.17, $p < 0.05$), and having a primary care physician (beta = 0.17, $p < 0.05$) were significant predictors of cancer literacy after controlling other factors including sex, marital status, years in the US, English proficiency; monthly income, subjective economic status, healthcare insurance, healthcare motivation, family support, exercise frequency, subjective self-rated health status, number of chronic diseases, family cancer history, smoking frequency.

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
H. Y. Lee et al., 2015a	(a) gender (female); (b) cultural modesty; (c) educational attainment; (d) English proficiency; (e) having a primary care physician	Age; marital status, health status; number of chronic diseases	Multiple regression analysis showed that among Korean American immigrants, participants' gender (beta = 0.14, p < 0.05), cultural modesty (beta = 0.13, p < 0.05), educational attainment (beta = 0.13, p < 0.05), English proficiency (beta = 0.18, p < 0.01), and having a primary care physician (beta = 0.19, p < 0.01) were found to be significant predictors in determining their health literacy level. Adjusted covariates included gender, age, marital status, and cultural modesty, education attainment, English proficiency, primary care physician, health status, number of chronic diseases.
Mantwill et al., 2017	(a) length of stay; (b) age at the time of immigration	Language dependent health literacy was not significantly associated with length of stay in Serbian-speakers. Language independent health literacy was not significantly associated with age at immigration across all three groups.	Unadjusted regression analysis showed that health literacy as measured by the Chew items (a language-dependent measure, LDM) increased with length of stay (except for Serbian-speakers) and decreased with age at the time of immigration, supporting that the longer an immigrant has been staying in Switzerland, the higher her/his functional health literacy will be as measured by the LDM, while no such association will appear in the S-TOFHLA (a language-independent measures, LIM) and the younger an immigrant was at the time of arrival in Switzerland, the higher her/his functional health literacy will be as measured by the LDM, while no such association will appear in the LIM. [Length of stay with LIM (the S-TOFHLA): in Albanian-speakers, $\beta = -0.439$, 95% CI (-0.707, -0.159), $p \leq 0.01$, in Portuguese-speakers $\beta = -0.263$, 95% CI (-0.428, -0.102), $p \leq 0.01$, in Serbian-speakers $\beta = -0.598$, 95% CI (-0.763, -0.435), $p \leq 0.01$; Length of stay with LDM (the Chew items): in Albanian-speakers, $\beta = 0.055$, 95% CI (0.015, 0.097), $p \leq 0.01$, in Portuguese-speakers $\beta = 0.071$, 95% CI (0.041, 0.099), $p \leq 0.001$; Age at immigration with LDM (the Chew items): in Albanian-speakers, $\beta = -0.153$, 95% CI (-0.182, -0.125), $p \leq 0.001$, in Portuguese-speakers $\beta = -0.110$, 95% CI (-0.142, -0.078), $p \leq 0.001$, in Serbian-speakers $\beta = -0.105$, 95% CI (-0.135, -0.070), $p \leq 0.001$]

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
Ng et al., 2014	(a) literacy practices at work; (b) own education; (c) maternal education; (d) participation in adult education; (e) household income; (f) literacy practice at home	Immigration subgroups (established, recent, recent European and American vs. established European and American), age, sex, similar home and survey languages, employment status, Census Metropolitan Area	Multiple logistic regression showed that the literacy practices at home (OR = 5.05, 95% CI (1.88, 13.58)), literacy practice at work (OR = 2.29, 95% CI (1.42, 3.68)), as well as their own (OR = 2.76, 95% CI (1.59, 4.80)) and maternal education (OR = 2.48, 95% CI (1.71, 3.59)), participation in adult education (OR = 1.57, 95% CI (1.11, 1.22)) and household income (OR = 1.82, 95% CI (1.18, 2.81)) had substantial and significant effects on the odds of having adequate health literacy in immigrants to Canada. Covariates included socio-demographic factors (such as age, sex, immigrant subgroups, literacy and educational factors including literacy practices at home, literacy practice at work, own and maternal education, participation in adult education), socioeconomic factors including Employment status, Census Metropolitan Area, Household income.
Thomson et al., 2011	Predictors of health literacy (measured by S-TOFHLA representing comprehension of general health information): (a) age, (b) acculturation, (c) media-Internet use, (d) media-television use (hours of Internet and television use in English); Predictors of health literacy (measured by REALM representing comprehension of general health information): (e) acculturation, (f) media-television use; Predictors of health literacy (measured by a modified Cloze procedure representing comprehension of colon cancer information): (g) acculturation, (h) years of Spanish education	Health literacy (measured by S-TOFHLA): years of Spanish education, employment, duration of residency; Health literacy (measured by REALM): age, years of Spanish education, employment, duration of residency, media-Internet use; Health literacy (measured by a modified Cloze procedure): age, employment, duration of residency, media-internet use.	Multiple regression showed that among immigrant women to Canada, age ($\beta = -0.34$, 95% CI (-0.57, -0.11), $p = 0.00$), acculturation ($\beta = 5.10$, 95% CI (2.38, 7.81), $p = 0.00$), Internet ($\beta = 1.32$, 95% CI (0.23, 2.41), $p = 0.02$) and television ($\beta = -1.25$, 95% CI (-2.01, -0.49), $p = 0.00$) use predicted health literacy via S-TOFHLA. Acculturation ($\beta = 2.93$, 95% CI (0.21, 5.65), $p = 0.04$) and television use ($\beta = -0.67$, 95% CI (-1.25, -0.01), $p = 0.04$) predicted health literacy via REALM. Acculturation ($\beta = 2.04$, 95% CI (0.64, 3.44), $p = 0.01$) and years of Spanish education ($\beta = 0.36$, 95% CI (0.09, 0.63), $p = 0.01$) predicted health literacy via the Cloze test. Covariates included age, acculturation, years of Spanish education, employment, duration of residency, media-internet use, and media-television use.

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
Todd et al., 2011b	Predictors of health literacy measured by S-TOFHLA: (a) acculturation, (b) age, (c) self-reported proficiency reading English, (d) education; Predictors of health literacy measured by the colon cancer Cloze test (English and Chinese versions): (e) acculturation for the English version of the test, (f) education for the Chinese version of the test	Age, education, and English proficiency were not significantly associated with health literacy measured by the colon cancer Cloze test (English version). Age, acculturation, and English proficiency were not significantly associated with health literacy measured by the colon cancer Cloze test (Chinese version).	The regression analysis showed that health literacy, as measured by S-TOFHLA, was significantly predicted by acculturation ($\beta = 0.438$, 95% CI (5.219, 11.974), $p = 0.0001$), age ($\beta = -0.215$, 95% CI (-0.445, -0.109), $p = 0.001$), self-reported proficiency reading English ($\beta = 0.250$, 95% CI (1.744, 9.457), $p = 0.005$), and education ($\beta = 0.202$, 95% CI (1.599, 7.497), $p = 0.003$); these factors accounted for 56.6% of the variation in scores among Chinese immigrant women to Canada. Health literacy, reflected by scores on the colon cancer Cloze test, was significantly predicted by acculturation ($\beta = 0.514$, 95% CI (7.787, 27.903), $p = 0.001$) for the English version of the test and by education ($\beta = 0.337$, 95% CI (3.590, 19.703), $p = 0.005$) for the Chinese version of the test.
Wangdahl et al., 2014	Predictors of functional health literacy (a) low education (b) being born in Somalia; Predictors of comprehensive health literacy (c) having inadequate functional health literacy	Predictors of functional health literacy: sex, age, years of resident permit, long term sickness, country of origin except Somalia. Predictors of comprehensive health literacy: sex, age, education, years of resident permit, long term sickness, country of origin	Multivariate analysis showed that among refugees in Sweden, low education (OR = 2.25, 95% CI (1.20, 4.20), $p < 0.05$) and being born in Somalia (OR = 2.89, 95% CI (1.28, 6.53), $p < 0.05$) were factors associated with an increased risk of having inadequate functional health literacy. Having inadequate functional health literacy was associated with an increased risk of having inadequate comprehensive health literacy (OR = 3.97, 95% CI (1.23, 12.89), $p < 0.05$). Covariates included sex, age, education, years of resident permit, long term sickness, country functional health literacy.
Wister et al., 2010	None	Years of immigration	The association between years of immigration to Canada and health literacy were not found to be statistically significant in the fully adjusted logistic regression model.

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
Qualitative Studies			
Carroll et al., 2007	Community-based health education programs	Not applicable	Theme "Need for Somali health care workers and health education programs" – A promising and potentially effective venue for improving health literacy raised by Somali refugee women in the US was to develop community-based health education programs for African women.
Leung et al., 2014	(a) high regard for authority; (b) a desire to avoid being burdensome to others; (c) a desire to be together or follow a collective approach; (d) insurance; (e) transportation issues; (f) limited information in Chinese-speaking communities; (g) namely unawareness of self-care responsibility; (h) age related limitations	Not applicable	Eight key themes were found to potentially affect different components of health literacy among Chinese immigrants to the US with diabetes: cultural factors (three themes), structural barriers (three themes), and personal barriers (two themes). Three cultural factors, namely high regard for authority, a desire to avoid being burdensome to others, and a desire to be together or follow a collective approach, were identified. Beliefs or perceptions among participants seemed consistent across interviews and were grounded in Chinese culture. Structural barriers, such as insurance, transportation issues and limited information in Chinese-speaking communities, were identified.
McMichael et al., 2009	Experiences prior to arrival at destination country	Not applicable	Theme "Sources of Information Prior to Arrival" – Common experiences prior to arrival to Australia as a refugee included disrupted schooling, long periods of time living in refugee camps, limited or no access to primary health care, experiences or threats of sexual violence and fragmented family life, all of which impacted on sexual health literacy.

Included Studies	Antecedents Statistically Significant	Antecedents Not Significant	Narrative Summary (or Quotes)
Murry et al., 2018	(a) involving the person's support network; (b) providing instruction not only in written format	Not applicable	Theme "involve the person's support network" and "less reliance on written instruction" were two solutions proposed by former refugees to build individual functional health literacy for medication management.
Taiwo, 2013	(a) religio-cultural beliefs and assumptions on health; (b) challenges with language and communication issues, (c) affordability concerns to healthcare literacy and access care	Not applicable	Theme "Barriers to health literacy" – Consistencies in experiences and expectations were the Nigerian immigrant's religio-cultural beliefs and assumptions on health, their challenges with language and communication issues in a high-tech environment like in U.S., and the affordability concerns to healthcare literacy and access care.

Appendix V. Summary of key findings regarding consequences of health literacy or numeracy from the included studies

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Quantitative Studies			
Bekker et al., 2004	Subscales of Medical Outcomes Study-Short Form 36 (MOS-SF36): (a) physical functioning; (b) general health perception	Role limitation due to physical problem; pain; vitality; social functioning; role limitation due to emotional problem; mental health	Analysis of variance (ANOVA) showed that within first-generation Berber women to the Netherlands, illiterates, compared with literates, scored significantly lower on the MOS-SF36 Physical functioning ($F(1, 50) = 5.79, p < 0.05$) and General health perception ($F(1, 50) = 7.06, p < 0.05$). Adjusted covariates included number of children, age, and own and partner's employment status.
Borges et al., 2017	Relapse rate in children with nephrotic syndrome	None	The scatterplot showed an inverse correlation of the S-TOFHLA overall health literacy score with relapse rate by immigration status. The line of best fit was steeper among those born outside of Canada than those born in Canada. Numeracy subscore was not associated with outcomes in the total population including both immigrants and non-immigrants.
Calvasina et al., 2016	(a) not visiting a dentist in the preceding year; (b) not having a dentist as the primary source of dental information; (c) not participating in shared dental treatment decision making	None	Among Brazilian immigrants to Canada, inadequate/marginal oral health literacy was associated with not visiting a dentist in the preceding year ($OR = 3.61; p = 0.04$), not having a dentist as the primary source of dental information ($OR = 5.55; p < 0.01$), and not participating in shared dental treatment decision making ($OR = 1.06; p = 0.05$) in multivariate logistic regressions controlling for covariates including age, time living in Canada, self-rated oral health, oral health knowledge, oral health-related quality of life, medical health literacy, oral health self-efficacy, access to dental care.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Calvo, 2016	Reporting high Quality of Care (QoC)	None	Multivariate logistic regression showed that among Latino immigrants to the US, health literacy (assessed by the Chew items) was significantly associated with QoC even after controlling for immigrants' education and income, English proficiency, health insurance coverage, and regular place of care. To be specific, the associations between reporting high QoC and the Chew item1 (read hospital materials), item2 (filling out medical forms alone), and item3 (understanding written medical information) were [OR 1.23, 95% CI (0.99, 1.53), $p < 0.05$], [OR 1.56, 95% CI (1.25, 1.94), $p < 0.001$], and [OR 1.39, 95% CI (1.11, 1.73), $p < 0.01$], respectively.
S. E. Choi et al., 2013	(a) waist to hip ratio; (b) blood glucose level	BMI; systolic blood pressure; diastolic blood pressure	Bivariate correlation analysis showed that higher dietary health literacy was associated with lower waist to hip ratio ($r = -0.19$, $p < 0.05$). Higher dietary health literacy was associated with lower blood glucose level only among the Korean immigrants to the US who were above the mean score on the acculturation scale ($r = -0.39$, $p < 0.001$), suggesting that diet health literacy may have different implications for highly acculturated individuals than for less acculturated individuals.
Y. J. Choi et al., 2016	Levels of mental health	None	Bivariate correlation analysis showed that among immigrant women to South Korea, as levels of mental health literacy increased, levels of mental health (Pearson $r = -0.262$, $p < 0.005$) increased.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Coffman et al., 2007	Health care service use	None	Multiple regression analysis showed that among Latino immigrants to the US, health care service use was significantly associated with health literacy. Although health literacy alone was not a predictor, the interaction between income and health literacy was a significant predictor of health care resources ($\beta = -0.004$, $F [1, 83] = 6.11$, $p = .015$). Participants with lower health literacy used more health care resources as incomes increased. Adjusted covariates included age, illness, income, education, health literacy, health status, age*education, illness*education, income*education.
Coffman et al., 2010	Depression score of the 20-item Spanish Language Center for Epidemiologic Studies Depression Scale	None	Regression analysis showed that lower health literacy levels ($\beta = -0.2$; $p = 0.048$) predicted higher depression scores among Latino immigrants to the US, after controlling demands of immigration.
Coffman et al., 2012	Health care use	None	Multiple regression analysis showed that health literacy was significantly associated with health care use ($B = 0.02$, $\beta = 0.18$; $p = 0.03$) among Latino immigrants to the US. Adjusted covariates included glucose level, diabetes knowledge.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Geltman et al., 2013	(a) decayed, missing, and filled teeth; (b) risk of periodontal disease	Untreated decay	In the adjusted multivariate analysis, among Somali refugees living in the US for 0 to 4 years, decayed, missing, and filled teeth count (lifetime history of dental decay) was 1.3 times as high among participants with high as those with low S-TOFHLA scores (for the newer immigrants with low literacy, adjusted rate ratio (RR) = 0.78; p = 0.02). By contrast, among participants living in the US for 5 to 10 years, those with low S-TOFHLA scores had 1.4 times the lifetime history of dental decay of those with higher scores (adjusted RR = 1.37; p = 0.01). Adjusted covariates included acculturation, age, sex, ethnicity, education, income, dental insurance, dental exam in past, treatment in past, brush teeth, use stick brush, Posttraumatic Stress Disorder Checklist–Civilian Version, Medical Outcomes Study Short-Form 12-item survey. In adjusted analyses, participants with low S-TOFHLA scores had 0.22 the odds of periodontal disease as those with higher S-TOFHLA scores (adjusted OR = 0.22, p = 0.047). Adjusted covariates included acculturation, sex, education, years in US, brush teeth, Oral Health Quality of Life Measure.
Geltman et al., 2014	None	Preventive dental care	Controlling for acculturation in multivariate analysis, English functional health literacy was not associated with preventive dental care in Somali refugees to the US.
Hernandez-Mekonnen et al., 2016	Child risk for developmental delay	Use of early intervention services	Regression analysis showed that Mexican immigrant women to the US with low maternal health literacy (scoring <4 on the PHLAT) had an increased odd of having a child at risk for developmental delay even after adjusting for a mother’s total number of children (adjusted OR = 4.4; 95 % CI (1.3, 15.4), p value not reported).

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Idehen et al., 2017	Likelihood of Pap test participation (among Russians)	Among Somalis or Kurds, literacy in Finnish/Swedish was not associated with likelihood of screening participation.	Multiple logistic regression analysis showed that among Russians to Finland, one of the significant factors associated with higher likelihood of screening participation was literacy in Finnish/Swedish (OR = 3.80, 95% CI (1.60, 9.01), p = 0.002). Adjusted covariates included socio-demographic (age, high school in any country, married/common law partnership, number of family members, employment status, living outside of the metropolitan area of Finland), immigration category (age of migration, living in Finland for 10 years and longer, Has experienced any discrimination), health care services category (at least one general health check-up (previous five years), at least one gynecological check-up (previous five years), distrust in public healthcare services), health category (very good or good self-rated health status, has ever given birth, has experienced problems due to female genital mutilation).
Igarashi et al., 2013	(a) Quality of Care for Pregnancy, Delivery and Postpartum Questionnaire scores, including Respect (feeling respected), Understanding (feeling being understood by health care providers) and Cold (feeling distant and unhelpful attitude) subscores; (b) care satisfaction	None	Regression showed that among immigrant women to Japan, there were statistically significant correlations (all p < 0.01) between literacy level and all Quality of Care for Pregnancy, Delivery and Postpartum Questionnaire subscores or between literacy and care satisfaction. [During pregnancy: Respect subscores r = -0.397, Understanding subscores r = -0.349, Cold subscores r = 0.459, Care satisfaction r = -0.087; During Delivery: Respect subscores r = -0.194, Understanding subscores r = -0.477, Cold subscores r = 0.220, Care satisfaction r = -0.105; During Postpartum: Respect subscores r = -0.353, Understanding subscores r = -0.425, Cold subscores r = 0.209, Care satisfaction r = -0.107]

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Kankou et al., 2017	HIV disclosure to a steady partner	HIV disclosure to relatives and to others	Multivariable logistic regression analysis showed that among immigrants from sub-Saharan countries to France, a higher literacy level (adjusted OR = 5.7; 95 % CI (2.0, 16.1), p not reported) was associated with HIV disclosure to one's steady partner. Adjusted covariates included gender, age at the time of the study (<41 vs. >41 years), follow-up duration since HIV diagnosis (<5 vs. >5 years), follow-up duration since ART initiation (<5 vs. >5 years), AIDS status, CD4 cell count (<350 vs. >350/mm ³), educational level (<secondary school vs. >secondary school), housing in France (owner or renter vs. living with others/unstable housing), employment (formal or informal sector vs. unemployed), marital status (married/living in a couple vs. single, divorced /separate /widower), region of HIV care (Paris area vs. outside Paris area), region of birth (West Africa vs. Central Africa/other).
Kim et al., 2013	None	Depression	Hierarchical multiple regression analysis showed that among immigrant women to Korea, the association between language literacy and depression was not statistically significant.
Kim et al., 2018	Pap test use (indirect association via mediators including (a) decisional balance and (b) self-efficacy)	Pap test use (direct association and indirect association via knowledge)	Structural equation modeling tested an indirect relationship between health literacy and Pap test use through the selected mediators. The model resulted in similar goodness-of-fit indices with the data: CFI = 0.934, RMSEA = 0.055, 90%CI (0.045, 0.065), TLI = 0.909, and $\chi^2/df = 158/59$. Among Korean American women, the indirect effect of health literacy on Pap test use was small, but significant, with a standardized path coefficient of 0.125 (p < 0.001). Decisional balance was the strongest mediator between HL and Pap test use, followed by self-efficacy (standardized path coefficients: 0.043 (p = 0.01) and 0.037 (p = 0.02), respectively).

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Ko, 2014	Having a sigmoidoscopy	Having a colonoscopy; having a fecal occult blood test	Bivariate correlation analysis showed that having a sigmoidoscopy was significantly associated colorectal cancer literacy ($r = 0.14$, $p = 0.026$) in Korean American immigrants.
Koch-Weser et al., 2006	None	Self-rated health	After taking age, sex, and disability into account in the multivariate analysis, none of the language and literacy variables were independent predictors of self-reported health in refugees from Southeast Asia to the US.
H. Y. Lee et al., 2015b	(a) self-reported health status; (b) depression symptoms	Health literacy was not associated with self-reported health status in Filipino, South Asian, and Vietnamese subgroups. Health literacy was not associated with depression symptoms in Chinese, Filipino, and Vietnamese.	Multiple regression analysis showed that health literacy was significantly associated with health status among non-Latino white and aggregated Asian immigrant groups in the US ($\beta = 0.034$, $p < 0.01$). However, when the latter group was disaggregated into five ethnic groups, only the Chinese ($\beta = 0.030$, $p < 0.05$) and Korean ($\beta = 0.057$, $p < 0.05$) groups showed health literacy as a significant predictor of self-rated health status. Health literacy was significantly associated with depression symptoms among non-Latino white and aggregated Asian immigrant groups ($\beta = -0.054$, $p < 0.01$). However, when the latter group was disaggregated, only the Korean ($\beta = 0.082$, $p < 0.05$) and South Asian ($\beta = 0.114$, $p < 0.05$) groups showed health literacy as a significant factor for depression symptoms. Covariates included sociodemographic factors (age, gender, education, marital status, poverty level rural); health access and health status factors (insurance, disability, frequency of doctor visits, number of diseases, self-rated health status, depression symptoms); immigration factors (limited English proficiency), years in US.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
H. Y. Lee et al., 2015c	None	Cancer screening	Hierarchical Binary Logistic Regression showed that cancer literacy was not significantly associated with cancer-screening among Hmong immigrants to the US.
Mantwill et al., 2017	General health status with language-dependent health literacy	Language-independent health literacy was not significantly associated with health status across all three groups of immigrants.	Multiple linear regression showed that the S-TOFHLA (LIM) was not significantly associated with health status. Instead, the Chew items (LDM) score was significantly associated with health status in all three language groups in Switzerland. [in Albanian-speakers, $\beta = 1.767$, 95% CI (1.039, 2.476), $p \leq 0.01$; in Portuguese-speakers, $\beta = 1.137$, 95% CI (0.501, 1.796), $p \leq 0.01$; in Serbian-speakers, $\beta = 1.980$, 95% CI (1.136, 2.774), $p \leq 0.01$] Covariates included age, gender, education, presence of chronic condition, where participants had spent most of their school years between the age of 6 and 16, length of stay in Switzerland, age when arriving in Switzerland, language dependent health literacy by the Chew items, language-independent health literacy including numeracy by S-TOFHLA.
Njeru et al., 2016	None	Diabetes outcomes	Two-sample t-test showed that there was no association between diabetes literacy scores and diabetes outcome measures among Somali migrants in the US.
Omariba et al., 2011	None	Self-rated health	Multivariate logistic regression analysis showed that high health literacy was not significantly associated with good self-rated health in first-generation immigrants to Canada.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Omariba et al., 2015	None	Disability	Multivariate logistic regression analysis showed that all the literacy- and education-related factors including health literacy were not significantly associated with disability in first-generation immigrants to Canada.
Prins et al., 2015	None	(a) literacy and self-rated health; (b) numeracy and self-rated health	Among Asians, Hispanics and other immigrants to the US, after controlling human capital characteristics (e.g., educational attainment, employment status, income, English proficiency), logistic regression analysis showed that neither literacy nor numeracy was statistically related to self-rated health.
P. C. Smith et al., 2012	(a) follow up with a physician as instructed (frequency); (b) understanding discharge instructions	Filling prescriptions as instructed	ANOVA showed that Spanish-speaking immigrants to the US with inadequate functional health literacy attended follow-up appointments less frequently than patients with marginal or adequate functional health literacy (36% vs 40% vs 59%, $p < 0.001$). Further, patients with low functional health literacy also reported less understanding of emergency department instructions than those with higher functional health literacy (60% vs 82% vs 95% understood instructions, $p < 0.001$).
Thabit et al., 2009	Glycemic control	None	Correlation analysis showed that among immigrants to Ireland, health literacy was found to be inversely correlated with glycemic control, as measured by HbA1c ($r = 0.35$, $p = 0.018$).

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Todd et al., 2011a	Colon cancer screening	Breast cancer screening	Health literacy scores, as determined by S-TOFHLA, were significantly higher in Chinese immigrants to Canada who had ever screened (mean=19.18) than those who had never screened (mean = 13.7; p = 0.038), or in immigrants who were current colon cancer screeners (completed a fecal occult blood test in the past 2 years or a colonoscopy or sigmoidoscopy in past 5 years) (mean=19.45) than non-current screeners (mean = 14.91; p = 0.042).
Tsoh et al., 2016	Self-rated health	None	Multivariable logistic regression analysis showed that among Chinese American immigrants, low written health literacy, compared to adequate health literacy, was significantly associated with poor self-rated health (adjusted OR = 1.64, 95% CI (1.003–2.68), p = 0.04). Covariates included health communication barriers including Spoken English proficiency, Ever needed medical interpreter, Written health literacy; Socio-demographics and acculturation factors including male, age ≥ 65, married/living with partner, education below high school, Employment status, Born in Mainland China, Years living in U.S.; enabling factors including Had health insurance, Annual household income >\$20,000; perceived health needs including Heart disease or stroke, Diabetes, Hypertension, Hyperlipidemia.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Wangdahl et al., 2015	Consequences of low comprehensive health literacy: (a) having experienced poor quality of communication, (b) having received little health care information, (c) not having received any new knowledge, (d) not having received any help	Functional healthy literacy was not associated with experiences of health examination for asylum seekers.	Multivariate logistic regression analysis showed that among refugees in Sweden, having inadequate as compared to sufficient comprehensive health literacy was associated with the experience of a poorer quality of communication (adjusted OR = 9.64, 95% CI (3.25, 28.58), p < 0.001) and the experience of receiving little health care information (adjusted OR = 6.54, 95% CI (2.45, 17.47), p < 0.001). Having inadequate as compared to sufficient comprehensive health literacy was associated with the experience of not having received any new knowledge (adjusted OR = 7.94, 95% CI (3.00, 21.06), p < 0.001) or not having received any help (adjusted OR = 8.07, 95% CI (2.50, 26.07), p < 0.001). Adjusted covariates included sex, age, education, country of origin, support by interpreter, functional health literacy, comprehensive health literacy.
Wangdahl et al., 2018	(a) having reported poor health status; (b) impaired psychological well-being; (c) having refrained from seeking healthcare	None	Multivariate logistic regression analysis showed that among refugees in Sweden, having inadequate as compared to sufficient comprehensive health literacy was associated with having reported poor health (adjusted OR = 2.93, 95% CI (1.58, 5.42), p < 0.01), impaired psychological well-being (adjusted OR = 4.86, 95% CI (2.74, 8.66), p < 0.001), and having refrained from seeking healthcare (adjusted OR = 2.12, 95% CI (1.20–3.73), p < 0.01). Adjusted covariates included sex, age, education, years with a residence permit in Sweden, reason for residence permit, and having participated in the health examination for asylum seekers (HEA), long term illness.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Yunusa Vakkai, 2016	General health status	Literacy and the interaction of literacy and numeracy were not predictors of variance in general health status. Numeracy and literacy were not found to have statistically significant associations with health care access and marketplace utilization	Regression analysis showed that among immigrants to the US, numeracy predicted 9.3% of the variance (Beta = 0.293, t = 5.881, p ≤ .001) of general health status. Covariates included age, gender, literacy, numeracy and their interactions.
Qualitative Studies			
Alzayer et al., 2017	Asthma control	None	Health literacy might affect asthma control in Arabic-speaking migrants in Australia. (Theme “health literacy”, including subthemes: “language barriers” and “health and health system awareness-asthma action plan”, and “no first-aid knowledge”)
Clark et al., 2014	Accessing primary health care and understanding GPs and pharmacists	None	Inadequate literacy may cause difficulties in accessing primary health care and understanding GPs and pharmacists among refugees in Australia. Theme “education and literacy” – Many of the women from Groups 1, 2 and 3 had little or no education, or were not literate in their own languages and had never lived in a town or city before coming to Australia, therefore there were multifactorial causes to the access and understanding barriers.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Cyril et al., 2017	Childhood obesity prevention	None	Culturally and linguistically diverse parents (migrants in Australia) prioritized low levels of health and food literacy as the fourth barrier to childhood obesity prevention. Theme “Health information/health literacy”
Filippi et al., 2014	Access to health care, information, and treatment	None	Literacy was perceived as the main obstacle to health care, information, and treatment. Theme “barrier to access”
Floyd et al., 2017	Accessing health care	None	Theme “isolation” – Language barriers, literacy challenges, and adaptation to their new environment left most of the refugee women in Canada feeling isolated while accessing healthcare in Canada. This isolation was experienced in a number of ways: rejection, fear, and shame.
Gele et al., 2017	Participation in cervical cancer screening	None	Theme “Sociocultural barriers: poor health literacy and economic constraints” – Some of the immigrants to Norway were of the view that they dealt with the economic problems of their families, which took priority over their health problem. They mentioned that they did not feel the need to seek any health care until they felt sick.
Gregory, 2015	Communicating gastrointestinal illness concerns	None	Theme “personal factors: low health care literacy” – Latino American immigrants' perceived barriers to communicating gastrointestinal illness concerns included undocumentedness, fear of deportation, personal healthcare beliefs and practices, lack of insurance, low health care literacy, and limited income.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Groenenberg et al., 2015	Likelihood of participation in the Dutch two-stage cardio-metabolic health check (health risk assessment and prevention consultations)	None	Theme “Awareness factors: health literacy (apply to all groups)” – Immigrants to the Netherlands with little or no education are more ignorant concerning health and disease. Theme “Barriers and ability factors: health illiteracy and language barrier (apply to all groups)” – Difficulty of information provided and translation barriers, leading to mistakes and misunderstandings.
Hurley et al., 2013	Community-based services use	None	Theme “Attitudes to formal services” – When asked about barriers to service use, Greek immigrants to Australia in both interviews and focus groups identified the barriers: language and literacy, costs, transport and the type of services offered.
Jafri, 2012	None	Participation in breast cancer screening	Theme “Health Literacy and Patient-Physician Communication” – The limited participation in breast screening behaviors does not appear to be due to poor health literacy levels or a nonunderstanding of written materials on breast cancer screening offered by their health providers. This study showed that Pakistani immigrants to the US did not fall into the path of barriers to breast cancer screening caused by the issues of modesty or low health literacy levels, but fell into the path of other barriers posed by components of culture and health communication.
Jung et al., 2017	Colorectal cancer screening	None	Among Chinese and Korean immigrants to the US, cultural barriers to colorectal screening included language (e.g. limited English proficiency and low health literacy).

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
F. H. Lee et al., 2014	Receiving cervical screening	None	Lack of health literacy is one barrier to receiving cervical screening in immigrants from Vietnam to Taiwan, China. Theme “Lack of health literacy”, including unaware of policy and the purposes of cervical screening; unaware of where the service is provided; lack of the concept of preventive care; belief that a cervical screening is not required.
Marshall et al., 2010	Unmet healthcare needs	None	Theme “identifying an unmet need” – immigrants from China and South Asia to Canada had unmet healthcare needs in relation to barriers to accessing care, their lack of health system literacy, and when the health system was less responsive than their expectations.
Murry et al., 2018	Medication management	None	Perceived barriers to managing medications included language barriers, low health literacy, and culturally unresponsive interactions with services.
Oliver, 2015	(a) limited engagement in preventive health behaviors; (b) challenges with post-resettlement health care system	None	Theme “Limited Health Literacy”, including subthemes: “limited engagement in preventive health behaviors among Bhutanese refugees in the US” and “challenges with post-resettlement health care system”

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Renzaho et al., 2017	Ability to get involved in childhood obesity prevention	None	Theme “Obesity literacy and childhood obesity prevention” – Firstly, among immigrants to Australia, parents’ low pre-migratory obesity literacy was negatively impacting their ability to get involved in obesity prevention programs. Additionally, due to the low pre-migration obesity literacy, childhood obesity was not an issue of concern in their home countries especially in comparison to other urgent health conditions such as infectious diseases. Secondly, parents did not recognize obesity as a disease and consequently could not identify obesity as a major issue in their community. Thirdly, parents did not believe that unhealthy eating was linked to childhood obesity and viewed that children required a lot of food for their growth and development and unhealthy eating was part of this phase. Finally, parents stated that the concept of obesity prevention was alien to them.
Schoenmakers et al., 2017	Immigrant social network not supporting immigrants to recognize depression symptoms	None	Theme “The social network does not recognize mental health problems” – The social network could often not support elderly migrants to the Netherlands to recognize symptoms of depression, because they also missed the ‘mental health literacy’ to do so.
Shedlin et al., 2004	Knowledge and understanding of HIV/AIDS	None	Theme “Knowledge of and experience with HIV/AIDS” – The lack of education and low literacy levels of many Latino immigrants to the US, especially from Mexico and Central America, was acknowledged as an obstacle to knowledge and understanding.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Smaland Goth et al., 2011	Understanding the health system	None	Theme “Factors enhancing participation” – Some migrant leaders in Norway mentioned that a generally low level of medical literacy could also pose difficulties for understanding the system.
Sriphanlop et al., 2014	Attaining preventative services and screenings for hepatitis B	None	Theme “Language and health literacy” – Health literacy and language were mentioned as barriers that hindered African immigrants to the US from attaining preventative services and screenings for hepatitis B.
Thomson et al., 2012	Needs and views in diet-related cancer prevention.	None	Immigrants to Canada with functional health literacy and interactive health literacy groups have different needs and views in diet-related cancer prevention. Four themes were identified: 1. “general information requests” – low functional health literacy (FHL) (n = 23/43): low FHL focused on basic requests for information. 2. “specific nutrition inquiries” – high FHL (n = 17/43): women with higher FHL expressed frustration adjusting to Canadian food habits. 3. “actions for healthy eating” – low interactive health literacy (IHL) (n = 8/43): low IHL described they used nutrition information to benefit themselves and family members. 4. “community communication issues” – high IHL (n = 3/43): they were interested in discussing recommendations for educating community members regarding diet-related colon cancer prevention.

Included Studies	Consequences Statistically Significant	Consequences Not Significant	Narrative Summary (or Quotes)
Todd et al., 2011c	None	Sources of cancer information; barriers to cancer information seeking; strategies used during cancer information seeking	Health literacy did not distinguish the Chinese immigrant women to Canada on any of the major themes, including “Sources of cancer information”, “Barriers to cancer information seeking”, “Strategies used during cancer information seeking.”
Watts et al., 2014	Attitudes towards and use of contraception	None	Attitudes towards and use of contraception among refugees in Australia were influenced by parental sexual health literacy and attitudes, gender roles and culturally informed attitudes around motherhood.
Woudstra et al., 2016	Informed participation in colorectal cancer screening	None	The language barrier and low literacy formed serious barriers to informed participation in colorectal cancer screening among immigrants to the Netherlands.