University of Alberta

Health Literacy and Health Outcomes in Diabetes

by

Fatima Al Sayah

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Public Health

Department of Public Health Sciences

Faculty of Nursing

Fall 2013
Edmonton, Alberta

Abstract

There is a general agreement that a relationship exists between health literacy (HL) and health outcomes. Nonetheless, there are critical gaps in the measurement of HL and in the evidence on the impact of inadequate HL on health outcomes, especially in the diabetes population. These gaps need to be addressed before any recommendations regarding HL screening or interventions are implemented.

To address these gaps, first we conducted two systematic reviews, one in which we reviewed the evidence on the relationship between HL and health outcomes in the diabetes population, and the other consisted of a review and evaluation of HL measures used in this population. Then we conducted a validation study that examined the measurement properties of a HL measure; a longitudinal study that examined the associations between HL and health outcomes in individuals with diabetes; and last, a qualitative study that examined the use of interactive communication loops and medical jargon in relation to HL in nurses' interaction with individuals with diabetes.

We found that the evidence on the impact of HL on health outcomes in the diabetes population is limited and inconclusive; measures of HL are not comprehensive enough with limited evidence on their measurement properties; the 3-brief screening questions are potentially a useful measure for screening for inadequate HL; inadequate HL was not associated with worse health outcomes in individuals with diabetes and depressive symptoms; and healthcare providers may place high demands on patients through their communication and interaction with them.

Despite the use of rigorous research methods and the robust evidence generated, the overall available evidence on these relationships is still inconsistent and thus inconclusive. Our work highlights two crucial questions that need to be examined “how to comprehensively measure HL?” and “whether HL is modifiable?” Until, these questions – and others – are answered and conclusive evidence is available, we believe that, outside of the study setting, it might be premature to invest in routinely screening for HL or to trying to improve HL for the purposes of improving patient-related outcomes in diabetes, although there might be other reasons to do so.

Acknowledgements

There are a number of people without whom this thesis might not have been written, and to whom I am greatly indebted.

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Jeff Johnson, for his exceptional guidance, caring, patience, and providing me with an excellent atmosphere for doing research. Without his supervision and constant support, this dissertation would not have been possible.

I would like also to thank the rest of my thesis committee: Dr. Bev Williams and Dr. Sumit Majumdar, for their continuous support, encouragement, guidance, insightful comments, and hard questions; Dr. Scot Simpson for kindly serving as an examiner during both my candidacy and final exams; and Dr. Leonard Egede for serving as an examiner on my candidacy exam and for providing research data for one of the main projects in this dissertation.

In addition to all of these people who contributed to my learning and growth, I'd like to thank Dr. Yutaka Yasui for not only teaching me everything I know about biostatistics, but also for always opening his door for me and listening and bearing the brunt of my frustration and rages against the world.

In my daily work I have been blessed with a friendly and cheerful group of fellow trainees at ACHORD, students at the School of Public Health and Faculty of Nursing, as well as administrative and support staff, and students who helped in a number of projects; many thanks to all of you for making the last five years a great experience.

And finally, many thanks to my parents "Hussein & Souraya", my brother & his wife "Mohamad & Amal", my sisters "Suha & Nisreen", my best friends "Najla & Hisham", and numerous friends who endured this long process with me, always offering support and love.

Table of Contents

List of Tables

Tables numbers prefixed by S indicate chapter supplements

List of Figures and Boxes

Figures

Text Boxes

List of Abbreviations

16-SQ	16 screening questions of inadequate health literacy
3-SQ	3 brief screening questions of inadequate health literacy
6-SQ	6 brief screening questions of inadequate health literacy
A1c	Glycated hemoglobin
BMI	Body mass index
DKQ	Diabetes Knowledge Questionnaire
DOR	Diagnostic Odds Ratio
EQ-5D	EuroQol Group health status index 5-Dimensions
HL	Health Literacy
HRQL	health-related quality of life
IALS	International Adult Literacy and Skills survey
LAD	Literacy Assessment for Diabetes
LDL	Low Density Lipoprotein
LOCF	last observation carried forward
LR-	Negative Likelihood Ratio
LR+	Positive Likelihood Ratio
MART	Medical Achievement Reading Test
MCS	mental component score
NAAL	National Assessment of Adult Literacy
NRI	Net Reclassification Index
PCN	primary care network
PCS	physical component score
PDSMS	Perceived Diabetes Self-Management Scale
PHQ-9	Patient Health Questionnaire-9

REALM	Rapid Assessment of Adult Literacy in Medicine
RMSEA	root mean square error of approximation
ROC	Receiver Operating Curve
s-TOFHLA	Short form – Test of Functional Health Literacy in Adults
SAHLSA	Short Assessment of Health Literacy for Spanish speaking Adults
SBP	systolic blood pressure
SEM	Structural Equation Modeling
SF-12	Medical Outcomes Study 12-item Short-Form Health Survey
SILS	Single Items Literacy Screener
SMBG	Self-monitoring of Blood Glucose
T2DM	type 2 diabetes
TOFHLA	Test of Functional Health Literacy in Adults

Chapter 1

Introduction

1.1. Overview

Health literacy (HL) is the degree to which individuals can obtain, process, understand, and communicate about health-related information needed to make informed health decisions [1]. According to the International Adult Literacy and Skills survey (IALS), more than half the adults in Canada have inadequate HL [2]. Inadequate HL has been recognized as a stronger predictor of a persons' health than age, income, education level, employment status, and race [3]. It is associated with a wide range of adverse direct and indirect effects on care processes and health outcomes [4, 5], and additional healthcare expenditures that range from $143 to $7,798 per person per year [6]. Additionally, inadequate HL is more common among older adults, ethnic minorities for whom English is a second language, people with low levels of income and education, and people with an already compromised health status [7-9]; the same populations that carry the greatest burden of chronic conditions [3].

In the twentieth century, chronic conditions replaced infectious diseases as the dominant health threat. Diabetes, the 6th leading cause of death in Canada [10], affects around 2.4 million Canadians [11]; the prevalence is expected to reach 2.6 million in 2016 [12]. Diabetes is characterized by a high level of care complexity that requires extensive education and self-care management, where HL is anticipated to play a key role. Inadequate HL is common among individuals with diabetes [2], and is suggested to have adverse effects on health outcomes in these individuals. These include general outcomes such as misunderstandings about medical conditions [13], less compliance with medical instructions, decreased self-management skills [14], poor self-rated health [15], and higher health care costs [16], as well as diabetes-specific outcomes such as worse glycemic control [17, 18] and higher risk of significant hypoglycemia [19]. Despite the growing evidence of the effects of inadequate HL on health outcomes in the diabetes population, the evidence has not been previously systematically reviewed, and thus we do not have a comprehensive understanding of these associations.

Several instruments were developed to directly or indirectly assess HL skills. The most commonly used direct measures are the Test of Functional Health Literacy in Adults (TOFHLA) and its shorter form s-TOFHLA [20, 21], and the Rapid Estimate of Adult Literacy in Medicine (REALM) and its shorter and revised forms (REALM-SF, REALM-R) [20-23]. The most commonly used indirect measures are the Single Item Literacy Screener [24], and the 3 screening questions of inadequate HL [25], which are a brief version of 16 screening questions [26]. The development of these instruments has lagged behind the attention more recently paid to defining and conceptualizing HL, and hence their utility befall questionable. Additionally, there is limited evidence on the measurement properties of these instruments. Thus, we lack an understanding of the measurement scope of these instruments and their psychometric properties.

As the challenges associated with inadequate HL are becoming more evident, it is increasingly apparent that the healthcare system has not evolved to serve those with inadequate HL [27, 28]. Meanwhile, healthcare providers play an important role. Patient-provider communication has been suggested as a potential pathway through which HL might impact health outcomes, particularly in individuals with chronic diseases where self-care is a cornerstone of disease management [29, 30]. A few studies have investigated the communication style of care providers with T2DM patients with inadequate HL [31-33]; however, none of these studies examined whether care providers tailor their communication style to accommodate patient's HL needs. Further, none of these studies involved nurses who usually provide self-care management education to T2DM patients within primary care settings. Therefore, we lack an understanding of whether care providers, and particularly primary care nurses, apply communication strategies to ensure patient understanding, and whether they tailor their communication style to patients' HL needs.

This dissertation encapsulates a body of research intended to address these issues. First, we performed a systematic review to examine the research evidence on the relationship between HL and health outcomes in the diabetes population. Second, we conducted another systematic review to identify and evaluate HL measures that have been used with the diabetes population. Third, using cross-sectional data from a study of predominantly African-American individuals with T2DM, we examined the factor structure of the most commonly used indirect measure of HL, and investigated its measurement properties in comparison to the most commonly used direct measure. Fourth, we examined the longitudinal associations of HL with several health outcomes using data from a clinical trial of 154 Canadians with T2DM who recently screened positive for depression. Finally, and in order to investigate the impact of HL on patient-provider communication, we conducted a study using qualitative research methods and examined the use of interactive communication loops and medical jargon in relation to HL in nurses' communication with T2DM patients in Canadian primary care settings.

1.2. Defining Health Literacy and Health Outcomes

As a relatively new construct, the definition of HL is evolving and has not been consistently applied [1]. First and foremost, an important distinction between literacy and HL is to be made. Literacy refers to a person's ability to read and write (this is referred to as functional literacy) [34]. For example, an illiterate person is someone who cannot use print language to perform activities of daily living such as reading a bus schedule. On the other hand, HL refers to people's ability to understand and use health information to effectively manage their health [35]. HL is broadly defined as a set of skills that people need to function effectively in the healthcare environment [36]. These include the ability to read and understand written text, locate and interpret information in documents, and write or complete forms (functional); the ability to speak and listen effectively and communicate about health-related information (interactive); and the ability to navigate the healthcare system and make appropriate health decisions (critical) [36].

The Canadian Public Health Association defines HL 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" [37]. Other definitions characterize HL as a product of both an individual's capabilities and the demands of the health care system [3, 38]. In this dissertation, we adopted the following definition of HL proposed by Berkman and colleagues: "health literacy is the degree to which individuals can obtain, process, understand, and communicate about health-related information needed to make informed health decisions" [1].

“Health outcomes” has become a widely-used term in health care policy over the past ten years; however, the meaning behind the term differs greatly depending on the user and the context. A health outcome is usually defined as a change in the health of an individual, group of people or population which is attributable to an intervention or series of interventions [39]. The term typically refers to the impact healthcare activities have on people — on their symptoms, ability to do what they want to do, and ultimately on whether they live or die. Health outcomes might include, for example, changes in their self-perceived health status or changes in the distribution of health determinants, or factors that are known to affect health, well-being and quality of life.

Whatever the outcome of interest, the goals of medical care are the maximization of good outcomes and the minimization of poor outcomes. This general approach to defining outcomes may be considered in a framework that facilitates the assessment of the quality of care. For example, the Donabedian framework is a conceptual model for examining health services and evaluating quality of care. According to the model, information about quality of care can be drawn from three categories: “structure,” “process,” and “outcomes” [40].

Structure includes all the factors that affect the context in which care is delivered, process is the sum of all actions that make up healthcare, and outcomes contain all the effects of healthcare on patients or populations including morbidity, mortality, pain, functional status, satisfaction, and costs [41]. Outcomes could be classified into immediate or short-term, intermediate, or long-term outcomes [42]. In this dissertation, we examined HL, which is considered a distal factor, as a component of this structure in relation to morbidity- and health status-related short- and intermediate-term outcomes.

1.3. Scope of the Problem

In 2007, the Canadian Council on Learning published “Health Literacy in Canada: Initial Results from the International Adult Literacy and Skills survey” (IALS) [2], which estimated 55% of Canadians aged 16 to 65 to have inadequate HL [2]. This prevalence is lower than that of the U.S. population reported in a similar survey [2]. Only one in eight adults (12%) over age 65 appears to have adequate HL skills. This is particularly significant since seniors are more likely to have more chronic diseases and to use medication than younger age groups, facing a higher level of health information demands. Additionally, the IALS reported interesting results with respect to the relationship between HL and the prevalence of chronic conditions such as diabetes, hypertension, arthritis, and asthma. The survey reported that the strongest and only significant association was with diabetes; as HL scores increase (i.e. better HL), prevalence of diabetes decreases [2]. A similar, but weaker, correlation was found between HL and the prevalence of hypertension. These findings indicate the importance of HL effects in chronic conditions, particularly those that are influenced by life style and health behaviors.

In 2003, more than 50% of Canadians aged 12 or older reported at least one chronic condition and by age 65, 77% of men and 85% of women had at least one chronic condition [43]. In 2009, almost 6.8% of Canadians were living with diabetes and 36.5% of them reported having two or more other chronic conditions [11]. Depression, for example, commonly affects 10-30% of individuals with diabetes [44], which makes it one of the most common comorbidities in this patient population. One of the ways to address the anticipated escalation in the rates of diabetes and other chronic diseases and the subsequent demands this will place on the health care system is to engage patients in more effective self-management. Self-management includes all of the tasks that an individual must undertake to live well with one or more chronic conditions [45] in which HL plays a crucial role. In order to manage a chronic condition on a day-to-day basis, individuals must be able to understand and assess health information, which often includes a complex medical regimen, plan and make lifestyle adjustments and informed decisions, and understand how to access health care services when necessary. A lack of skill in these areas prevents many patients from engaging in effective self-management, which could have adverse effects on their health.

1.4. Health Literacy and Health Outcomes in the Diabetes Population

Research demonstrates that individuals with inadequate HL often cannot read medication labels accurately, may take medication incorrectly, may not understand consent forms, and generally have difficulty understanding print instructions for follow-up care and reading health advisories or warnings [37]. In a systematic review on HL and health outcomes in different patient groups, DeWalt and colleagues found insufficiently low and inconsistent evidence on the relationship between HL level and health outcomes in different patients groups including those with diabetes [4]. In 2011, the same investigators updated the 2004 review, and reported growing, but yet inconsistent low evidence on these relationships [5].

A thorough review of this literature suggests that individuals with diabetes who have limited HL have worse glycemic control [17, 18], higher rates of retinopathy [18], less comprehension of medication instructions, dosing, timing, and warnings [46], higher risk of significant hypoglycemia [19], poorer disease knowledge [13, 17], and poorer patient-physician communication [33, 47, 48] than those with high HL levels. On the other hand, recent studies did not find a direct relationship between HL level and health outcomes in people with diabetes [49] [50] [51]. Further, recent studies suggest that HL could have an indirect effect on outcomes in people with diabetes through intermediate factors such as self-efficacy and social support [51, 52]. Accordingly, the evidence on the relationships between HL levels and diabetes-related health outcomes is not consistent across studies, and thus the effect of HL on the health of people with diabetes is not clear.

The mechanisms by which HL influences health outcomes and the direct and indirect pathways of how HL affects several components of diabetes care and management have not been comprehensively investigated. A few generic models were developed to explain the pathways linking HL and health outcomes. However, these models are not comprehensive enough and do not accommodate the complexity of diabetes care and management, and thus they are not entirely useful to understand and conceptualize the pathways through which HL could influence the health of people with diabetes. The "Paasche-Orlow and Wolf model" is the most commonly used model that outlines the possible causal pathways between HL and health outcomes [53]. Despite its comprehensiveness, the model did not address the role of health care system-related factors in determining HL skills of individuals. Further, the model did not illustrate all of the potential relationships and interactions between the different pathways.

Based on this basic model and the HL literature, we proposed a model presenting an integrative comprehensive view of how HL might influence health outcomes. The proposed model (Figure 1.1) is comprised of three main pillars, each consisting of several components: (1) individual and health care system characteristics, (2) health care and management, and (3) individual and system level outcomes. This model addresses the bi-directional interactive relationship between the different components within and among the three pillars. It also conceptualizes HL as an intermediary rather than a risk factor and shows the reciprocal interaction between the various components of the model. Additionally, this model aligns with the social cognitive theory framework. Social cognitive theory is an expectancy value theory that focuses on the interaction between the individual and the environment [54]. It emphasizes the reciprocal interplay between self-regulatory and environmental determinants of health behavior such as knowledge, self-efficacy, outcome expectations, health goals, and social and structural facilitators and impediments to taking health actions [55]. From this perspective, the proposed HL model could be viewed as an extension and application of the components of the social cognitive learning theory to the "Paasche-Orlow and Wolf" model of HL. This proposed integrative model was the framework that guided the conceptualization of the projects included in this dissertation.

1.5. Measurement of Health Literacy in the Diabetes Population

Since individuals often read several grade levels lower than the highest grade achieved in school [56], educational attainment cannot be used as a proxy for health literacy; this made the development of HL measures a necessity. As the definitions and conceptualizations of HL have changed over the last decades, they became more representative of the skills needed to function successfully in the current society. Nonetheless, it seems that HL has been in large part limited by progress in developing measurement tools, primarily as a result of lack of a shared conceptual framework of what HL encompasses.

Health literacy is a complicated construct that depends on the individual capacity and skills and the demands posed by society and the health care system [38, 57]. At issue is whether individuals' level of HL would be considered higher or lower based on variation in the complexity of the information they encounter. One could argue that an individual's HL would be higher if health-related materials and communication more universally integrated principals of clear language, making them easier to understand and a closer match to individuals' skill level. We believe that HL of individuals in a given health care context is relative to the complexity and demands of that context, and its measurement should account for the latter. Nonetheless, caution is warranted to ensure that this comprehensive conceptualization of HL is immeasurable with the available tools.

To date, all of the available HL tools measure only individual capacity and skills irrespective of the demands of the society and healthcare system. In particular, these measures are mainly focused on assessing what individuals can read and understand in clinical contexts [38]. Therefore, they are considered surrogate measures of HL under the assumption that all public health and health care systems place similar reading and oral communication demands on individuals. Existing HL measures include various versions of the Rapid Estimate of Adult Literacy in Medicine (REALM), the Test of Functional Health Literacy in

Adults (TOFHLA), Health Activities Literacy Scale (HALS), Newest Vital Sign (NVS), Medical Achievement Reading Test (MART), Literacy Assessment for Diabetes (LAD), and the Short Assessment of Health Literacy for Spanish speaking Adults (SAHLSA). Additionally, several indirect measures of HL were developed and primarily applied in survey-based research such as the 3 brief screening questions of inadequate HL [25, 26] and the 3-level HL scale [58]. The National Assessment of Adult Literacy (NAAL) is considered the most comprehensive indirect measure; however, it is not publically available and thus cannot be used in research or intervention studies.

These instruments vary in their development, structure, measurement scope, and subsequently their psychometric properties. Although these instruments have been used with several patient populations, their usefulness and applicability for people with diabetes remains challenging. The reason is the complexity of tasks and skills that are required by people who have diabetes, and the postulation that the available instruments do not address that complexity and all of the important components of HL altogether. Additionally, the continuous adjustment of the meaning and components of HL makes the available instruments questionable in what they actually measure.

1.6. Health Literacy and Patient-Provider Communication

For a significant period, the focus pertaining to HL has been on identifying deficits of the general public and/or patients [59]. The importance of this focus will not diminish nor should it be depreciated. However, there is an increasing and imperative need to address the demands and complexity of healthcare environments including providers. Rudd has noted: “if we forge a better match between the expectations and processes of the system and the skills of US adults, we can solve the problem of health literacy and reduce barriers to good health" [59].

Healthcare providers are key components of every healthcare system, and medical encounters are a key aspect of patient care. Verifying and evaluating patients' understanding during these encounters is one of the most critical elements of effective communication. This is important for patients with chronic conditions who require intensive self-care education and support, and especially important for those with learning challenges such as individuals with inadequate HL.

On the other hand, and as the healthcare systems and environments have become increasingly complex, accessing health information, navigating the healthcare system, and self-care managing is becoming more challenging especially for those with inadequate HL. These individuals are more likely to have difficulty communicating with their health care providers and following up with self-care instructions due to poor understanding of basic health vocabulary, limited background health knowledge, and trouble assimilating new information and concepts [60]. Incompetent and incomplete patient-provider communication may result in misinformation, misunderstandings and mistakes. Patients with inadequate HL who have chronic diseases, e.g. diabetes, hypertension, or asthma have less understanding of their disease [14], report less adherence with medications, and may not be aware of important treatment side effects or the need for follow-up testing [56, 61].

While a number of studies have investigated several aspects of patient-provider communication, only a few examined whether healthcare providers address patient's HL level and needs, and whether they modify their communication style and the terminology they use accordingly. Nonetheless, these studies only involved family physicians. Since patient-provider interaction is crucial for those with chronic diseases, it is imperative to understand how healthcare professionals who provide self-care education and counseling communicate with these individuals, particularly those with inadequate HL.

1.7. Objectives and Program of Research

Although the literature on health literacy is still in its early stages of development, it is nonetheless a vast and rapidly growing body of knowledge. There is general agreement that a relationship exists between health literacy and health outcomes. Nonetheless, there are critical gaps, especially in the diabetes population, that need to be addressed before any recommendations regarding HL screening or interventions are implemented. In particular, we lack a comprehensive understanding of the available evidence on the associations between HL and health outcomes in the diabetes population, and of the measurement scope and measurement properties of HL measures and their applicability in this population. Additionally, we lack longitudinal studies that provide rigorous evidence on the associations between HL and health outcomes in this patient population. We also lack evidence on the level of complexity of the healthcare system, particularly with respect to provider's communication with diabetes patients with different HL skills and needs within Canadian primary care settings.

This dissertation intends to fill some of these evidence gaps through the following objectives:

- To identify, appraise, and synthesize research evidence on the relationships between HL (functional, interactive, and critical) or numeracy and health outcomes (i.e., knowledge, behavioral and clinical) in people with diabetes.
- To identify instruments used to measure HL in people with diabetes; evaluate their use, measurement scope, and properties; discuss their strengths and weaknesses; and propose the most useful, reliable, and applicable measure for use in research and practice settings.
- To examine the measurement properties of the most commonly used indirect measure of HL (16-Screening Questions and their brief version, the 3-Screening Questions) in greater detail, and identify the best set of items to screen for inadequate HL.
- To explore the longitudinal associations of inadequate HL with health outcomes in patients with T2DM who recently screened positive for depression, including depressive symptoms, health-related quality of life, and cardio-metabolic outcomes.
- To investigate whether primary care nurses addressed all components of the interactive communication loop, particularly with respect to assessing recall and comprehension, and avoided the use of jargon while providing self-management education and counseling to individuals with T2DM, and to explore whether these aspects of nurses' communication are influenced by patient's HL level.

The first two objectives were addressed in two separate systematic reviews of the literature intended to provide the most up-to-date comprehensive review of evidence on the associations between HL and health outcomes in the diabetes population, and the measures used to assess HL in this population. The third objective was achieved in a validation study that examined the factor structure of the 16 screening questions of inadequate HL, through which a set of 6 questions was identified, and then used ROC analysis to compare these different sets with the 3-brief screening questions and the s-TOFHLA. This study was conducted using cross-sectional data from a study of 378 predominantly African-American individuals with T2DM. The fourth objective was accomplished in a longitudinal study that examined the associations between HL and depressive symptoms, health-related quality of life, and cardio-metabolic outcomes including A1c, LDL, and SBP. This study was conducted using baseline, 6-months and 12-months data from a clinical trial of 154 Canadians with T2DM who recently screened positive for depression. The last objective was achieved in a qualitative research study that examined the use of interactive communication loops and medical jargon in relation to HL. This study involved audio-recording interactions of nurses' with T2DM patients in Canadian primary care settings.

Figure 1-1: An Integrative Model of Health Literacy

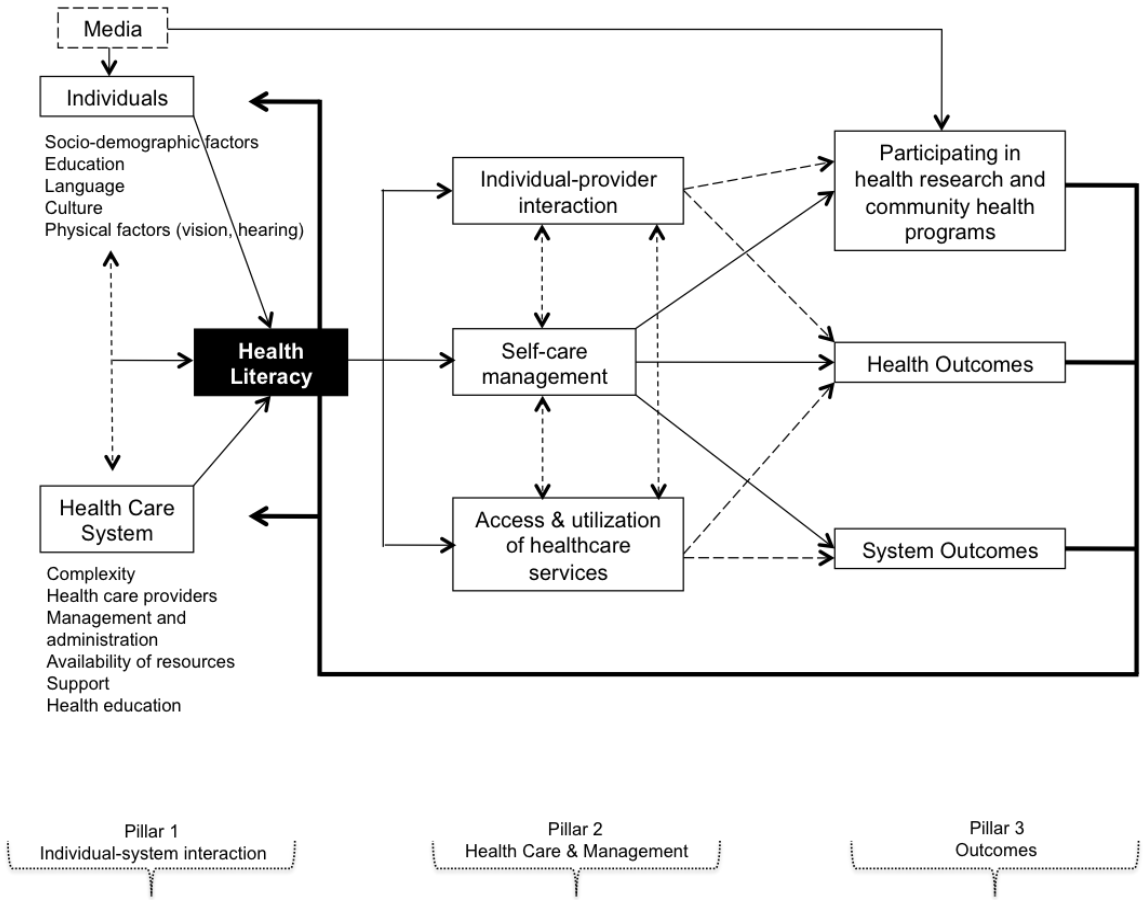


Figure 1: An Integrative Model of Health Literacy

References

1. Berkman N, Terry D, and McCormack L, *Health literacy: What is it?* Journal of Health Communication, 2010. **15**: p. 9-19.

2. Learning, C.C.o., *Health Literacy in Canada: Initial results from the International Adult Literacy and Skills Survey*, 2007: Ottawa.

3. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, A.M.A., *Health literacy: Report on the Council of Scientific Affairs.* JAMA, 1999. **282**: p. 526-527.

4. DeWalt DA, et al., *Literacy and Health Outcomes: A Systematic Review of the Literature.* Journal of General Internal Medicine, 2004 **19**(12): p. 1228–1239.

5. Berkman ND, et al., *Low health literacy and health outcomes: an updated systematic review* Annals of Internal Medicine 2011. **155**(2): p. 97-107.

6. Eichler K, Wiese S, and Brugger U, *The costs of limited health literacy: a systematic review.* Int J Public Health, 2009. **54**: p. 313–324.

7. Wolf M, Gazmararian J, and Baker D, *Health literacy and functional health status among older adults* Archives of Internal Medicine, 2005. **165**: p. 1946-1952.

8. Schultz S and Kopec J, *Impact of chronic conditions. Health Reports.* Statistics Canada catalogue no. 82-003-XIE, 2003. **14**: p. 41-56.

9. Sudore R, et al., *Limited literacy in older people and disparities in health and healthcare access.* Journal of the American Geriatrics Society, 2006. **54**: p. 770-776.

10. Canada, S., *Leading causes of death*, 2012, Statistics Canada.

11. Canada, P.H.A.o., *Diabetes in Canada: Facts and figures from a public health perspective*, 2011, Public Health Agency of Canada

12. Ohinmaa A, et al., *The Projection of Prevalence and Cost of Diabetes in Canada: 2000 to 2016.* Canadian Journal of Diabetes, 2004. **28**.

13. Gazmararian JA, et al., *Health literacy and knowledge of chronic disease.* Patient Education and Counseling 2003. **51**: p. 267-275.

14. Williams M, et al., *Relationship of functional health literacy to patients' knowledge of chronic disease: A study of patients with hypertension and diabetes* Archives of Internal Medicine 1998. **158**: p. 166-172.

15. Bennett IM, et al., *The contribution of health literacy to disparities in self-rated health status and preventive health behaviors in older adults.* Ann Fam Med, 2009. **7**(3): p. 204-11.

16. Howard D, Gazmararian JA, and Parker R, *The impact of low health literacy on the medical costs of Medicare managed care enrollees.* American Journal of Medicine, 2005. **118**: p. 371-377.

17. Powell CK, Hill EG, and Clancy DE, *The relationship between health literacy and diabetes knowledge and readiness to take health actions.* The Diabetes Educator 2007. **33**: p. 144-151.

18. Schillinger D, et al., *Association of health literacy with diabetes outcomes.* JAMA, 2002. **288**: p. 475-482.

19. Sarkar U, et al., *Hypoglycemia is more common among type 2 diabetes patients with limited health literacy: The diabetes Study of Northern California (DISTANCE).* Journal of General Internal Medicine 2010. **25**: p. 962-967.

20. Baker D, et al., *Development of a brief test to measure functional health literacy.* Patient Education and Counseling 1999. **38**: p. 33-42.

21. Parker R, et al., *The test of functional health literacy in adults (TOFHLA): a new instrument for measuring patients' literacy skills* Journal of General Internal Medicine 1995. **10**: p. 537-545.

22. Davis TC, et al., *Rapid assessment of literacy levels of adult primary care patients* Family Medicine 1991. **23**: p. 433-435.

23. Davis TC, et al., *Rapid estimate of adult literacy in medicine: a shortened screening instrument.* Family Medicine, 1993. **25**: p. 391-395.

24. Morris NS, et al., *The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability.* BMC Family Practice 2006. **7**.

25. Chew LD, et al., *Validation of screening questions for limited health literacy in a large VA outpatient population.* Journal of General Internal Medicine 2008. **23**: p. 561-566.

26. Chew LD, Bradley KA, and Boyko EJ, *Brief questions to identify patients with inadequate health literacy.* Family Medicine 2004. **36**: p. 588–594.

27. Paasche-Orlow MK, et al., *How health care systems can begin to address the challenge of limited literacy.* J Gen Intern Med 2006. **21**(8): p. 884-7.

28. Volandes AE and Paasche-Orlow MK, *Health literacy, health inequality and a just healthcare system.* Am J Bioeth, 2007. **7**(11): p. 5-10.

29. Hironaka LK and Passche-Orlow MK, *The implications of health literacy on patient-provider communication.* Arch Dis Child, 2008. **93**: p. 428-432.

30. Al Sayah F and Williams B, *An Integrated Model of Health Literacy using Diabetes as an Exemplar* Canadian Journal of Diabetes 2012 **36**: p. 27-31.

31. Ishikawa H, et al., *Patient health literacy and patient–physician information exchange during a visit.* Family Practice 2009

32. Schillinger D, et al., *Functional health literacy and the quality of physician–patient communication among diabetes patients.* Patient Education and Counseling, 2004. **52**: p. 315–323.

33. Schillinger D, et al., *Closing the loop: Physician communication with diabetic patients who have low health literacy.* Archives of Internal Medicine 2003. **163**: p. 83-90.

34. Kirsch J S, et al., *Adult Literacy in America: A first look at the results of the National Adult Literacy Survey (NALS)*, 1993: Washington, DC.

35. Nielsen-Bohlman L, Panzer AM, and Kindig DA, *Health literacy: A Prescription to end confusion* 2004: Washington DC.

36. Nutbeam, D., *The evolving concept of health literacy.* Soc Sci Med, 2008. **67**(12): p. 2072-8.

37. Rootman I and Gordon-El-Bihbety D, *A vision for health literate Canada: Report of the Expert Panel on Health Literacy*, 2008, Canadian Public Health Association Ottawa.

38. Baker D, *The meaning and measure of health literacy.* Journal of General Internal Medicine, 2006. **21**: p. 878-883.

39. Parrish GR, *Measuring Population Health Outcomes.* Preventing Chronic Disease, 2010. **7**(4).

40. Donabedian A, *The role of outcomes in quality assessment and assurance.* QRB Qual Rev Bull., 1992(18): p. 356–360.

41. Donabedian A, *The Definition of Quality and Approaches to its Assessment*1980, Ann Arbor, Michigan: Health Administration Press.

42. Krousel-Wood MA, *Practical Considerations in the Measurement of Outcomes in Healthcare.* The Ochsner Journal, 1999. **1**(4): p. 187–194.

43. Schultz S and Kopec J, *Impact of chronic conditions*, in *Statistics Canada catalogue*2003, Statistics Canada. p. 41–56.

44. Ali S, et al., *The prevalence of co-morbid depression in adults with type 2 diabetes: a systematic review and meta-analysis.* Diabetic Medicine, 2006. **23**(11): p. 1165-1173.

45. Bodenheimer T, et al., *Patient self-management of chronic disease in primary care.* JAMA, 2002. **288**: p. 2469-2475.

46. Wolf MS, et al., *Patients' shame and attitudes toward discussing the results of literacy screening.* Journal of Health Communication, 2007. **12**: p. 721-732.

47. Schillinger D, et al., *Functional health literacy and the quality of physician-patient communication among diabetes patients.* patient Education and Counseling 2004. **52**: p. 315-323.

48. Castro C, et al., *Babel Babble: Physicians' use of unclarified medical jargon with patients.* American Journal of Health Behavior 2007. **31**(Suppl 1).

49. Bains SS and Egede LE, *Associations between health literacy, diabetes knowledge, self-care behaviors, and glycemic control in a low income*

population with type 2 diabetes. Diabetes Technology and Therapeutics 2011. **13**: p. 335-341.

50. Mancuso JM, *Impact of health literacy and patient trust on glycemic control in an urban USA population.* Nursing and Health Sciences, 2010. **12**: p. 94-104.

51. Osborn C, et al., *Self-Efficacy Links Health Literacy and Numeracy to Glycemic Control.* Journal of Health Communication, 2010. **15**: p. 146-158.

52. Osborn CY, Bains SS, and Egede LE, *Health literacy, diabetes self-care, and glycemic control in adults with type 2 diabetes* Diabetes Technology and Therapeutics, 2010. **12**: p. 913-919.

53. Paasche-Orlow M and Wolf M, *The causal pathway linking health literacy to health outcomes* American Journal of Health Behavior 2007. **31**: p. S19-S26

54. Bandura A, *Health promotion from the perspective of social cognitive theory.* Psychology & Health, 1998. **13**(4).

55. Bandura A, *Health Promotion by Social Cognitive Means.* Health Education & Behavior, 2004. **31**(2): p. 143-164.

56. Baker DW, et al., *The health care experience of patients with low literacy.* Archives of Family Medicine, 1996. **5**: p. 329-334.

57. Pleasant A, McKinney J, and Rikard RV, *Health Literacy Measurement: A Proposed Research Agenda.* Journal of Health Communication: International Perspectives, 2011. **16**(3): p. 11-21.

58. Ishikawa H, Takeuchi T, and Yano E, *Measuring functional, communicative, and critical health literacy among diabetic patients.* Diabetes Care, 2008. **31**: p. 874-879.

59. Rudd RE, *Improving Americans' health literacy.* N Engl J Med, 2010. **363**(24): p. 2283-2285.

60. Parker R, *Health literacy: a challenge for American patients and their health care providers.* Health Promotion International, 2000. **15**(4): p. 277-283

61. Gazmararian JA, et al., *Factors associated with medication refill adherence in cardiovascular-related diseases.* Journal of General Internal Medicine 2006. **21**: p. 1215-1221.

Chapter 2

Health Literacy and Health Outcomes in Diabetes: A Systematic Review [*]

Abstract

Background: Low health literacy (HL) is considered a potential barrier to improving health outcomes in people with diabetes and other chronic conditions although the evidence has not been previously systematically reviewed.

Objective: To identify, appraise, and synthesize research evidence on the relationships between HL (functional, interactive, and critical) or numeracy and health outcomes (i.e., knowledge, behavioral and clinical) in people with diabetes.

Methods: English-language articles that addressed the relationship between HL or numeracy and at least one health outcome in people with diabetes were identified by two reviewers through searching six scientific databases, and hand-searching journals and reference lists.

Findings: 723 citations were identified and screened, 196 considered, and 34 publications reporting data from 24 studies met the inclusion criteria and were included in this review. Consistent and sufficient evidence showed a positive association between HL and diabetes knowledge (8 studies). There was a lack of consistent evidence on the relationship between HL or numeracy and clinical outcomes, e.g., A1C (13 studies), self-reported complications (2 studies), and achievement of clinical goals (1 study); behavioral outcomes, e.g., self-monitoring

---

[*] A version of this chapter has been published. Al Sayah F, Majumdar SR, Williams B, Robertson S, Johnson JA. Health literacy and health outcomes in Diabetes: A systematic review. *Journal of General Internal Medicine* 2013 Mar; 28(3): 444-52.

of blood glucose (1 study), self-efficacy (5 studies); or patient-provider interactions (i.e., patient-physician communication, information exchange, decision-making, and trust), and other outcomes. The majority of the studies were from US primary care setting (87.5%), and there were no randomized or other trials to improve HL.

Conclusion: Low HL is consistently associated with poorer diabetes knowledge. However, there is little sufficient or consistent evidence suggesting that it is independently associated with processes or outcomes of diabetes-related care. Based on these findings it may be premature to routinely screen for low HL as a means for improving diabetes-related health-related outcomes.

2.1. Introduction

Health literacy (HL) is a set of skills that people need to function effectively in the healthcare environment [1]. These include functional, interactive, critical and numeracy skills. Functional skills are the ability to read and understand written text, locate and interpret information in documents, and write or complete forms. Interactive skills are ability to speak and listen effectively and communicate about health-related information. Critical skills are the ability to navigate the healthcare system and make appropriate health decisions. Numeracy skills are ability to use numeric information for tasks, such as interpreting medication dosages and food labels [1, 2].

Low HL is recognized as a stronger predictor of a persons' health than age, income, employment status, education level, or race [3], and is associated with a wide range of adverse effects on care processes and health outcomes [4, 5]. Low levels of HL are more prevalent in minority populations, among persons for whom English is a second language, people with low levels of income and education, and people with a compromised health status and elder communities - the very same populations that carry the greatest burden of chronic conditions [3, 6-8].

Diabetes is a prototypical chronic disease, characterized by a high level of complexity that requires extensive self-care education and management. The demands on individuals with diabetes are complicated by the fact that self-care often relies on printed educational materials and verbal instructions and requires advanced HL skills [9]. There is a growing body of literature that explores the relationship between HL and health outcomes in people with diabetes. Older studies of low HL reported adverse impacts on diabetes-related health outcomes

[10-12]; however, more recent studies showed no association between HL levels and intensity, frequency or incidence of outcomes [13, 14], and thus the effect of HL on the health of people with diabetes is yet unclear.

To better understand the relationship between HL and numeracy and health outcomes in diabetes, we conducted a systematic review of the literature. Since HL and numeracy are measured as separate constructs, in synthesizing the literature, we considered the association between HL (functional, interactive, and critical components) and health outcomes, and between numeracy (computational component) and health outcomes separately in the diabetes population.

2.2. Methods

2.2.1. Data Sources, Search strategy and Study selection

We conducted a systematic review of six databases: CINAHL, Embase, ERIC, Medline, psycINFO, and SCOPUS. The searches were not limited to any time period, language, or type of published paper. No Medical Subject Heading terms specifically identify HL-related articles, so we conducted the searches using different combinations of the following keywords: literacy, numeracy, HL, diabetes, diabetic, type 2 diabetes, type 1 diabetes, and the names of HL instruments. Keywords were matched to database specific indexing terms (detailed information about the search strategy are available upon request from FAS). Electronic searches were supplemented by hand searches, review of the reference lists of the included articles, and contact with content experts.

Additionally, authors were contacted when additional information to determine eligibility for inclusion was needed.

Two reviewers screened the identified records, reviewed the full text of the included articles, and performed data abstraction and assessment of quality and strength of evidence. Criteria for inclusion were studies that: addressed the relationship between HL or numeracy and health outcomes in people with type 1 or type 2 diabetes, involved the use of valid direct or indirect measure of HL or numeracy, addressed at least one health outcome, and were written in English. Cohen's kappa was used to assess inter-rater reliability in selecting studies for inclusion. Initial discrepancies were addressed through consensus and, if necessary, resolved by the senior author (JAJ).

2.2.2. Data Abstraction and Quality Assessment

Data on the general characteristics of the studies including aim, design, methods, sample size and characteristics, HL measurement, outcomes measurement, analysis, results and conclusions were extracted (Tables S2-1 and S2-2). Data abstraction was done by one reviewer, and confirmed by the second reviewer. We were interested in summarizing evidence on the following outcomes: 1) clinical, 2) behavioral, 3) patient-provider relationship, and 4) other outcomes. If a study produced multiple publications on the relationship between a particular outcome and HL, estimates from only one publication, the most recent one, were considered in the assessment of the strength of evidence.

Since all the identified studies were observational, with no intervention intended to affect the outcome, each article was rated based on the criteria in the quality rating for observational studies by the Agency for Healthcare Research (AHRQ) [15]. Quality ratings presented in Tables S2-1 and S2-2. Two reviewers independently rated the quality of studies as good, fair, or poor based on the criteria that focuses on detecting bias in an observational study and precision of measurement. Because analysis techniques differed among multiple publications of the same study, overall quality was assessed for each publication, not each study. Only studies that were of fair to good quality were included in this review. Cohen's kappa was used to assess inter-rater reliability in quality rating between the two reviewers.

2.2.3. Strength of Evidence Assessment

The strength of the evidence for each outcome was determined using the AHRQ guidelines [16] for publications measuring HL and numeracy. Two reviewers independently graded the strength of evidence as high, moderate, low or insufficient on the basis of potential risk of bias of the included publications, consistency of effect across publications, directness of evidence, and precision of the estimate (Table 2-1). Cohen's kappa was used to assess inter-rater reliability in rating the strength of evidence between the two reviewers.

We attempted to perform a meta-analysis to quantitatively summarize the evidence for each outcome. In both fixed and random effects models, heterogeneity was large (I^2 = 80-90%), which meant that a meta-analysis was not feasible, and pooled estimates of effects were not reported. In comparing the general characteristics of the studies, we noticed methodological variations between studies with respect to sample size, sample characteristics, HL measurement, and in the adjustment of potential confounders. These factors might have contributed to the observed heterogeneity. For these reasons, the evidence from included studies was summarized qualitatively in this review.

2.3. Results

2.3.1. Literature Search

Across all databases, the search yielded 2138 citations (Figure 2-1). After removal of duplicates, 723 remained. Titles and abstracts were screened for relevance, and based on that, 196 publications were included for full-text review. Of these, 32 publications met our eligibility criteria and were included. The most common reasons for exclusion were: review and conceptual articles (35.2%), studies that did not address any of the outcomes of interest (31.5%), and studies that did not involve individuals with diabetes (17.3%). Studies that addressed HL in caregivers of individuals with diabetes (N = 4), and in women with gestational diabetes (N = 1) were also excluded. The hand search of reference lists of included studies resulted in the inclusion of two additional publications. Thus, the total number of eligible publications was 34. Inter-rater agreement was 88% and Cohen's kappa was 0.70 (95%CI: 0.59 - 0.84). The 34 articles were multiple publications from 24 studies. We reported the findings of this review based on the 24 studies and evaluated the evidence based on estimates from the most recent publication of each study to avoid bias due to multiple publications.

2.3.2. Characteristics of the Included Studies

The 34 identified publications of the 24 studies were of fair to good quality (Tables S2-1 and S2-2), and thus they were all included in this review. Inter-rater agreement on quality rating was 97% and Cohen's kappa was 0.91 (95%CI: 0.76-0.98). The majority of publications (29 from 21 studies) addressed only HL and health outcomes, while only 5 publications (from 4 studies) addressed health numeracy. All studies were observational; only two performed longitudinal data analysis (one using data from a randomized controlled trial and one from a prospective cohort study). There were no experimental studies that addressed the effect of HL and/or numeracy on health outcomes in this population. Half of the studies included only individuals with type 2 diabetes (12 studies) and 12 included individuals with type 1 or type 2; no studies focused solely on type 1 diabetes. The majority of the identified studies were conducted in the United States, primarily on minority or vulnerable populations.

2.3.3. Health Literacy Measures

Several instruments were used to measure HL and numeracy in these studies. The most commonly used measure of HL was the Test of Functional Health Literacy in Adults-short form (s-TOFHLA) (10 studies), followed by the Rapid Estimate of Adult Literacy in Medicine (REALM) (7 studies), 3-brief screening questions (3 studies), the original long form of the TOFHLA (2 studies), the revised REALM-R (1 study), self-rated Health Literacy 3-item scale (1 study). Diabetes Numeracy Test was used to measure diabetes numeracy (3 studies), and the Wide Range Achievement Test (3 studies) and Subjective Numeracy Scale (1 study) to measure general numeracy.

The included studies differed in how investigators distinguished between levels or thresholds of HL, either as a continuous measure or categories (e.g., inadequate, marginal, adequate; or high versus low). When categorized, the majority of the studies focused on the differences between the lowest and highest groups. Additionally, studies differed in which domain of HL they addressed, where most used measures of functional HL (23 studies), and only one study addressed all components (functional, interactive, and critical).

2.3.4. Health literacy and Health Outcomes

Overall, inter-rater agreement for strength of evidence ratings for HL and health outcomes was 92% and Cohen's kappa was 0.85 (95%CI: 0.71-0.98). For HL, there were 23 outcomes grouped into four categories: Clinical, behavioral, patient-provider communication, and other outcomes (Table 2-2).

2.3.4.1.Health literacy and clinical outcomes

Glycemic control: The relationship between HL and A1c was explored in 13 studies; 12 of which were cross-sectional and one longitudinal [17]. Some of the identified studies showed that higher levels of HL were associated with better glycemic control [10, 12, 18, 19], and this relationship was mostly observed in studies that adjusted for age, sex, race, education, and treatment regimen. Moreover, few studies reported that HL has indirect effect on glycemic control, but did not show any direct associations [20]. Overall, evidence on the relationship between HL and glycemic control was inconsistent across studies, and the heterogeneity did not permit the estimation of an overall effect. Therefore, this evidence was rated insufficient.

We were able to observe a trend between confounders controlled for and reported estimates in studies that explored that relationship between HL and A1c. The association between HL and A1c was significant in studies that did not adjust for diabetes knowledge [10, 12, 19]; however, that association was not observed in studies that controlled for diabetes knowledge [13, 14, 17]. Similar trends were not observed for other outcomes because of the small number of available studies.

Hypoglycemia: Two studies explored the relationships between low HL and self-reported hypoglycemia and one of these studies reported that lower HL was associated with higher frequency of self-reported hypoglycemia [11]. The quality of evidence from these two studies was rated low.

Blood pressure: Two studies explored the relationship between HL and blood pressure in people with diabetes [21, 22], although only one adjusted for potential confounders [22]. Both studies reported that HL was not associated with blood pressure control. This evidence was rated low.

Diabetes complications: The relationship between HL and self-reported complications was explored in two studies both adjusted for potential confounders [12, 22]. One study reported that lower HL was associated with retinopathy and stroke, but not with nephropathy, lower extremity amputation or ischemic heart disease [12]. The other study showed no association between HL and self-reported complications [22]. The evidence from these studies was inconsistent and rated insufficient.

Low Density Lipoprotein (LDL): One study that adjusted for potential confounders explored the relationship between HL and LDL and showed that HL was not associated with LDL levels [22].

2.3.4.2. Health literacy and behavioral indicators and patient-reported outcomes

Diabetes knowledge: Nine studies, six of which adjusted for potential confounders [10, 13, 14, 23-25] and three did not [21, 26, 27], provided high evidence that higher HL levels were associated with better diabetes knowledge.

Self-efficacy: Five studies provided evidence on the relationship between HL and self-efficacy. Three studies reported adjusted results and showed no association between HL levels and self-efficacy [18, 21, 24]. One study that adjusted for confounders [28] showed no association between HL and self-efficacy, and the unadjusted study [29] showed that higher HL levels were associated with higher self-efficacy scores. The evidence was inconsistent and rated insufficient.

Self-care: Four studies that reported adjusted results showed no association between HL and self-care behaviors namely diet, exercise, blood sugar testing, foot care, smoking cessation, and medication adherence [13, 14, 21, 25]. Since the evidence on this relationship was consistent, it was rated moderate.

Self-monitoring of Blood Glucose (SMBG) and other self-management: Three studies explored the relationship between HL and SMBG and self-management support and all adjusted for potential confounders. One study did not show an association between HL and the frequency of SMBG [29], and another study did not show an association with medication adherence [13]. The third study showed that higher HL levels were associated with higher self-management support ratings [30]. The evidence from each of these studies was rated low.

2.3.4.3. Health literacy and patient-provider interaction indicators

Patient-provider communication: Two studies [31, 32], where only one reported adjusted results, showed that higher HL levels were associated with better patient-physician communication [32]. The evidence from these studies was rated low.

Patient trust: Two studies explored the relationship between HL and patient trust, where only one reported adjusted results and showed that higher HL levels were associated with higher scores on patient trust scores [14]. The other study did not show an association between HL and patient trust [29]. The evidence from the two studies was rated insufficient.

Information exchange and involvement in decision-making: One study reported adjusted results and showed that higher HL was associated with better information exchange between patients and their physicians [32]. The evidence from this study was rated low. Another study that did not adjust for confounders, showed no association between HL and patient's involvement in decision making with their physicians [29]. The evidence from this study was rated insufficient.

Use of computers and Internet: Two studies, where only one reported adjusted results [33], explored the relationship between HL and patient's use of computers and Internet for health-related learning. These studies provided low evidence that higher HL was associated with more frequent use of computers and Internet [33, 34].

Other outcomes: Six studies explored the relationship between HL and prevalence of heart failure, prevalence of depressive symptoms, health-related quality of life (HRQL), diabetes health-related beliefs, medication beliefs, and healthcare discrimination respectively, where all but one reported adjusted findings. In these studies there was no association between HL and prevalence of heart failure [35], prevalence of depressive symptoms [14], HRQL [29], or diabetes health-related beliefs [10]. One study showed that lower HL was associated with medication beliefs, particularly with more concern about the harmfulness of medications [36], and another showed that lower HL was associated with higher reporting of healthcare discrimination [37]. The evidence on these relationships was rated low, except for HRQL which was rated insufficient.

2.3.5. Numeracy and Health Outcomes

Overall, inter-rater agreement for strength of evidence ratings for numeracy and health outcomes was 94% and Cohen's kappa was 0.86 (95% CI: 0.73-0.99). There were four outcomes grouped into three categories: Clinical, behavioral, and other outcomes (Table 2-3).

2.3.5.1. Numeracy and clinical outcomes

Glycemic control: The relationship between numeracy and A1c was explored in four studies. Two studies adjusted for potential confounders; one showed that higher numeracy was directly associated with better A1c [38] and the other showed indirect effect of numeracy on A1c [39]. The two other studies that reported unadjusted results did not show an association between numeracy and A1c levels [34, 40]. The evidence was rated insufficient.

2.3.5.2. Numeracy and behavioral indicators and patient-reported outcomes

Self-efficacy: Two studies explored the relationship between numeracy and self-efficacy. One study reported adjusted results and showed that higher numeracy was associated with better self-efficacy [28], and the other study reported unadjusted results and showed no association between numeracy and self-efficacy [40]. This evidence was rated insufficient.

Self-care: One study that did not adjust for confounders reported no association between numeracy and self-care [40]. This evidence was rated insufficient.

Other outcomes: One unadjusted study explored that relationship between numeracy and the use of computers and internet, and reported that higher numeracy was associated with more frequent use of computers and Internet for health-related learning [34]. This evidence was rated insufficient.

2.4. Discussion

Our systematic review showed a discrepancy among studies regarding the relationship between HL or numeracy and several health outcomes in people with diabetes. Consistent evidence suggested a positive association between HL and diabetes knowledge but even this evidence was only rated sufficient. Likewise, there is likely sufficient evidence to support a positive relationship between HL and self-care activities. On the other hand, the evidence for an association between HL and clinical indicators was weak. We found little evidence to support (or refute) an association between HL and important clinical events (such as mortality, cardiovascular disease), other than self-reported hypoglycemia and presence of diabetic complications. The majority of this evidence comes from cross-sectional studies, however, limiting causal inference.

It is important to note that substantial discrepancies exists in the literature, which could be due to methodological issues and challenges in the identified studies. One potential source of discrepancy could be the different tools used to measure HL [41] and differences in thresholds used to distinguish between HL levels [42]. This variation in estimates and thresholds, in addition to the fact that these instruments measure different aspects of HL and thus reflect different skills [43], could have influenced the magnitude and the precision of the observed estimates in these studies.

Another potential reason for discrepancy could be adjustment for confounders. Most studies adjusted for age, sex, race, and educational level in the analyses; however, few studies also adjusted for other factors such as diabetes duration, diabetes knowledge, self-care, self-efficacy, health status, treatment regimen, and many others, where some of these were included as mediators in the pathway between HL and outcomes [28]. Adjusting for these confounders that are possibly intermediate variables could have induced over-adjustment bias in estimating direct effects of HL on outcomes [44].

Another equally interesting observation in this review was from recent studies that explored factors that mediate the relationship between HL and diabetes-related health outcomes. Osborn and colleagues [39] found that HL was not directly associated with self-care and A1c, however, was indirectly associated with these outcomes through social support. The same investigators in subsequent analysis of the same study found that HL was indirectly associated with A1c through self-efficacy [28]. Future research should further investigate these mediators and others to better understand the relationship between HL and health outcomes and what factors should be the target of intervention, HL or the mediators.

Other methodological issues that might have introduced the inconsistent results include the lack of power in some studies. Additionally, the heterogeneity of participants across studies could have also led to the observed inconsistency. This could indicate that HL might be related to certain outcomes in particular diabetes populations but not in others. These speculations are hard to examine, however, with the limited data and available studies.

A recent review by Berkman and colleagues explored the relationship between HL and health outcomes in patients of all ages and was not limited to any patient groups [4]. Their findings were similar to ours particularly on the relationship between HL and disease knowledge, and they reported inconsistent results regarding other outcomes that were not addressed in our review such as healthcare utilization and costs. Other reviews focused on specific populations such as emergency room patients [45], working-age adults [46], children [47, 48], and ambulatory care patients [49]. These reviews also had similar results to ours with respect to disease knowledge; most were not able to provide firm conclusions on other outcomes due to insufficient evidence.

This review, as any other systematic review, reflects the quality of the published literature. Although the quality of the included studies ranged from fair to good, these ratings did not reflect the limitations imposed by the cross-sectional design of the majority of the studies, using different measures of HL across studies, choosing different cut-points for analysis, the inconsistent and potentially inappropriate control for confounders, and poor reporting, which made comparisons between studies difficult. In addition to methodological limitations, the majority of the studies were conducted in primary care clinics in the US; only a few were population-based and very few conducted outside of the US (Japan, China, and Ireland).

2.5. Conclusion

Our review indicates that the current understanding of the effect of low HL on the health of people with diabetes is limited. We found that low HL is consistently associated with poorer diabetes knowledge. However, there is little sufficient or consistent evidence suggesting that it is independently associated with processes or outcomes of diabetes-related care. Given how important the topic is, we were surprised by the paucity of high-quality evidence.

Certainly our findings suggest that it might be premature to embark on randomized trials or controlled interventions to improve HL in those with diabetes given how little we actually know. Until better evidence is available, we believe that, outside of the study setting, it might be premature to routinely screen for HL or to try to improve HL for the purposes of improving patient-related outcomes in diabetes – although there may be other reasons to do so.

Tables and Figures

Table 2-1: Strength of Evidence Grades and Definitions [16]

Grade	Definition
High	High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of the effect.
Moderate	Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of the effect and may change the estimate.
Low	Low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of the effect and is likely to change the estimate. The evidence was graded as low if findings were limited to only 1 or a few studies that controlled for potential confounding or the preponderance of evidence was based on studies that did not control for potential confounding.
Insufficient	Evidence is unavailable or does not permit estimation of an effect. Inconsistent findings across studies were generally graded as insufficient, as was evidence limited to 1 study that did not control for potential confounding.

Table 2-2: Health Literacy and Outcomes in Diabetes: Strength of Evidence and Summary of Findings

Category	Outcome	Total # of studies by study design = N (# of studies NOT controlling for confounding = n1; # of studies controlling for confounding = n2)		Summary of Results				Strength of Evidence
				Studies NOT controlling for confounding (n1)		Studies controlling for confounding (n2)		
		Design	**N (n1; n2)**	**Positive results***	**Negative results†**	**Positive results***	**Negative results†**	
Clinical outcomes	HbA1c	Cross-sectional	12 (4; 8)	1	3	4	4	Insufficient
		Longitudinal	1 (0; 1)	-	-	-	1	
	LDL	Cross-sectional	1 (0; 1)	-	-	-	1	Low
	Blood pressure	Cross-sectional	2 (1; 1)	-	1		1	Low
	Self-reported significant hypoglycemia	Cross-sectional	1 (0; 1)	-	-	1	-	Low
	Self-reported complications	Cross-sectional	2 (0; 2)	1	-	-	1	Insufficient
	Achievement of clinical outcome goals (A1c)	Longitudinal	1 (0; 1)	-	-	1	-	Low
Behavioral outcomes	Self-care	Cross-sectional	4 (0; 4)	-	-	4	-	Moderate
	Self-monitoring of blood glucose	Cross-sectional	1 (0; 1)	-	-	-	1	Low
	Self-efficacy	Cross-sectional	5 (1; 4)	1	-	1	3	Insufficient
	Self-management support	Cross-sectional	1 (0; 1)	-	-	1		Low
	Medication adherence	Cross-sectional	1 (0; 1)	-	-	-	1	Low
	Diabetes knowledge	Cross-sectional	8 (3; 5)	3	-	5	-	High
		Longitudinal	1 (0; 1)		-	1	-	
Patient-provider	Patient-physician communication	Cross-sectional	2 (1; 1)	1	-	1	-	Low

interaction	Information exchange	Cross-sectional	1 (0; 1)	-	-	1	-	Low
	Involvement in decision making	Cross-sectional	1 (1: 0)	1	-	-	-	Insufficient
	Trust	Cross-sectional	2 (1; 1)	-	1	1	-	Insufficient
Other outcomes	Prevalence of heart failure	Cross-sectional	1 (0; 1)	-	-	-	1	Low
	Prevalence of depressive symptoms	Cross-sectional	1 (0; 1)	-	-	-	1	Low
	Health-related quality of life	Cross-sectional	1 (1; 0)	-	1	-		Insufficient
	Diabetes health-related beliefs	Cross-sectional	1 (0; 1)	-	-	-	1	Low
	Medication beliefs	Cross-sectional	1 (0; 1)	-	-	1	-	Low
	Healthcare and general discrimination	Cross-sectional	1 (0; 1)	-	-	1	-	Low
	Use of computers and internet	Cross-sectional	2 (1; 1)	1	-	1	-	Low

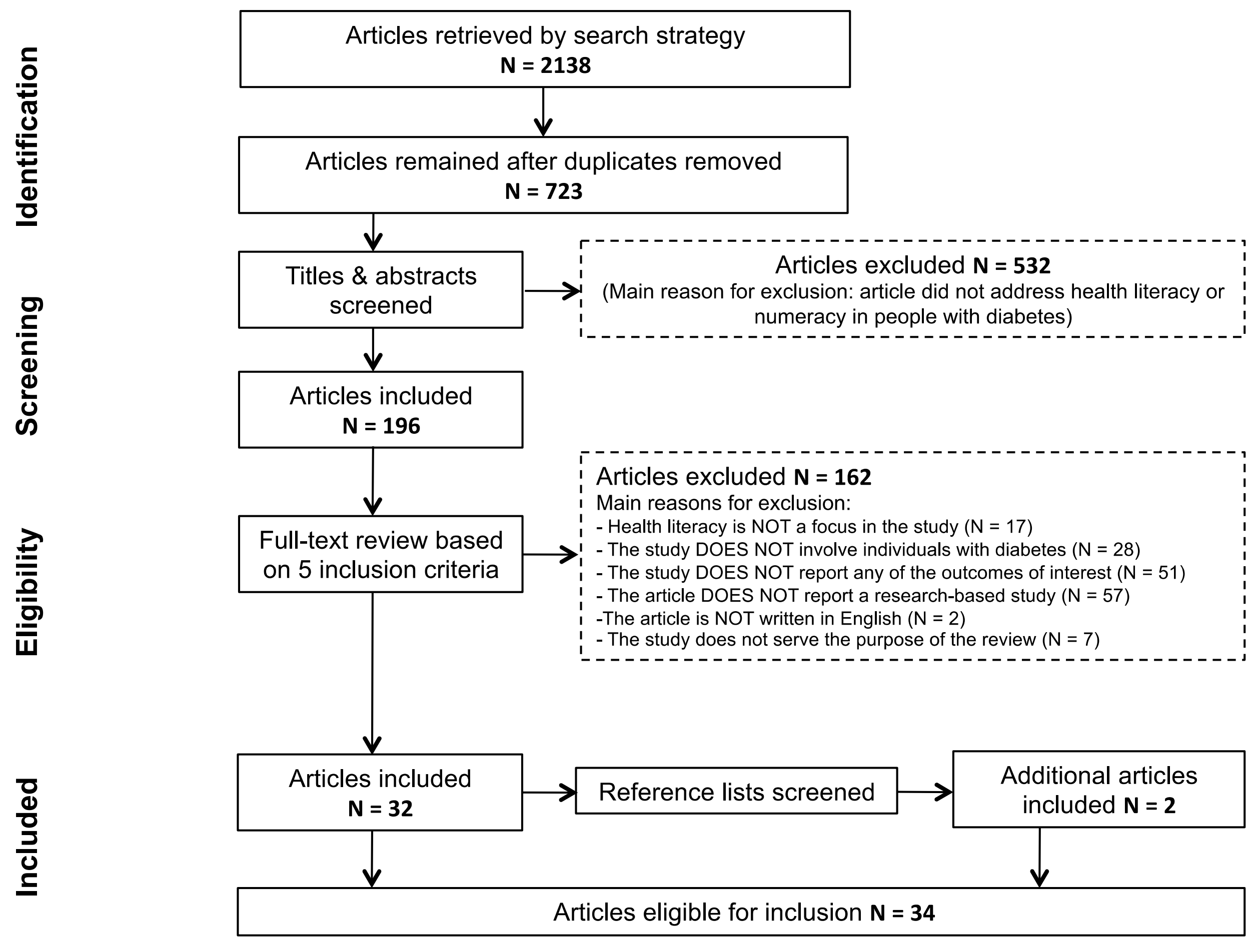
* Indicates that the study found a significant association between health literacy level and the outcomes
† Indicates that the study did not find a significant association between health literacy level and the outcomes

Table 2-3: Numeracy and Outcomes in Diabetes: Strength of Evidence and Summary of Findings

Category	Outcome	Total # of studies by study design = N (# of studies NOT controlling for confounding = n1; # of studies controlling for confounding = n2)		Summary of results: Studies NOT controlling for confounding (n1)		Summary of results: Studies controlling for confounding (n2)		Strength of evidence
		Design	N (n1; n2)	Positive results*	Negative results†	Positive results*	Negative results†	
Clinical outcomes	A1c	Cross-sectional	4 (1; 3)	-	1	2	1	Insufficient
Behavioral outcomes	Self-care	Cross-sectional	1 (1; 0)	-	1	-	-	Insufficient
	Self-efficacy	Cross-sectional	2 (1; 1)	-	1	1	-	Insufficient
Other outcomes	Use of computers and internet	Cross-sectional	1 (1; 0)	1	-	-	-	Insufficient

* Indicates that the study found a significant association between health literacy level and the outcomes
† Indicates that the study did not find a significant association between health literacy level and the outcomes

Figure 2-1: PRISMA Diagram of the Search and Retrieval Process



References

1. Berkman N, Terry D, and McCormack L, *Health literacy: What is it?* Journal of Health Communication, 2010. **15**: p. 9-19.

2. Nutbeam D, *The evolving concept of health literacy.* Social Science and Medicine 2008 **67**: p. 2072-2078.

3. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, A.M.A., *Health literacy: Report on the Council of Scientific Affairs.* JAMA, 1999. **282**: p. 526-527.

4. Berkman ND, et al., *Low health literacy and health outcomes: an updated systematic review* Annals of Internal Medicine 2011. **155**(2): p. 97-107.

5. DeWalt DA, et al., *Literacy and Health Outcomes: A Systematic Review of the Literature.* Journal of General Internal Medicine, 2004 **19**(12): p. 1228–1239.

6. Schultz S and Kopec J, *Impact of chronic conditions. Health Reports.* Statistics Canada catalogue no. 82-003-XIE, 2003. **14**: p. 41-56.

7. Sudore R, et al., *Limited literacy in older people and disparities in health and healthcare access.* Journal of the American Geriatrics Society, 2006. **54**: p. 770-776.

8. Wolf M, Gazmararian J, and Baker D, *Health literacy and functional health status among older adults* Archives of Internal Medicine, 2005. **165**: p. 1946-1952.

9. White R, et al., *Addressing Health Literacy and Numeracy to Improve Diabetes Education and Care.* Diabetes Spectrum 2010. **23**: p. 238-243

10. Powell CK, Hill EG, and Clancy DE, *The relationship between health literacy and diabetes knowledge and readiness to take health actions.* The Diabetes Educator 2007. **33**: p. 144-151.

11. Sarkar U, et al., *Hypoglycemia is more common among type 2 diabetes patients with limited health literacy: The diabetes Study of Northern California (DISTANCE).* Journal of General Internal Medicine 2010. **25**: p. 962-967.

12. Schillinger D, et al., *Association of health literacy with diabetes outcomes.* JAMA, 2002. **288**: p. 475-482.

13. Bains SS and Egede LE, *Associations between health literacy, diabetes knowledge, self-care behaviors, and glycemic control in a low income population with type 2 diabetes.* Diabetes Technology and Therapeutics 2011. **13**: p. 335-341.

14. Mancuso JM, *Impact of health literacy and patient trust on glycemic control in an urban USA population.* Nursing and Health Sciences, 2010. **12**: p. 94-104.

15. Berkman ND, et al., *Health literacy interventions and outcomes: an update of the literacy and health outcomes systematic review of the*

literature. Evidence Report/Technology Assessment no.199. , 2011, Agency for Healcare Research and Quality: Rockville, MD.

16. Owens DK, et al., *AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions – agency for healthcare research and quality and the effective health-care programs.* Journal of Clinical Epidemiology, 2010. **63**: p. 513-23.

17. Rothman R, et al., *Influence of patient literacy in the effectiveness of a primary care-based diabetes disease management program.* JAMA, 2004. **292**: p. 1711-1716.

18. Ishikawa H and Yano E., *The relationship of patient participation and diabetes outcomes for patients with high vs. low health literacy.* Patient Education and Counseling 2011. **84**(3): p. 393-7.

19. Tang YH, et al., *Health literacy, complication awareness, and diabetic control in patients with type 2 diabetes mellitus* AJAN, 2008. **62**: p. 74-83.

20. Osborn CY, et al., *Diabetes Numeracy An overlooked factor in understanding racial disparities in glycemic control* Diabetes Care 2009. **32**(9).

21. McCleary-Jones V, *Health literacy and its association with diabetes knowledge, self efficacy and diabetes self-management among African Americans with Diabetes Mellitus* The ABNF Journal 2011. **22**(2): p. 25-32.

22. Morris NS, et al., *Literacy and health outcomes: A cross-sectional study in 1002 adults with diabetes.* BMC Family Practice 2006. **7:** p. 49.

23. Gazmararian JA, et al., *Health literacy and knowledge of chronic disease.* Patient Education and Counseling 2003. **51**: p. 267-275.

24. Gerber BS, et al., *The assessment of diabetes knowledge and self-efficacy in diverse population using Rasch Measurmeent* Journal of Applied Measurement 2006 **7**(1): p. 55-73.

25. Kim S, et al., *Association of health literacy with self-management behavior in patients with diabetes.* Diabetes Care, 2004. **27**: p. 2980-2982.

26. Ishikawa H, Takeuchi T, and Yano E, *Measuring functional, communicative, and critical health literacy among diabetic patients.* Diabetes Care, 2008. **31**: p. 874-879.

27. Williams M, et al., *Relationship of functional health literacy to patients' knowledge of chronic disease: A study of patients with hypertension and diabetes* Archives of Internal Medicine 1998. **158**: p. 166-172.

28. Osborn C, et al., *Self-Efficacy Links Health Literacy and Numeracy to Glycemic Control.* Journal of Health Communication, 2010. **15**: p. 146-158.

29. DeWalt DA, Boone RS, and Pignone MP, *Literacy and Its Relationship With Self-efficacy, Trust, and Participation in Medical Decision Making.* American Journal of Health Behavior, 2007 **31**(Suppl 1): p. S27-S35.

30. Wallace AS, et al., *The influence of literacy on patient-reported experiences of diabetes self-management support.* Nursing Research, 2010. **59**: p. 356-363.

31. Arthur SA, et al., *Health literacy and control in the medical encounter: A mixed-methods analysis.* J National Med Assoc, 2009. **101**: p. 677-683.

32. Ishikawa H, et al., *Patient health literacy and patient-physician information exchange during a visit.* Family Practice 2009. **26**: p. 517-523.

33. Sarkar U, et al., *The literacy divide: Health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE).* Journal of Health Communication 2010. **15**(Suppl 2): p. 183-196.

34. Mayberry LS, et al., *Bridging the digital divide in diabetes: Family support and implications for health literacy* Diabetes Technology and Therapeutics, 2011. **in press**.

35. Laramee AS, Morris N, and Littenberg B, *Relationship of literacy and heart failure in adults with diabetes.* BMC Health Serv Res., 2007. **7**.

36. Aikens J and Piette J, *Diabetic patients' medication underuse, illness outcomes, and beliefs about antihyperglycemic and antihypertensive treatments.* Diabetes Care, 2009. **32**: p. 19-24.

37. Lyles C, et al., *Correlates of patient-reported racial/ethnic health care discrimination in the diabetes study of northern California (DISTANCE)* Journal of Health Care for the Poor and Underserved, 2011. **22**: p. 211-225.

38. Cavanaugh K, et al., *Association of Numeracy and Diabetes Control. Annals of Internal Medicine*, 2008. **148**(10): p.737-46.

39. Osborn CY, Bains SS, and Egede LE, *Health literacy, diabetes self-care, and glycemic control in adults with type 2 diabetes* Diabetes Technology and Therapeutics, 2010. **12**: p. 913-919.

40. White RO, et al., *Development and validation of a Spanish diabetes-specific numeracy measure: DNT-15 Latino* Diabetes Technology and Therapeutics, 2011. **13**(9).

41. Griffin JM, et al., *Variation in Estimates of Limited Health Literacy by Assessment Instruments and Non-Response Bias* Journal of General Internal Medicine, 2010. **25**(7): p. 675-681.

42. Wolf MS, et al., *In search of 'low health literacy': threshold vs. gradient effect of literacy on health status and mortality.* Social Science and Medicine 2010. **70**(9): p. 1335-41.

43. Al Sayah F, Williams B, and Johnson JA, *Measuring Health Literacy in Individuals With Diabetes: A Systematic Review and Evaluation of Available Measures* Health Education and Behavior 2012. **40**( ).

44. Schisterman EF, Cole SR, and Platt RW, *Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies.* Epidemiology, 2009. **20**(4): p. 488-495.

45. Herndon JB, Chaney M, and Carden D, *Health literacy and emergency department outcomes: a systematic review.* Annals of Emergency Medicine, 2011. **57**(4): p. 334-45.

46. Easton P, Entwistle VA, and Williams B, *Health in the 'hidden population' of people with low literacy. A systematic review of the literature.* BMC Public Health, 2010. **5**(10).

47. DeWalt DA and Hink A, *Health literacy and child health outcomes: a systematic review of the literature.* Pediatrics, 2009. **124**(Suppl 3): p. S265-74.

48. Sanders LM, et al., *Literacy and child health: a systematic review.* Archives of Pediatrics and Adolescent Medicine 2009. **163**(2): p. 131-40.

49. Keller DL, Wright J, and Pace HA, *Impact of health literacy on health outcomes in ambulatory care patients: a systematic review.* The Annals of Pharmacotherapy, 2008. **42**(9): p. 1272-81.

50. Mbaezue N, et al., *The impact of health literacy on self-monitoring of blood glucose in patients with diabetes receiving care in an inner-city hospital.* J Natl Med Assoc., 2010. **102**: p. 5-9.

51. Schillinger D, et al., *Closing the loop: Physician communication with diabetic patients who have low health literacy.* Archives of Internal Medicine 2003. **163**: p. 83-90.

52. Schillinger D, et al., *Functional health literacy and the quality of physician-patient communication among diabetes patients.* patient Education and Counseling 2004. **52**: p. 315-323.

53. Schillinger D, et al., *Does literacy mediate the relationship between education and health outcomes? A study of a low-income population with diabetes.* Public Health Rep, 2006. **121**: p. 245-254.

54. Thabit H, et al., *Globalization, immigration and diabetes self-management: an empirical study amongst immigrants with type 2 diabetes mellitus in Ireland* QJM: An International Journal of Medicine, 2009 **102**(10): p. 713-720.

Chapter 2 Supplementary Tables

Table S2-1: Overview of Health Literacy Studies

Author(s), year Source	Design	Participants' characteristics N, diabetes type, recruitment site [age, sex, race, educational level]	HL measure (levels) / [domain addressed]	Outcomes	Type of analysis performed	Covariates*:	Main findings††	Quality Rating
Aikens & Piette, 2009 USA [36]	Cross-sectional	1376 patients with T2DM recruited from general medicine clinics [antihyperglycemic group: Mean age (SD): 55.3(11.8); 61.6% females; Race: 58.3% African American & 41.7% other races; years of education: 21.6% < 12 years; 35.7% = 12 years, & 42.7% > 12 years; antihyperglycemic + antihypertensive group: Mean age (SD): 57.2(10.7); 68.3% females; Race: 61.9% African American & 38.1% other races; years of education: 22.8% < 12 years; 36.6% = 12 years, & 40.6% > 12 years]	3-brief* SQs (problem learning, needing help reading, not confident with forms) [Functional HL]	Medication beliefs (BMQ)	Spearman correlation, multiple linear regression analysis	Age, sex, race, income, number of medical conditions, number of prescribed medications, insulin use, satisfaction with medication information	HL was associated with concerns about medications harmfulness	Good
Arthur et al., 2009 USA [31]	Cross-sectional	31 patients with T2DM recruited from a primary care clinic in an urban public hospital [79.4% females; Race: 100% African American; Mean years of education (SD):	REALM Categorical (2 levels: inadequate vs. marginal-adequate)	Patient-physician communication	Chi-square difference test	None	HL was associated with paternalistic interactions with care	Fair

		11.6 (2.3)]	[Functional HL]				providers	
Bains & Egede, 2011 USA Ψ [13]	Cross-sectional	125 patients with T2DM recruited from a primary care clinic [Age: 50.7% were under 65 years old & 49.3% 65 years and above; 72.5% females, Race: 71.4% black & 28.6% white; Education: 68.2% ≤ high school & 31.8% > high school]	REALM-R Continuous [Functional HL]	Glycemic control (A1c), diabetes knowledge (DKQ), self-care (SDSCA), medication adherence (Morisky score)	Spearman correlation Multiple linear regression	Age, sex, race, educational level, income, health status	HL was associated with diabetes knowledge, but not with A1c, self-care or medication adherence	Good
DeWalt et al., 2007 USA [29]	Cross-sectional	268 patients with T2DM recruited from general internal medicine clinics [High HL group: Mean age (SD): 58 (11); 69% females; Race: 48% African American, 49% white; Mean education years (SD): 12(3); Low HL group: Mean age (SD): 62 (10); 57% females; Race: 74% African American, 25% white, & 1% other races; Mean education years (SD): 8(3)]	REALM Categorical (2 levels: low vs. high) Continuous [Functional HL]	Glycemic control (A1c), health-related quality of life, trust (WFPTS), self-efficacy (DMSES), involvement in decision making (DPMD, FPI), knowledge	Pairwise correlation	None	HL was associated with self-efficacy, and involvement in decision making, but not with A1c, health-related quality of life, or trust	Fair-good
Gazmarari an et al., 2003 USA [23]	Cross-sectional	266 patients with T1DM or T2DM recruited from a national managed care organization [NR]	s-TOFHLA Categorical (3 levels: inadequate, marginal, adequate) Continuous [Functional HL]	Diabetes knowledge (11 questions)	Multiple linear regression	Age, diabetes duration, attended a diabetes class, education, cognitive health,	HL was associated with diabetes knowledge	Fair-good

						physical functioning, mental functioning		
Gerber et al. 2006 USA [24]	Cross-sectional	255 patients with T1DM or T2DM recruited from 5 community-based clinics [Mean age (SD): 55.2(12.3); 65.9% females; Race: 30.2% African-American, 64.7% Hispanic, 3.1% Caucasian, & 2% other races; Education: 45% < high school, 34.5% some or completed high school, & 20.4% > high school]	s-TOFHLA Categorical (3 levels: inadequate, marginal, adequate) [Functional HL]	Diabetes knowledge (KDQ), self-efficacy	Partial correlation	Age, diabetes duration, treatment regimen	HL was associated with knowledge, but not with self efficacy	Good
Ishikawa & Yano, 2011 Japan ϖ [18]	Cross-sectional	143 patients with T2DM recruited from a university hospital [Mean age (SD): 65(9.1); 42.7% females; Education: 23.8% middle school, 48.9% high school, 13.3% vocational school, 14% university or higher]	Self-rated HL scale**: Communicative HL-subscale [Interactive HL]	Glycemic control (A1c), self-efficacy	Pearson correlation Multiple linear regression Logistic regression	Age, sex, education, duration of diabetes, A1c, and visit length	Communicative HL was associated with self-efficacy and with A1c	Good
Ishikawa et al., 2009 Japan ϖ [32]	Cross-sectional	134 patients with T2DM recruited from a university hospital [Mean age (SD): 65(9); 44% females; Education: 23.9% middle school, 48.5% high school, 14.2% vocational school, 13.4% university or higher]	Self-rated HL scale, 3 sub-scales** [Functional HL, Interactive HL, Critical HL]	Patient-physician information exchange (RIAS)	Logistic regression	Age, sex, education, duration of diabetes, A1c, visit length	Functional HL was associated with asking close-ended questions & information giving; Critical HL was	Good

							associated with psychosocial information giving; communicative HL was associated with counseling, and perception of physician's explanation.	
Ishikawa et al., 2008 Japan ϖ [26]	Cross-sectional	138 patients with T2DM recruited from outpatient family medicine clinics [Mean age (SD): 65(9.9); 47.1% females; Education: 21.7% middle school, 43.5% high school, 13.8% vocational school, 12.3% university or higher]	Self-rated HL scale, 3 sub-scales** [Functional HL, Interactive HL, Critical HL]	Glycemic control (A1c), self efficacy, diabetes knowledge	Spearman correlation	None	HL was associated with diabetes knowledge and self efficacy, but not with A1c	Fair-good
Kim et al. 2004 USA [25]	Longitudinal	92 patients with T1DM or T2DM, recruited from diabetes education classes at a hospital [High HL group: mean age = 58.2 years; 58.6% females; Race: 36.2% white, 60% black, & 2.9% other; Mean education years = 14. Low HL group: mean age =	s-TOFHLA Categorical (2 levels: adequate vs. limited) [Functional HL]	Glycemic control (A1c), Self-care (SDSCA), diabetes Knowledge (DKQ)	Paired t-test, ANCOVA	Age, years of education, income	HL was associated with diabetes knowledge, but not with A1c or self-care	Good

		67.2; 81% females; Race: 20% white, 75% black, & 5% other; Mean education years = 10.2]						
Laramee et al., 2007 USA [35]	Cross-sectional	998 patients with T1DM or T2DM recruited from family medicine and general internal medicine clinics [Mean age (range): 65(22-93); 54% females; Race: 97% white & 3% other races; Education: 75% high school graduate & 25% other educational level]	s-TOFHLA Continuous [Functional HL]	Heart failure	Logistic regression	Age, sex, race, income, marital status, health insurance	HL was not associated with heart failure	Good
Lyles et al., 2011 USA ω [37]	Cross-sectional	17,795 patients with T1DM or T2DM from Kaiser Permanente Northern California [Age: 21% < 50 years, 50% 50-64 years, 29% ≥ 65 years; 49% females; Race: 20% Black, 23% Latino, 27% White, 13% Asian, 11% Filipino, 7% other races; Education: 46% ≤ high school, 24% some college, 29% ≥ college]	3-brief* SQs (problem learning, needing help reading, not confident with forms) [Functional HL]	Healthcare and general discrimination	Logistic regression	Age, sex, race, educational level, income, English proficiency	HL was associated with healthcare and general discrimination	Good
Mancuso, 2010 USA [14]	Cross-sectional	102 patients with T1DM or T2DM recruited from two urban primary care clinics [Mean age (SD): 52(9.1); 60.8% females; Race: 12.7% non-Hispanic Caucasian, 79.4% non-Hispanic Black/African American, 5.9%	TOFHLA Continuous [Functional HL]	Glycemic control (A1c), diabetes knowledge (DKT), self-care (SDSCA), depression (CES-D), patient	Spearman correlation Multiple linear regression	Patient trust, depression, diabetes knowledge, performance of self-care activities	Hl was associated with diabetes knowledge, but not with A1c, self-care, or depression symptoms	Good

		Hispanic/Latino American, 2% other races; Education: 70.6% ≤ high school, 29.4% > than high school]		trust (HCR Trust Scale)				
Mbaezue et al., 2010 USA [50]	Cross-sectional	189 patients with T1DM or T2DM recruited from one large hospital-based clinic [Mean age (SD): 51.2 (10); 58.7% females; Race: 94.7% African American & 5.3% white; Education: 32.3% <high school; 33.9% high school graduate & 33.9% some college]	s-TOFHLA Categorical (2 levels: inadequate vs. adequate) [Functional HL]	SMBG (Diabetes Care Profile)	Logistic regression	Age, education, insurance status, diabetes duration	HL was not associated with SBMG	Good
McCleary-Jones, 2011 USA [21]	Cross-sectional	50 patients with T1DM or T2DM recruited from a community health center [Mean age (SD): 58.6 (11.5); 76% females; Education: 6%: 7-11 grade, 28%: HS/GED, 24% some college, 20% college grad, 22% grad degree]	REALM Continuous [Functional]	Diabetes knowledge (DKT), self-care (SDSCA)	Pearson correlation	None	HL was associated with diabetes knowledge, but not with self-efficacy	Fair-good
				Self-efficacy (Diabetes Self-efficacy Scale)	Multiple linear regression	Diabetes knowledge, self-efficacy	HL was not associated with self-care	
Morris et al., 2006 USA [22]	Cross-sectional	1002 patients with T1DM or T2DM recruited from the Vermont Diabetes Information System [Median age (IQR): 66 (57-74); 54% females; Race: 97% white; Education: 25% ≤ high school, 36% high school, 31% some college or graduated college,	s-TOFHLA Categorical (3 levels: inadequate, marginal, adequate) & continuous [Functional HL]	Glycemic control (A1c), LDL, blood pressure, self-reported diabetes complications	Multiple linear regression, logistic regression	Age, sex, race, marital status, insurance, income, duration of diabetes, diabetes education,	HL was not associated with A1c, LDL, blood pressure, or self-reported diabetes complications	Good

		& 9% had graduate education]				depression, alcohol use, medication use		
Osborn et al., 2010 USA Ψ [39]	Cross-sectional	130 patients with T2DM, recruited from an outpatient primary care clinic [Mean age (SD): 62.7 (11.8); 72.5% females; Race: 28.6% non-Hispanic white & 71.4% non-Hispanic black; Mean education years (SD): 12.4 (5.2)]	REALM-R Categorical (2 levels: low vs. high) [Functional HL]	Glycemic control (A1c), self-care (SDSCA),	SEM	Diabetes Knowledge, diabetes fatalism, social support	There was no direct association between HL and diabetes self-care or A1c. HL had an indirect effect on diabetes self-care and A1c through social support	Good
Powell et al., 2007 USA [10]	Cross-sectional	68 patients with T2DM recruited from an academic general internal medicine clinic [Median age (IQR): 55 (51-60); 79.4% females; Race: 66.2% African American & 33.8% other races; Education: 4.4% < 4^{th} grade, 10.3% 4^{th} – 6^{th} grade, 13.2% 7^{th} – 8^{th} grade, 72.1% high school]	REALM Categorical (4 levels) [Functional HL]	Glycemic control (A1c), diabetes knowledge (DKT), diabetes-related health beliefs (DHBM)	one-way ANOVA Multiple linear regression	Age, race, educational level, treatment regimen	HL was associated with A1c and diabetes knowledge, but not with diabetes-related health beliefs	Good
Rothman et al., 2004 USA [17]	Longitudi-nal	217 patients with T2DM recruited from the US academic general internal medicine practice [Control group Low HL mean age =	REALM Categorical (2 levels: low vs. high) [Functional HL]	Achievement of goal A1c levels	Logistic regression	Age, sex, race, income, insulin status at	HL was an effect modifier for reaching goal levels of A1c	Good

		59(10.4) years; 53% females; 68% African American; 82% < high school education. Control group High HL mean age = 56(10.9) years; 58% females; 55% African American; 26% < high school education. Intervention group Low HL mean age = 57(10.5) years; 55% females; 94% African American; 82% < high school education. Intervention group High HL mean age= 51(13.1) years; 65% females; 51% African American; 59% < high school education]				enrollment, duration of disease, baseline A1c		
Sarkar et al., 2010 USA ω [11]	Cross-sectional	14,357 patients with T2DM recruited from Kaiser Permanente Northern California [Mean age (SD): 58 (10); 49% females; Race: 17% African American, 22% non-Hispanic white, 18% Latino, 23% Asian, & 20% other races]	3-brief* SQs (problem learning, needing help reading, not confident with forms) [Functional HL]	Patient reported frequency of significant hypoglycemia	Logistic regression	Age, sex, race, English proficiency, medication type, diabetes duration, A1c, glomerular filtration rate, income, dementia, and history of stroke	HL was associated with significant hypoglycemia	Good

Sarkar et al., 2010 USA ω [33]	Cross-sectional	14,102 patients with T1DM or T2DM from Kaiser Permanente Northern California [Mean age: 59; 49% females; Race: 21% African-American, 28% non-Hispanic white, 14% Latino, 9% Asian, 12% Filipino, & 17% other races; Education: 10% no degree, 30% high school, 27% some college, & 33% college grad/post grad]	3-brief* SQs (problem learning, needing help reading, not confident with forms) [Functional HL]	Use of internet portal	Logistic regression	Age, sex, race, educational attainment, income	HL was associated with the use of the internet portal	Good
Schillinger et al., 2002 USA ψ [12]	Cross-sectional	408 patients with T2DM recruited from two primary care clinics at a US public hospital [Mean age (SD): 58.1 (11.4); 58% females; Race: 18% Asian, 25% Black, 42% Latino; & 15% White; Education: 46% ≤ high school, 23% = high school, 28% some college &/or technical school, & 3% had a graduate degree]	s-TOFHLA Continuous [Functional HL]	Glycemic control (A1c), self-reported diabetes complications	Multiple linear regression, logistic regression	Age, sex, race, education, language, insurance, social support, depression, treatment regimen, diabetes duration, diabetes education	HL was associated with A1c and self-reported complications	Good
Schillinger et al., 2003 USA ψ [51]	Cross-sectional	61 patients with T2DM recruited from two primary care clinics at a US public hospital [Age: 55.7% ≥ 65 years 44.3% < 65 years; 49% females; Race: 55.7% African	s-TOFHLA Categorical (2 levels: inadequate vs. marginal) [Functional HL]	Glycemic control (A1c)	Logistic regression (using data for patients with low and marginal HL	Age, sex, race, diabetes duration, insulin use, number of	HL was associated with glycemic control	Good

		American & 44.3% non-African American]			only)	new concepts conveyed		
Schillinger et al., 2004 USA ψ [52]	Cross-sectional	408 patients with T2DM recruited from two primary care clinics at a US public hospital [Mean age: 58.1; 58% females; Race: 18% Asian, 25% Black, 42% Latino; & 15% White; Education: 46% ≤ high school, 23% = high school, 28% some college &/or technical school, & 3% had a graduate degree]	s-TOFHLA Categorical (3 levels: inadequate, marginal, adequate) [Functional HL]	Quality of physician-patient communication (IPC)	Logistic regression (one model for people with inadequate HL & one for people with adequate HL)	Age, sex, race, education, language, insurance, treatment regimen, A1c, depression score, diabetes duration, Spanish fluency, length of time in the physician's care	HL was associated with explanatory/ participatory dimensions of patient-physician communication, but not with the listening dimensions	Good
Schillinger et al., 2006 USA ψ [53]	Cross-sectional	395 patients with T2DM recruited from two primary care clinics at a US public hospital [Mean age (SD): 57.9(11.4); Race: 18.5% Asian/Pacific Islander, 25.3% black, 42.3% Hispanic, & 13.9% white; Education: 46.8% ≤ high school, 24.1% = high school, & 29.1% some college &/or technical school]	s-TOFHLA Continuous [Functional HL]	Glycemic control (A1c)	SEM	Age, race, educational level, primary language, insurance status	HL was associated with A1c	Good

Tang et al., 2007 China [19]	Cross-sectional	149 patients with T2DM recruited from a diabetes education management centre of a public hospital [Mean Age (range): 59.8 (27-90); 45.6% females; Education: 55.8% ≤ primary education, 39.6%: secondary education, & 4.7% ≥ college education]	s-TOFHLA Continuous [Functional HL]	Glycemic control (A1c)	Spearman correlation Multiple linear regression	Sex, insurance status, duration of diabetes, patient awareness score, C-SDSCA score	HL was associated with A1c	Good
Thabit et al., 2009 Ireland [54]	Cross-sectional	100 patients with T2DM recruited from diabetes outpatient service [Immigrant patients: Mean age (SD): 45.8(11.8); Irish patients: Mean age (SD): 60.1(11)]	REALM Continuous [Functional HL]	Glycemic control (A1c)	Pairwise correlation (N=50)	None	HL was associated with A1c	Fair-good
Wallace et al., 2010 USA [30]	Cross-sectional	208 patients with T2DM recruited from a diabetes management program at an internal medicine practice [Mean age (range): 58 (23-85); 64% females; Race: 48% African American; Education: 34% < high school]	s-TOFHLA Categorical (2 levels: Adequate vs. inadequate), & continuous [Functional HL]	Self-management support (PACIC)	Pearson correlation Multiple linear regression	Sex, race, duration of diabetes	HL was associated with self management support	Good
Williams et al., 1998 USA [27]	Cross-sectional	114 patients with T1DM or T2DM recruited from 2 urban public hospitals [Inadequate HL group: Mean age (SD): 57.4(9.3); 76% females; Race: 18% black, 2% white, 80% Latino; Years of schooling: 78% ≤ 6 years, 16% 7-11	TOFHLA Categorical (3 levels: inadequate, marginal, adequate) [Functional HL]	Glycemic control (A1c), blood pressure, diabetes knowledge	Multiple linear regression, Chi-square test	None	HL was associated with diabetes knowledge, but not with A1c or blood pressure	Good

		years, 4% 12 years, & 2% > high school; Marginal HL group: Mean age (SD): 53.2(8.8); 69% females; Race: 31% black, 69% Latino; Years of schooling: 39% ≤ 6 years, 39% 7-11 years, 15% 12 years, & 8% > high school; Adequate HL group: Mean age (SD): 49.2(10.3); 67% females; Race: 37% black, 33% white, 29% Latino; Years of schooling: 2% ≤ 6 years, 29% 7-11 years, 37% 12 years, & 31% > high school]						

† Variables used in multivariate analysis, or adjusted for in bivariate analysis; if both analyses were performed, the reported variables are the ones included in multivariate analysis.

‡ Adjusted analyses results are reported. In case adjustment was not performed, un-adjusted analyses findings are reported. If both bivariate and multivariate analysis were done; findings based on multivariate analysis were reported.

Ψ, ψ, ω Indicates that the identified record is one of multiple publications of the same study; publications that have the same symbol are from the same study.

* Each question is analyzed separately as a categorical variable (2 levels). ** Each subscales is analyzed separately as a categorical variable (2 levels). *** Self-care was assessed by asking "overall, what grade would you give yourself on your diabetes self-care in the past 6 months?"; response categories ranges from 1 to 9.

SD: Standard Deviation; OR: Odds Ratio; IQR: Interquartile range; SEM: Structural Equation Modeling; T1DM: Type 1 Diabetes Mellitus; T2DM: Type 2 Diabetes Mellitus; HL: Health Literacy; NR: Not Reported; BMQ: Beliefs about Medications Questionnaire; C-SDSCA: Chinese – Summary of Diabetes Self Care Activities; CESD-10: Center for Epidemiologic Studies Depression Scale-10; DBP: Diastolic Blood Pressure; DFS: Diabetes Fatalism Scale; DHBM: Diabetes Health Belief Model; DKQ: Diabetes Knowledge Questionnaire; DKT: Diabetes Knowledge Test; DMSES: Diabetes Management Self-Efficacy Scale; DPMD: Desire to Participate in Medical Decision making; DSHLM: Diabetes specific health literacy measure; FPI: Facilitation of Patient Involvement; HCR: Health Care Relationship Trust Scale; HDL: High Density Lipoprotein; IPC: Interpersonal Processes of Care in Diverse Populations Questionnaire; KDQ: Knowledge Diabetes Questionnaire; LDL: Low Density Lipoprotein; MOS: Medical Outcomes Study; PACIC: Patient Assessment of Chronic Illness Care; PFI: Facilitation of Patient Involvement scale; REALM-R: Rapid Estimate of Adult Literacy in Medicine- Revised; RIAS: Roter Interaction Analysis System; s-TOFHLA: Test of Functional Health Literacy in Adults – Short form; 3-brief SQs: 3 brief screening questions; SBP: Systolic Blood Pressure; SCQ: Self-administered Comorbidity Questionnaire; SDSCA: Summary of Diabetes Self Care Activities; SKILLD: Spoken Knowledge in Low Literacy in Diabetes Scale; SMBG: Self-Monitoring of Blood Glucose; TIBI: Total Illness Burden Index; WFPTS: Wake Forest Physician Trust Scale; TOFHLA: Test of Functional Health Literacy in Adults

Table S2-2: Overview of Health Literacy and Numeracy Studies

Author(s), year Source	**Design**	**Participants' characteristics N, diabetes type, recruitment site [age, sex, race, educational level]**	**Numeracy assessment**	**HL measure (levels) / [domain addressed]**	**Outcomes**	**Type of analysis performed**	**Covariates†:**	**Main findings‡**	**Quality Rating**
Cavanaugh et al., 2008 USA [38]	Cross-sectional	398 patients with T1DM or T2DM recruited from 2 primary care and 2 diabetes clinics at 3 medical centers [Median age (IQR): 55 (46-64); 51% females; Race: 63% White & 37% nonwhite; Education: 43% ≤ high school & 57% > high school]	DNT (Continuous) WRAT-3R (Categorical: 2 levels)	REALM Categorical (2 levels: high vs. low) [Functional HL]	Glycemic control (A1c)	Multiple linear regression	Age, sex, race, annual income, type of diabetes, diabetes duration, clinic site	Diabetes numeracy was associated with A1c; HL was not associated with A1c	Good
Mayberry et al., 2011 USA [34]	Cross-sectional	75 patients with T2DM recruited from a primary care clinic at an academic medical center [Mean age (SD): 56.9 (8.8); 68% females; Race: 65% white & 35% non-white]	SNS	3-brief* SQs (problem learning, needing help reading, not confident with forms) [Functional HL]	Use of computer and patient web portal (PWP)	Spearman correlation	None	HL was associated with computer use but not with PWP, and numeracy was not associated with either	Fair-good
Osborn et al., 2009 USA Ψ [20]	Cross-sectional	383 patients with T2DM recruited from primary care clinics at 3 medical centers [Median age (IQR): 56(47-64); 50% females; Race: 65% white, & 35% nonwhite; Education: 43% < high school & 57% ≥ high school]	DNT (Continuous) WRAT-3 (Continuous)	REALM-R Continuous [Functional HL]	Glycemic control (A1c)	SEM	Age, sex, race, educational level, income, diabetes type, diabetes	Diabetes-related numeracy was associated with A1c, HL and general numeracy were not associated	Good

							duration, insulin use	with A1c	
Osborn et al., 2010 USA Ψ [28]	Cross-sectional	383 patients with T1DM or T2DM recruited from 2 primary care clinics and 2 diabetes clinics at 3 medical centers [Mean age (SD): 54.4(13); 50% females; Race: 65% white, & 35% nonwhite; Education: 8% < 9th grade, 35% 10th grade-high school, 57% > high school]	WRAT-3R (Continuous)	REALM-R Continuous [Functional HL]	Glycemic control (A1c), self-efficacy (PDSMS)	SEM	Age, sex, race, educational level, income, diabetes type, diabetes duration, insulin use	Numeracy, but not HL, was directly associated with self efficacy and indirectly associated with A1c	Good
White et al., 2011 USA [40]	Cross-sectional	144 patients with T1DM or T2DM recruited from two primary care clinics [Mean age (SD): 47.8 (12.1); 62% females; Education: 56% < high school, 17%: some high school, 18%: high school graduate, 9%: some college or more]	DNT-15 (continuous) WRAT-4 (continuous)	s-TOFHLA Categorical (3 levels: inadequate, marginal, adequate) [Functional HL]	Glycemic control (A1c), self-efficacy (PDSMS), self-care (SDSCA),	Spearman correlation	None	Numeracy was not associated with A1c, self-care, or self-efficacy	Good

* Variables used in multivariate analysis, or adjusted for in bivariate analysis; if both analyses were performed, the reported variables are the ones included in multivariate analysis.
† Adjusted analyses results are reported. In case adjustment was not performed, un-adjusted analyses findings are reported. If both bivariate and multivariate analysis were done; findings based on multivariate analysis were reported.
Ψ, ψ, ω Indicate that the identified record is one of multiple publications of the same study; publications that have the same symbol are from the same study.
SD: Standard Deviation; IQR: interquartile range; T1DM: Type 1 Diabetes Mellitus; T2DM: Type 2 Diabetes Mellitus; HL: Health Literacy; DKT: Diabetes Knowledge Test; DNT: Diabetes Numeracy Test; PDSS/PDSMS: Perceived Diabetes Self-Management Scale; REALM: Rapid Estimate of Adult Literacy in Medicine; s-TOFHLA: Test of Functional Health Literacy in Adults – Short form; 3-brief SQs: 3 brief screening questions; SDSCA: Summary of Diabetes Self Care Activities; WRAT: Wide Range Achievement; SNS: Subjective Numeracy Scale; NR: Not Reported

Chapter 3

Measuring Health Literacy in Individuals with Diabetes: A Systematic Review and Evaluation of Available Measures [*]

Abstract

Objective: To identify instruments used to measure health literacy (HL) and numeracy in people with diabetes, evaluate their use, measurement scope and properties, discuss their strengths and weaknesses, and propose the most useful, reliable, and applicable measure for use in research and practice settings.

Methods: A systematic literature review was conducted to identify the instruments. Nutbeam's domains of HL and a diabetes HL skill set were used to evaluate the measurement scope of the identified instruments and to evaluate their applicability in people with diabetes.

Results: 56 studies were included, from which one diabetes-specific (LAD) and eight generic measures of HL (REALM, REALM-R, TOFHLA, s-TOFHLA, NVS, 3-brief SQ, 3-level HL Scale, SILS) and one diabetes-specific (DNT) and two generic measures of numeracy (SNS, WRAT) were identified. These instruments were categorized into direct measures i.e. instruments that assess the performance of individuals on HL skills, and indirect measures that rely on self-report of these skills. The most commonly used instruments measure selective domains of HL, focus mainly on reading and writing skills, and do not address other important skills such as verbal communication, health care system navigation, health-related decision making, and numeracy. The structure, mode and length of administration, and measurement properties were found to affect

---

[*] A version of this chapter has been published. Al Sayah F, Williams B, Johnson JA. Measuring health literacy in individuals with diabetes: a systematic review and evaluation of available measures. *Health Education and Behavior* 2013 Feb; 40(1): 42-55.

the applicability of these instruments in clinical and research settings. Indirect self- or clinician-administered measures are the most useful in both clinical and research settings.

Conclusion: This review provides an evaluation of available HL measures and guidance to practitioners and researchers for selecting the appropriate measures for use in clinical settings and research applications.

3.1. Introduction

Health literacy (HL) is a set of skills that people need to function effectively in the healthcare environment [1]. These skills include the ability to read and understand written text, locate and interpret information in documents, and write or complete forms (functional); the ability to speak and listen effectively and communicate about health-related information (interactive); the ability to navigate the healthcare system and make appropriate health decisions (critical); and the ability to use numeric information for tasks, such as interpreting medication dosages, food labels and blood glucose measurements (numeracy) [1, 2].

Diabetes is a prototypical multifactorial chronic condition, characterized by a high level of complexity that requires extensive self-care education and management. The demands on individuals with diabetes is complicated by the fact that diabetes self-management often relies on printed educational material and verbal instructions and requires advanced HL skills [3]. With the emerging evidence on the adverse effects of inadequate HL on health care and outcomes in people with diabetes [4-7], the assessment of HL skills is becoming crucial in this population. Since individuals often read several grade levels lower than the highest grade achieved in school [8], educational attainment cannot be used as a proxy for HL, which made the development of HL measures a necessity.

As a result of this, several instruments were developed to assess skills or screen for inadequate HL [9-13]. These instruments vary in their development, structure, measurement scope, and subsequently psychometric properties. Although these instruments have been used with several patient populations, their usefulness and applicability for people with diabetes remains challenging.

The reason is the complexity of tasks and skills that are required by people who have diabetes, and the postulation that the available instruments do not address that complexity and all of the important components of HL altogether. Additionally, the continuous adjustment of the meaning and components of HL makes the available instruments questionable in what they actually measure. We therefore lack an understanding of the characteristics and measurement scope of HL measures used in people with diabetes, and their applicability in this population. In this review, we identified instruments used to measure HL in individuals with diabetes, evaluated their measurement scope and properties, and proposed recommendations for their use and applicability in different settings.

3.2. Methods

3.2.1. Data sources, search strategy and study selection

We conducted a systematic review of nine databases: Medline, PubMed, CINAHL, SCOPUS, Web of Science, ERIC, Nursing and Allied Health Source, Health Source (Nursing/Academic Edition), and Health and Psychosocial instruments. The searches were not limited to any time period, language, or type of published paper. The following search terms were used to identify eligible studies: literacy, numeracy, HL, diabetes, diabetic, type 2 diabetes, type 1 diabetes, instrument, measure, assessment, questionnaire, survey and screening. Keywords were matched to database specific indexing terms. The names of the identified instruments were used as keywords in further searches of databases. Electronic searches were supplemented by hand searching “HL” specific issues in journals and cross-referencing with identified articles (detailed information about the search strategy are available upon request). The search strategy and retrieval process are displayed in figure 3-1.

3.2.2. Data Extraction and Synthesis

Data on the general characteristics of the studies including design, methods, sample characteristics, and HL measures used in each study were extracted (Table S3-1). From these studies, we identified the HL instruments used for people with diabetes (Table 3-1). Data on the measurement scope and psychometric properties of the identified instruments were also extracted (Table 3-1 and Table 3-2).

Evaluation of the measurement scope and applicability of the identified instruments was based on Nutbeam's three components of HL (functional, interactive, critical) [2], and a diabetes-HL skill set. Functional HL focuses on reading and writing skills that enable an individual to function effectively in everyday situations; interactive (or communicative) HL includes advanced skills that allow a person to extract information, derive meaning from different forms of communication, and apply new information to changing circumstances; and critical HL encompasses more advanced skills for critically analyzing information and using information to exert greater control over life events and situations [2]. The HL skill set was developed based on a brief review of the literature on diabetes self-care and management, and it includes: 1) reading and understanding medication labels, information on medication bottles, blood glucose levels, insulin bottles and pens, and applying this information in taking medication and/or insulin; 2) reading and understanding diabetes education materials and apply the information and instructions in daily life activities; 3) understanding health care providers' instructions and applying it to daily living, such as diet management, physical activity, smoking cessation, monitoring of blood glucose, and self assessment; 4) completing medical forms, glucose monitoring logs, and dietary logs; 5) communicating with health care providers, explaining health concerns, asking questions, and obtaining needed information; 6) navigating the health care system; and 7) making appropriate health-related decisions. A score was assigned to each section based on the extent to which each domain and skill were addressed by the instrument (Table 3-2). Ratings were as follows: 0 = not addressed, 1 = partially addressed, 2 = fully addressed. Two reviewers independently rated the instruments. Cohen's kappa was used to assess inter-rater reliability in the ratings between the two reviewers. Scores were used to help in the qualitative comparison of the instruments, and not to rank the instruments with respect to ratings.

3.3. Results

3.3.1. Literature Search and Identification of Instruments

Across all databases, the search yielded 1120 publications. After removal of duplicates, 412 remained. Titles and abstracts were screened for relevance, and based on that, 161 publications were included for full-text review. The full texts of the 161 publications were screened for eligibility, and 53 publications that involved the use of a HL measure in individuals with diabetes were included. Only articles written in English were identified in this review. The hand search of reference lists of included studies resulted in the inclusion of three additional publications. Thus, the total number of eligible studies was 56.

We identified one diabetes-specific and eight generic measures of HL (Table 3-1): Literacy Assessment in Diabetes (LAD), Rapid Estimate of Adult Literacy in Medicine (the original REALM and the revised form REALM-R), Test of Functional Health Literacy in Adults (TOFHLA and the shorter form s-TOFHLA), Newest Vital Sign (NVS), 3-brief Screening Questions (3-brief SQ), the 3-level health literacy scale (3-level HL scale), and Single Item Literacy Screener (SILS). In addition, we identified one diabetes-specific and two generic measures of numeracy: Diabetes Numeracy Test (DNT; 15-item & 43-item versions), Wide Range Achievement Test (WRAT; the 3-item version WRAT-3 and the revised version WRAT-R), and the Subjective Numeracy Scale (SNS). In assessments of HL in individuals with diabetes, the s-TOFHLA was the most commonly used (26 studies), followed by the REALM (15 studies), 3-brief SQ (7 studies), TOFHLA (4 studies), 3-level HL scale (3 studies), SILS (2 studies), LAD (2 studies), and REALM-R (2 studies), NVS (1 study), WRAT (4 studies), DNT (3 studies), and SNS (1 study).

Several other instruments have been developed to measure general HL, such as the Medical Term Recognition Test (METER) [14], the News Skills Based Instrument [15], the Mandarin Health Literacy Scale (MHLS) [16], but their use in people with diabetes was not reported, and thus not included in our review. The “Diabetes Specific Health Literacy Measure” (DSHLM) [17] was presented as a HL measure, but in fact resembles a diabetes knowledge test and thus was not considered a HL measure in this review.

3.3.2. Description of the Identified Instruments

The identified instruments were categorized into those instruments that directly measure HL skills i.e. assessment of the performance of these skills and these include: REALM, REALM-R, TOFHLA, s-TOFHLA, NVS, and LAD, and instruments that indirectly assess HL skills i.e. instruments that rely on self-report of these skills, and these include: 3-brief SQ, 3-level HL Scale, and SILS. Additionally numeracy instruments were also categorized into direct measures and these include: DNT and WRAT, and indirect measures including the SNS (Table 3-1). There are considerable differences among the identified instruments used to measure HL and numeracy. These instruments vary in their structure, number of items/questions/domains, administration time and mode, scoring system, available languages, and their measurement scope and properties, which entails a number of implications for their use in research and clinical settings. Since the WRAT was not developed to be a measure of the numeracy component of HL, it was not evaluated based on the identified criteria, nor compared with other instruments.

3.3.3. Measurement Properties of the Identified Instruments

The REALM, TOFHLA, and s-TOFHLA have been validated in several populations (including people with diabetes) and have been used in most of the validation studies of newer HL instruments such as the 3-brief SQ, NVS and the SILS (Table 3-1). Among the identified instruments, the TOFHLA, s-TOFHLA, and the REALM have established their reliability and validity through several applications in different patient groups including the diabetes population. However, these instruments differ in their measurement scope and underlying constructs from newer instruments, and thus, their typical use as a gold standard in validation studies might be reconsidered. For example, the 16-brief Screening Questions (16-brief SQs) [10], which is the long version of the 3-brief SQ [11], were comprehensively developed to address all domains of HL, and were validated against the s-TOFHLA, which only measures functional HL. For instance, the 3-brief SQs performed the best based on the s-TOFHLA in screening for inadequate "functional" HL [10, 11]. This does not imply that the other questions of the 16-brief SQs that address interactive and critical HL have poor measurement properties; however, it does suggest that the validity of these questions should be evaluated against an instrument that measures the same underlying concepts. Further validation of the 16-brief SQ is therefore recommended. Similarly, the 3-level HL Scale which addresses all aspects of HL, except numeracy, was also compared to the REALM and TOFHLA as 'gold standards', where its functional domain performed the best when compared to these instruments [18, 19]. This applies to numeracy measurement instruments as well. Careful attention should be given to revising and developing HL and numeracy instruments, particularly with respect to what these instruments are constructed to measure, and accordingly demonstrating their performance using appropriate 'gold standards' or comparisons.

3.3.4. Measurement Scope of the Identified Instruments

Cohen's kappa between the two reviewers for the rating of measurement scope was 0.78. Based on the evaluation of the measurement scope of the identified instruments, we found that the most commonly used instruments (s-TOFHLA, REALM) are not sufficiently comprehensive, i.e. they measure selective domains of HL namely reading comprehension and writing ability, thereby tackling only functional aspects of HL (Table 3-2). Other instruments (3-brief SQ, 3-level HL Scale, NVS) address functional HL as well as critical HL such as decision-making, navigating the health care system, and following instructions and applying health information to daily life situations (Table 3-2). The 3-level HL Scale was the only instrument that was found to address interactive HL in addition to functional and critical HL (Table 3-2). The TOFHLA and NVS measure computational skills and thus address the numeracy component of HL. DNT focuses on diabetes numeracy in addition to assessing functional aspects of HL, where SNS only focuses on general numeracy.

The 3-level HL scale appears to have the broadest measurement scope and the one that addresses all the identified skills and the functional, interactive and critical aspects of HL but not numeracy. However, this instrument was developed in Japanese and was not validated in the English language, and it does not have a brief version that would be more applicable than the long version in clinical settings. Overall, the identified instruments varied widely in their measurement scope and the component of HL they measure.

3.4. Discussion

Although their measurement scope is limited to aspects of functional HL, the REALM and the s-TOFHLA were found to be the most commonly used instruments to measure HL amongst individuals with diabetes; however, due to their limited measurement scope and properties, their typical use as a gold standard in validation studies of other instruments might be reconsidered. In a study that compared the estimates of poor HL using the s-TOFHLA and the

REALM [20], Griffin and colleagues reported that estimates of poor HL varied by the assessment tool used, especially after adjusting for non-response bias. The reason for this discrepancy in the estimates could be due to the fact that these instruments measure different aspects of HL and thus reflect different entities. This should always be considered while selecting an instrument for the assessment of HL and in validation studies of other HL instruments.

The identified instruments have inherent strengths and weaknesses as a result of their structure, measurement scope and properties. ***First***, the REALM (all versions), TOFHLA/s-TOFHLA, NVS, and the LAD were designed to directly measure specific skills. The fact that these measures directly assess the skill level of individuals imposes many limitations on the applicability of these measures especially in clinical setting, where this approach might impose discomfort and embarrassment particularly for those who have inadequate HL skills [1, 21, 22]. This also applies to DNT, which is a direct measure of numeracy compared to the SNS as an indirect less burdensome measure of numeracy.

Second, direct measures require good visual acuity (particularly word recognition tests), good writing ability (such as the TOFHLA, NVS, DNT), and enough concentration to be able to complete the test. These limitations make direct measures less convenient for most clinical settings, and for survey-based research. On the other hand, indirect or self-reported HL measures (3-brief SQ, 3-level HL scale, SILS, SNS) provide information about confidence with certain skills without directly assessing these skills, and therefore they are less burdensome and do not impose discomfort and embarrassment, which makes them more suitable for most clinical settings and research applications.

Third, the mode of administration of identified instruments plays an important role in their applicability and use. The REALM (all versions), the LAD, and part of the DNT are word-recognition tests and are only administered by a clinician or researcher. This limits their use for research purposes particularly in survey-based studies, and makes them less practical for most clinical settings as they would require time from the care provider and could impose discomfort and embarrassment. On the other hand, instruments that are only self-administered (TOFHLA/s-TOFHLA, 3-level HL Scale, SNS) do not impose a lot of discomfort, but may have limited use since they require good visual acuity and writing skills. The 3-brief SQ, 3-level HL scale, SILS, and SNS could be self-administered or clinician/researcher-administered, which provides flexibility in their application for research purposes and in most clinical settings.

Fourth, administration time is also a factor that affects the use and applicability of the identified instruments. The TOFHLA, DNT, and WRAT require a long administration time, which makes them less practical for use in most clinical settings. The administration times of the REALM (all versions), s-TOFHLA, NVS, LAD, 3-brief SQ, the 3-level HL scale, SILS and SNS are relatively short, making these instruments useful in research and most clinical settings. It could be useful to use a briefer or shortened version of an instrument due to potential time constraints in a specific application or setting; however, it is important to recognize that there could be a trade-off with measurement scope, where briefer versions usually have a narrower scope than longer ones (16-brief SQ vs. 3-brief SQ, TOFHLA vs. s-TOFHLA).

Finally, it is important to note that direct measures (REALM, s-TOFHLA, NVS, LAD, DNT) use terms from the medical field and texts from real medical forms used in clinical settings, which implies that these instruments measure HL based on the health system demands of skill level. However, indirect measures (3-brief SQs, 3-level HL scale, SILS, SNS) assess HL skills by asking about personal abilities that are not related to specific medical forms or context. In other words, the level of HL skills required by medical forms and texts that are part of the instrument influence the HL score of individuals; individuals would score higher as the level of HL skills required by the forms is lower. This has a direct implication on measuring HL, where it was reported that measuring HL using different instruments yields different estimates [20].

Considering the measurement scope of these instruments, their psychometric properties, and their strengths and limitations collectively, the 3-level HL can be considered the most useful and comprehensive instrument to screen for inadequate HL. However, this instrument has not been validated in English. For English-based instruments, the 3-brief SQs (and their longer version 16-brief SQs) have the broadest measurement scope, demonstrated good measurement properties, have many advantages over other instruments, and could be considered the best available instrument to measure functional HL. The SNS, although minimally used, has good characteristics that make it very applicable; however, it requires further testing and validation with people who have diabetes.

With the escalating evidence on the adverse effects of inadequate HL on health outcomes in people with diabetes [5, 7, 23, 24], measuring HL skills is becoming imperative. Health care professionals tend to overestimate patients' literacy because they are very accustomed to the medical field and its terminology and because some patients who have inadequate HL skills often deny or conceal their deficit. Additionally, patients are often ashamed of their low HL, and many adults will attempt to conceal their reading impairments from others [21, 25]. For that, understanding the components of HL measurement and screening in general and in people with diabetes in particular is crucial to the planning and delivery of comprehensive individualized diabetes care and interventions.

The findings of this review are applicable to other chronic conditions with similar HL demands on individuals. Additionally, this review did not only address the applicability and usefulness of these instruments in individuals with diabetes, but also provided an evaluation of these instruments and their strengths and weaknesses, which are transferable to determining their applicability in other health conditions and situations. Researchers and clinicians could use this review as a guide to the selection of the most suitable instrument for a particular research application or in clinical settings.

3.5. Conclusion

We based this evaluation on the most collective and comprehensive description of HL; however, without a final consensus on what the underlying constructs of HL are, we will continue to fail in using and developing adequate measures of HL and in conducting valid measurements. This evaluation of HL instruments used in people with diabetes showed that the most commonly used instruments measure selective domains of HL and are not sufficiently comprehensive. Each of the identified instruments has strengths and limitations in its measurement scope, properties, applicability, and feasibility. It appears that indirect self- or clinician-administered measures are the most useful in both clinical and research settings. We found that the 3-level HL Scale [18] and the 3-brief SQs [11] as the most comprehensive, applicable and useful among the available instruments of HL measurement in individuals with diabetes.

Tables and Figures

Table 3-1: Characteristics of Identified Health Literacy and Numeracy Instruments

Measure	Type of measure	Number of items/ questions	Administration mode/time	Scoring	Measurement Properties†‡	Languages available
Direct Health Literacy Measures						
REALM [74]	Word recognition and pronunciation test	66 words	Clinician/ researcher administered < 3 min	Grade is assigned based on total score that ranges from 0-66: 0-18 <= 3rd grade; 19-44 = 4th - 6th grade; 45-60 = 7th – 8th grade; 61-66 >= 9th grade	*Reliability:* Cronbach's alpha: 0.99; ICC: 0.92 *Validity:* 0.97 (with PIAT-R); 0.96 (with WRAT-R); 0.88 (with SORT-R)	English
REALM-R [75]	Word recognition and pronunciation test	8 words	Clinician/ researcher administered < 2 min	Grade is assigned based on total score that ranges from 0-8. Score ≤ 6 corresponds to 6th grade and indicates poor HL	*Reliability:* Cronbach's alpha: 0.91 (Bass et al., 2003) *Validity:* 0.72 (with REALM), 0.64 (with WRAT-R)	English
TOFHLA [13]	Reading comprehension and numeracy test	50-item (reading comprehension) 17-item (numerical ability test)	Self administered 22 min	Scores range 0-100: < 60 = inadequate HL; 60 – 75 = marginal HL; > 75 = adequate HL	*Reliability:* Cronbach's alpha: 0.98 *Validity:* 0.84 (with REALM); 0.74 (with WRAT-R)	English Spanish Chinese
s-TOFHLA [9]	Reading comprehension	36-item reading assessment	Self administered 7 min	Scores range 0-36: 0-16: inadequate HL; 17-22: marginal HL; 23-36: adequate HL	*Reliability:* Cronbach's alpha: 0.98 *Validity:* 0.91 (with TOFHLA); 0.80 (with REALM)	English Spanish Chinese

NVS [76]	Reading comprehension Numeracy	6 questions on a ice cream nutrition label	Self administered 3 min	Each item answered correctly is given a score of 1. Scores range: 1 – 6 (score < 4: limited HL)	*Reliability:* Cronbach's alpha: 0.76 *Validity:* 0.59 (with TOFHLA) *AUROC:* 0.88 (based on TOFHLA) *Screening sensitivity:* 72% (based on TOFHLA) *Screening specificity:* 87% (based on TOFHLA)	English Spanish
LAD [50]	Word recognition and pronunciation test	3-graded word lists including 60 words	Clinician/ researcher administered 3 min	Grade is assigned based on total score that ranges from 0-60: 0-20: < = 4^{th} grade; 21-40: 5^{th} – 9^{th} grade; 41-60: > = 9^{th} grade	*Reliability:* ICC: 0.86 *Validity:* 0.90 (with REALM); 0.81 (with WRAT-3)	English
Indirect Health Literacy Measures						
3-brief SQ [11]	Self report of confidence in HL skills	3 questions	Self administered or clinician/ researcher administered Time: 1-2 min	Values of 0, 1, 2, 3, & 4 are assigned to each response option for each question Score ranges from 0-12; High scores = high HL skills; Low scores = low HL skills	*AUROC:* 0.66-0.87 (based on s-TOFHLA) 0.72-0.84 (based on REALM) *Screening sensitivity:* 48-60 (based on s-TOFHLA) *Screening specificity:* 79-83 (based on s-TOFHLA)	English
3-level HL scale [18]	Self report of HL skills	14 items, 3 domains: functional (5 items), communicative (5 items), and critical (4 items)	Self administered or clinician/ researcher administered 5-6 min	Each item is scored on a 4-point scale ranging from 1 (never) to 4 (often) The scores of the items are summed up and divided by the number	*Reliability:* Cronbach's alpha: Overall scale: 0.78; Functional domain: 0.84; Communicative domain: 0.77; Critical domain: 0.65	Japanese

				of the items in the scale. Higher scores indicate higher health literacy level		
SILS [48]	Self report of HL skills	1 question	Self administered or clinician/ researcher administered Time: ½ min	Possible scores are: 0 (never), 1 (rarely), 2 (sometimes), 3 (often), 4 (always). Score > 2 = difficulty with reading printed health material	*AUROC:* 0.64-0.87 (based on s-TOFHLA) *Screening sensitivity:* 0.47-0.73 (based on s-TOFHLA) *Screening specificity:* 83 (based on s-TOFHLA)	English
Direct Numeracy Measures						
DNT [77]	Reading recognition, spelling & arithmetic computation	43 questions in 5 domains, & 8 math problem types	Researcher/ clinical or self-administered 30 min	Items are scored as binary outcomes – correct or incorrect – Scores are reported as percent correct (with a possible range of 0% to 100%	*Reliability:* Cronbach's alpha: 0.95 (43-item version) 0.90 (15-item version) *Validity:* 0.96 (DNT-15 with DNT-43)	English Spanish
WRAT [78]	Reading recognition, spelling & arithmetic computation	Reading test: 15 letters & 42 individual words; Spelling test: 13 letters & 40 words; Arithmetic test: two parts: Part I: counting, reading number symbols, & solving simple arithmetic	Self-administered 20-30 min	3 scores: spelling, arithmetic, and reading Scoring: 1 = correct answer, 0 = incorrect answer, scores are standardized, and the average score = 100	---	English

		problems. Part II: 40 arithmetic problems				
Indirect Numeracy Measures						
SNS [79]	Ability to perform mathematical tasks	8 questions	Researcher/ clinical or self-administered 2 min	Responses on 6-point Likert scale (from *Not at all good to Extremely good)* Score range: 1 – 6	*Validity* 0.63 – 0.68 (with Lipkus & other numeracy scales)	English

†All the reported estimates are statistically significant at P-value < 0.05

‡ The reported Validity parameters are either Pearson or Spearman correlations

PIAT-R: Peabody Individual Achievement Test-Revised; WRAT-R: Wide Range Achievement Test-Revised; SORT-R: Slosson Oral Reading Test-Revised; ICC: intraclass correlation coefficient; AUROC: Area Under the Receiver Operating Characteristic (ROC) curve; REALM: Rapid Estimate of Adult Literacy in Medicine; REALM-R: Rapid Estimate of Adult Literacy in Medicine-Revised; TOFHLA: Test of Functional Health Literacy in Adults; s-TOFHLA: Test of Functional Health Literacy in Adults-Short Form; 3-brief SQ; 3 brief screening questions; SILS: Single Item Literacy Screener; LAD: Literacy assessment in Diabetes; NVS: Newest Vital Sign; DNT: Diabetes Numeracy Test; WRAT: Wide Range Achievement Test; SNS: Subjective Numeracy Scale

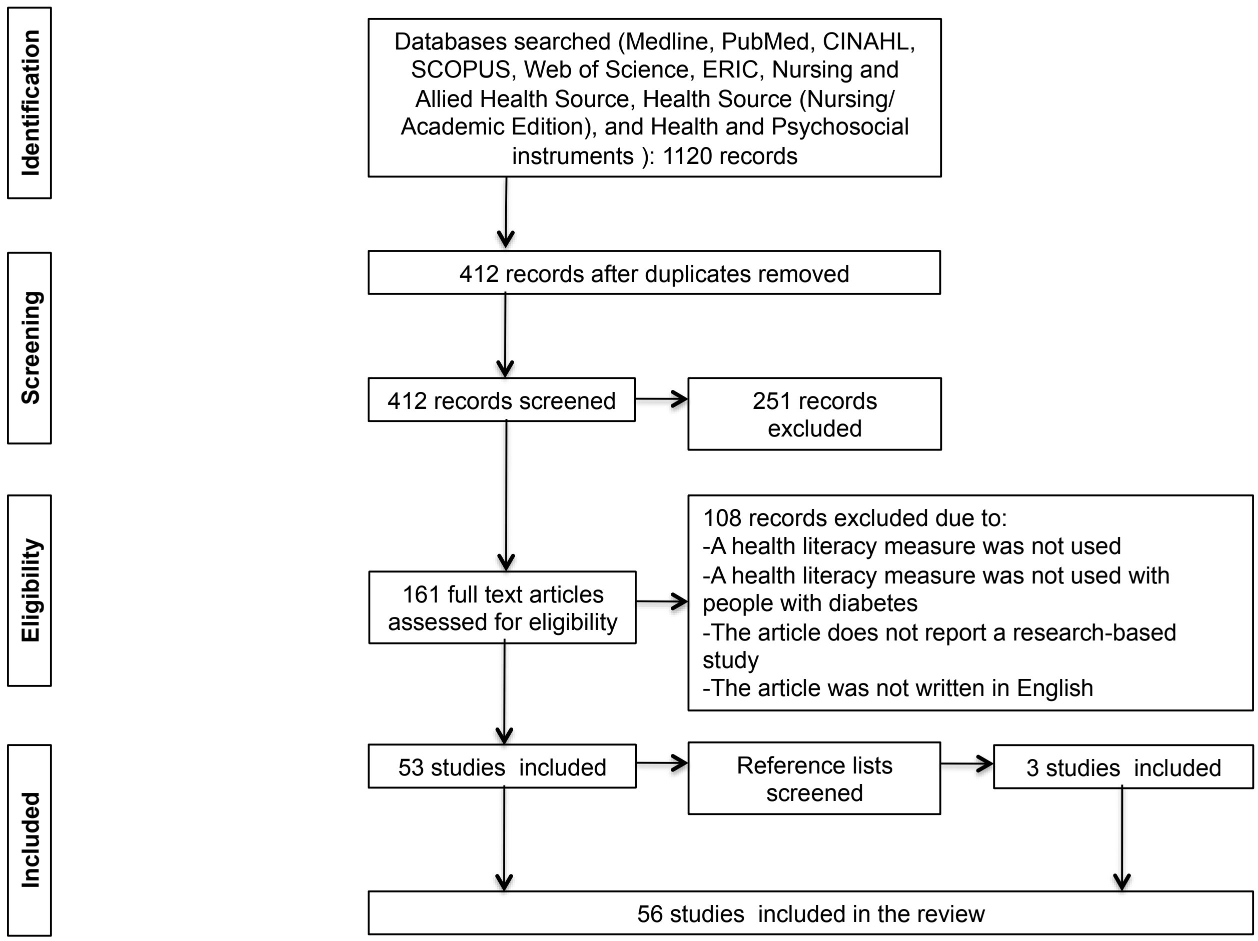
Table 3-2: Measurement Scope of the Identified Health Literacy and Numeracy Instruments[†]

Instrument	Diabetes Health Literacy Skill Set							Health literacy Domains			
	Reading & understanding health information	Applying health information	Understanding & applying health care providers' instructions	Completing medical & health care forms	Communicating with health care providers	Navigating the health care system	Making appropriate health-related decisions	Functional HL	Interactive HL	Critical HL	Numeracy
REALM	1	0	0	0	0	0	0	1	0	0	0
REALM-R	1	0	0	0	0	0	0	1	0	0	0
TOFHLA	2	2	0	0	1	0	0	2	0	0	2
s-TOFHLA	0	2	0	0	1	0	0	2	0	0	0
NVS	2	2	0	0	0	0	0	2	0	1	2
LAD	1	0	0	0	0	0	0	1	0	0	0
3-brief SQ	2	0	0	2	0	0	0	2	0	1	0
SILS	2	0	0	2	0	0	0	2	0	0	0
3-level HL Scale	2	2	2	2	2	2	2	2	2	2	0
DNT	2	0	0	0	0	0	0	2	0	0	2
WRAT	2	0	0	0	0	0	0	2	0	0	2
SNS	0	0	0	0	0	0	0	1	0	0	2

[†]Ratings: 0 = Not addressed, 1 = partially addressed, 2 = fully addressed

HL: Health Literacy; REALM: Rapid Estimate of Adult Literacy in Medicine; REALM-R: Rapid Estimate of Adult Literacy in Medicine-Revised; TOFHLA: Test of Functional Health Literacy in Adults; s-TOFHLA: Test of Functional Health Literacy in Adults-Short Form; 3-brief SQ; 3 brief screening questions; SILS: Single Item Literacy Screener; LAD: Literacy assessment in Diabetes; NVS: Newest Vital Sign; DNT: Diabetes Numeracy Test; WRAT: Wide Range Achievement Test; SNS: Subjective Numeracy Scale

Figure 3-1: Search Strategy and Retrieval Process



References

1. Berkman N, Terry D, and McCormack L, *Health literacy: What is it?* Journal of Health Communication, 2010. **15**: p. 9-19.

2. Nutbeam D, *The evolving concept of health literacy.* Social Science and Medicine 2008 **67**: p. 2072-2078.

3. White R, et al., *Addressing Health Literacy and Numeracy to Improve Diabetes Education and Care.* Diabetes Spectrum 2010. **23**: p. 238-243

4. Powell CK, Hill EG, and Clancy DE, *The relationship between health literacy and diabetes knowledge and readiness to take health actions.* The Diabetes Educator 2007. **33**: p. 144-151.

5. Sarkar U, et al., *Hypoglycemia is more common among type 2 diabetes patients with limited health literacy: The diabetes Study of Northern California (DISTANCE).* Journal of General Internal Medicine 2010. **25**: p. 962-967.

6. Schillinger D, et al., *Functional health literacy and the quality of physician-patient communication among diabetes patients.* patient Education and Counseling 2004. **52**: p. 315-323.

7. Schillinger D, et al., *Association of health literacy with diabetes outcomes.* JAMA, 2002. **288**: p. 475-482.

8. Baker DW, et al., *The health care experience of patients with low literacy.* Archives of Family Medicine, 1996. **5**: p. 329-334.

9. Baker D, et al., *Development of a brief test to measure functional health literacy.* Patient Education and Counseling 1999. **38**: p. 33-42.

10. Chew LD, Bradley KA, and Boyko EJ, *Brief questions to identify patients with inadequate health literacy.* Family Medicine 2004. **36**: p. 588–594.

11. Chew LD, et al., *Validation of screening questions for limited health literacy in a large VA outpatient population.* Journal of General Internal Medicine 2008. **23**: p. 561-566.

12. Davis TC, et al., *Rapid estimate of adult literacy in medicine: a shortened screening instrument.* Family Medicine, 1993. **25**: p. 391-395.

13. Parker R, et al., *The test of functional health literacy in adults (TOFHLA): a new instrument for measuring patients' literacy skills* Journal of General Internal Medicine 1995. **10**: p. 537-545.

14. Rawson KA, et al., *The METER: A Brief, Self-Administered Measure of Health Literacy* Journal of General Internal Medicine, 2009 **25**(1): p. 67-71.

15. McCormack L, et al., *Measuring health literacy: A pilot study of a New Skills-Based Instrument* Journal of Health Communication, 2010. **15**(1): p. 51-71.

16. Tsai TI, et al., *Methodology and validation of health literacy scale development in Taiwan* Journal of Health Communication 2011. **16**(1): p. 50-61.

17. Yamashita T and Kart C, *Is diabetes-specific health literacy associated with diabetes-related outcomes in older adults?* Journal of Diabetes 2011. **3**(2): p. 138–146.

18. Ishikawa H, Takeuchi T, and Yano E, *Measuring functional, communicative, and critical health literacy among diabetic patients.* Diabetes Care, 2008. **31**: p. 874-879.

19. Ishikawa H, et al., *Patient health literacy and patient-physician information exchange during a visit.* Family Practice 2009. **26**: p. 517-523.

20. Griffin JM, et al., *Variation in Estimates of Limited Health Literacy by Assessment Instruments and Non-Response Bias* Journal of General Internal Medicine, 2010. **25**(7): p. 675-681.

21. Baker D, *The meaning and measure of health literacy.* Journal of General Internal Medicine, 2006. **21**: p. 878-883.

22. Wolf MS, et al., *Patients' shame and attitudes toward discussing the results of literacy screening.* Journal of Health Communication, 2007. **12**: p. 721-732.

23. Kim S, et al., *Association of health literacy with self-management behavior in patients with diabetes.* Diabetes Care, 2004. **27**: p. 2980-2982.

24. Rothman R, et al., *Influence of patient literacy in the effectiveness of a primary care-based diabetes disease management program.* JAMA, 2004. **292**: p. 1711-1716.

25. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, A.M.A., *Health literacy: Report on the Council of Scientific Affairs.* JAMA, 1999. **282**: p. 526-527.

26. Aikens J and Piette J, *Diabetic patients' medication underuse, illness outcomes, and beliefs about antihyperglycemic and antihypertensive treatments.* Diabetes Care, 2009. **32**: p. 19-24.

27. Arthur SA, et al., *Health literacy and control in the medical encounter: A mixed-methods analysis.* J National Med Assoc, 2009. **101**: p. 677-683.

28. Bains SS and Egede LE, *Associations between health literacy, diabetes knowledge, self-care behaviors, and glycemic control in a low income population with type 2 diabetes.* Diabetes Technology and Therapeutics 2011. **13**: p. 335-341.

29. Carbone E, et al., *Testing the feasibility of an interactive learning styles measure for U.S. Latino adults with type 2 diabetes and low health literacy.* Quarterly Commun Health Educ., 2006. **25**: p. 315-335.

30. Castro C, et al., *Babel Babble: Physicians' use of unclarified medical jargon with patients.* American Journal of Health Behavior 2007. **31**(Suppl 1).

31. Cavanaugh K, et al., *Association of numeracy and diabetes control* Annals of Internal Medicine 2008. **148**: p. 737-746.

32. Cavanaugh K, et al., *Addressing literacy and numeracy to improve diabetes care: Two randomized controlled trials.* Diabetes Care, 2009. **32**: p. 2149-2155.

33. DeWalt DA, Boone RS, and Pignone MP, *Literacy and Its Relationship With Self-efficacy, Trust, and Participation in Medical Decision Making.* American Journal of Health Behavior, 2007 **31**(Suppl 1): p. S27-S35.

34. Endres L, et al., *Health literacy and pregnancy preparedness in pregestational diabetes* Diabetes Care 2004. **27**: p. 331-334.

35. Gazmararian JA, et al., *Health literacy and knowledge of chronic disease.* Patient Education and Counseling 2003. **51**: p. 267-275.

36. Gazmararian JA, et al., *Factors associated with medication refill adherence in cardiovascular-related diseases.* Journal of General Internal Medicine 2006. **21**: p. 1215-1221.

37. Gazmararian JA, Ziemer DC, and Barnes C, *Perception of barriers to self-care management among diabetic patients.* The Diabetes Educator 2009. **35**: p. 778-788.

38. Gerber B, et al., *Implementation of a low-literacy diabetes education computer multimedia application.* Diabetes Care, 2005. **28**: p. 1574-1580.

39. Gerber BS, et al., *The assessment of diabetes knowledge and self-efficacy in diverse population using Rasch Measurmeent* Journal of Applied Measurement 2006 **7**(1): p. 55-73.

40. Ishikawa H and Yano E., *The relationship of patient participation and diabetes outcomes for patients with high vs. low health literacy.* Patient Education and Counseling 2011. **84**(3): p. 393-7.

41. Jeppesen KM, Coyle JD, and Miser WF, *Screening questions to predict limited health literacy: A cross-sectional study of patients with diabetes mellitus* Annals of Family Medicine 2009. **7**: p. 24-31.

42. Kandula N, et al., *The relationship between health literacy and knowledge improvement after a multimedia type 2 diabetes education program.* Patient Education and Counseling, 2009. **75**: p. 321-327.

43. Laramee AS, Morris N, and Littenberg B, *Relationship of literacy and heart failure in adults with diabetes.* BMC Health Serv Res., 2007. **7**.

44. Mancuso JM, *Impact of health literacy and patient trust on glycemic control in an urban USA population.* Nursing and Health Sciences, 2010. **12**: p. 94-104.

45. Mayberry LS, et al., *Bridging the digital divide in diabetes: Family support and implications for health literacy* Diabetes Technology and Therapeutics, 2011; 13(10): 1005–1012.

46. Mbaezue N, et al., *The impact of health literacy on self-monitoring of blood glucose in patients with diabetes receiving care in an inner-city hospital.* J Natl Med Assoc., 2010. **102**: p. 5-9.

47. McCleary-Jones V, *Health literacy and its association with diabetes knowledge, self efficacy and diabetes self-management among African Americans with Diabetes Mellitus* The ABNF Journal 2011. **22**(2): p. 25-32.

48. Morris NS, et al., *The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability.* BMC Family Practice 2006. **7**.

49. Morris N, MacLean C, and Littenberg B, *Literacy and health outcomes: a cross-sectional study in 1002 adults with diabetes.* BMC Fam Prac., 2006. **7**.

50. Nath CR, et al., *Development and validation of a literacy assessment tool for persons with diabetes.* The Diabetes Educator, 2001. **27**: p. 857-864.

51. Ntiri DW and Stewart M, *Transformative learning intervention: Effect on functional health literacy and diabetes knowledge in older African Americans* Gerontol Geriatr Educ. , 2009. **30**: p. 100-103.

52. Nurss JR, et al., *Diabetes in urban African Americans: Functional health literacy of municipal hospital outpatients with diabetes.* THe Diabetes Educator 1997. **235**: p. 563-568.

53. Osborn CY, et al., *Diabetes Numeracy An overlooked factor in understanding racial disparities in glycemic control* Diabetes Care 2009. **32**(9).

54. Osborn C, et al., *Self-Efficacy Links Health Literacy and Numeracy to Glycemic Control.* Journal of Health Communication, 2010. **15**: p. 146-158.

55. Osborn CY, Bains SS, and Egede LE, *Health literacy, diabetes self-care, and glycemic control in adults with type 2 diabetes* Diabetes Technology and Therapeutics, 2010. **12**: p. 913-919.

56. Rees C, et al., *Correlates of patient-reported racial/ethnic health care discrimination in the diabetes study of northern California (DISTANCE)* Journal of Health Care for the Poor and Underserved, 2011. **22**: p. 211-225.

57. Rothman R, et al., *The relationship between literacy and glycemic control in a diabetes disease-management program.* The Diabetes Educator, 2004. **30**: p. 263-273.

58. Sarkar U, Fisher L, and Schillinger D, *Is self-efficacy associated with diabetes self-management across race/ethnicity and health literacy?* Diabetes Care, 2006. **29**: p. 832-839.

59. Sarkar U, et al., *Preferences for self-management support: Findings from a survey of diabetes patients in safety-net health systems.* Patient Education and Counseling 2008. **70**: p. 102-110.

60. Sarkar U, et al., *The literacy divide: Health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE).* Journal of Health Communication 2010. **15**(Suppl 2): p. 183-196.

61. Sarkar U, et al., *Validation of self-reported health literacy questions among diverse english and spanish-speaking populations.* Journal of general Internal Medicine 2011. **26**: p. 265-271.

62. Schillinger D, et al., *Closing the loop: Physician communication with diabetic patients who have low health literacy.* Archives of Internal Medicine 2003. **163**: p. 83-90.

63. Schillinger D, et al., *Does literacy mediate the relationship between education and health outcomes? A study of a low-income population with diabetes.* Public Health Rep, 2006. **121**: p. 245-254.

64. Schillinger D, et al., *Effects of self-management support on structure, process, and outcomes among vulnerable patients with diabetes.* Diabetes Care, 2009. **32**: p. 559-566.

65. Seligman HK, et al., *Physician notification of their diabetes patients' limited health literacy. A randomized, controlled trial.* Journal of General Internal Medicine, 2005. **20**: p. 1001-1007.

66. Shigaki C, et al., *Motivation and diabetes self-management.* Chronic Illness, 2010. **6**: p. 202–214.

67. Tang YH, et al., *Health literacy, complication awareness, and diabetic control in patients with type 2 diabetes mellitus* AJAN, 2008. **62**: p. 74-83.

68. Thabit H, et al., *Globalization, immigration and diabetes self-management: an empirical study amongst immigrants with type 2 diabetes mellitus in Ireland* QJM: An International Journal of Medicine, 2009 **102**(10): p. 713-720

69. Wallace A, et al., *Literacy-appropriate educational materials and brief counseling improve diabetes self-management* Patient Education and Counseling, 2009. **75**: p. 328-333.

70. Wallace AS, et al., *The influence of literacy on patient-reported experiences of diabetes self-management support.* Nursing Research, 2010. **59**: p. 356-363.

71. White RO, et al., *Development and validation of a Spanish diabetes-specific numeracy measure: DNT-15 Latino* Diabetes Technology and Therapeutics, 2011. **13**(9).

72. Whitten P, et al., *Health Education Online for Individuals with Low Health Literacy; Evaluation of the Diabetes and You Website.* Journal of Technology and Human Services, 2008. **26**: p. 77- 88.

73. Williams M, et al., *Relationship of functional health literacy to patients' knowledge of chronic disease: A study of patients with hypertension and diabetes* Archives of Internal Medicine 1998. **158**: p. 166-172.

74. Davis TC, et al., *Rapid assessment of literacy levels of adult primary care patients* Family Medicine 1991. **23**: p. 433-435.

75. Bass PF, Wilson J, and Griffith CH, *A shortened instrument for literacy screening.* Journal of General Internal Medicine 2003. **18**: p. 1036-1038.

76. Weiss BD, et al., *Quick assessment of literacy in primary care: the newest vital sign.* Annals of Family Medicine 2005. **3**: p. 514–522.

77. Huizinga MM, et al., *Development and validation of the Diabetes Numeracy Test (DNT).* BMC Health Services Research, 2008 **8**(96).

78. Stone MH, Jastak S, and Wilkinson G, *Wide Range Achievement Test - 3.* Rasch Measurement Transactions, 1995. **8**(4).

79. Fagerlin A, et al., *Measuring Numeracy without a Math Test: Development of the Subjective Numeracy Scale.* Medical Decision Making, 2007. **27**.

Chapter 3 Supplementary Tables

Table S3-1: Overview of the Included Studies

Author(s), year, Country	Study design	Sample Characteristics	HL/Numeracy measure(s) used
Aikens J & Piette J, 2009 [26] USA	Cross-sectional	N = 1376; two groups: group 1 mean age (SD): 55.3 (11.8), group 2 mean age (SD): 57.2 (10.7) T1DM and T2DM	3-brief SQ
Arthur SA, Geiser HR, Arriola KR, & Kripalani S, 2009 [27] USA	Cross-sectional	N = 31, Mean age (SD): NR T2DM	REALM
Bains SS & Egede LE, 2011 [28] USA	Cross-sectional	N = 125, Age: 50.7% < 65 years & 49.3% ≥ 65 years T2DM	REALM-R
Carbone E, Lennon K, Torres M, & Rosal M, 2006 [29] USA	Cross-sectional	N = 10, Mean age (range): 65 (55-82) T2DM	s-TOFHLA
Castro C, Wilson C, Wang F, & Schillinger D, 2007 [30] USA	Cross-sectional	N = 74, Two groups: audiotape group mean age (SD): 64 (10) & telephone follow-up group mean age (SD): 63 (9) T2DM	s-TOFHLA
Cavanaugh K et al., 2008 [31] USA	Cross-sectional	N = 398, Median age (IQR): 55 (46-64) T1DM & T2DM	REALM, DNT, WRAT-3R
Cavanaugh K et al., 2009 [32] USA	Randomized controlled trial	N = 198, Intervention group Median age (IQR): 53 (40-59.5), Control group Median age (IQR): 52 (45-59) T1DM & T2DM	REALM
DeWalt DA, Boone RS, & Pignone MP, 2007 [33] USA	Cross-sectional	N = 268, Two groups: High HL group mean age (SD): 58 (11) & low HL group mean age (SD): 62 (10) T2DM	REALM
Endres L, Sharp L, Haney E, & Dooley S, 2004 [34] USA	Cross-sectional	N = 74, Mean age (SD): 31 (6) Gestational diabetes	TOFHLA
Gazmararian JA, Williams MV, Peel J, & Baker DW, 2003 [35] USA	Cross-sectional	N = 653 (266 have diabetes); Age: NR T1DM & T2DM	s-TOFHLA
Gazmararian JA et al., 2006 [36] USA	Longitudinal	N = 1549, (Individuals who have diabetes and/or coronary heart disease and/or hypertension	s-TOFHLA

		and/or hyperlipidemia), sub-group analysis: NR	
Gazmararian JA, Ziemer DC, & Barnes C, 2009 [37] USA	Cross-sectional	N = 35, Group 1 mean age: 48, Group 2 mean age: 58, Group 3 mean age: 54 Diabetes type: NR	REALM
Gerber B et al., 2005 [38] USA	Randomized controlled trial	N = 244, Intervention-Low HL group mean age (SD): 57.7 (11.7), Intervention-High HL group mean age (SD): 49.4 (12), Control-Low HL group mean age (SD): 60.4 (10.8), Control-Low HL group mean age (SD): 51.8(11.3) T1DM & T2DM	s-TOFHLA
Gerber BS et al., 2006 [39] USA	Cross-sectional	N = 255, Mean age (SD): 55.2 (12.3) T1DM & T2DM	s-TOFHLA
Ishikawa H, et al., 2008 [18] Japan	Cross-sectional	N = 138, Mean age (SD): 65 (9.9) T2DM	3-level HL scale
Ishikawa H, et al., 2009 [19] Japan	Cross-sectional	N = 134, Mean age (SD): 65 (9) T2DM	3-level HL scale
Ishikawa H & Yano E., 2011 [40] Japan	Cross-sectional	N = 143, Mean age (SD): 65 (9.1) T2DM	3-level HL scale
Jeppesen KM, Coyle JD, & Miser WF, 2009 [41] USA	Cross-sectional	N = 225, Mean age (SD): 53.76 (12.8) Diabetes type not reported	s-TOFHLA SILS
Kandula N et al., 2009 [42] USA	Interventional: two-group pre-post	N = 190, Mean age (SD): 55.9 (9.3) Diabetes type: NR	s-TOFHLA
Kim S, et al., 2004 [23] USA	Longitudinal	N = 92, Mean age: 62 T1DM & T2DM	s-TOFHLA
Laramee AS, Morris N, & Littenberg B, 2007 [43] USA	Cross-sectional	N = 998, Mean age (range): 65 (22-93) T1DM & T2DM	s-TOFHLA
Mancuso JM, 2010 [44] USA	Cross-sectional	N = 102, Mean age (SD): 52 (9.1) T1DM & T2DM	TOFHLA
Mayberry LS, Kripalani S, Rothman RL, & Osborn CY, 2011 [45] USA	Cross-sectional	N = 61, Mean age (SD): 56.9 (8.8) T2DM	3-brief SQ, SNS
Mbaezue N et al., 2010 [46] USA	Cross-sectional	N = 189, Mean age (SD): 51.2 (10) T1DM & T2DM	s-TOFHLA

McCleary-Jones V, 2011 [47] USA	Cross-sectional	N = 50, Mean age (SD): 58.6 (11.5) T1DM & T2DM	REALM
Morris NS, et al., 2006 [48] USA	Cross-sectional	N = 999, Age: 22-93 T1DM& T2DM	SILS s-TOFHLA
Morris N, MacLean C, & Littenberg B, 2006 [49] USA	Cross-sectional	N = 1002, Median age (IQR): 66 (57-74) T1DM & T2DM	s-TOFHLA
Nath CR, et al., 2001 [50] USA	Longitudinal	N = 203, Mean age (SD): 43.6 (15.04) Diabetes type: NR	LAD, REALM, WRAT-3
Ntiri DW & Stewart M, 2009 [51] USA	Interventional: two-group pre-post	N = 20, Mean age (range): 68.1 (57-84) Diabetes type not reported	s-TOFHLA LAD
Nurss JR et al., 1997 [52] USA	Cross-sectional	N = 131, Mean age (SD): 54.7 (14) Diabetes type not reported	TOFHLA
Osborn CY, Cavanaugh K, Wallston KA, White RO, & Rothman RL, 2009 [53] USA	Cross-sectional	N = 383, Median age (IQR): 56 (47-64) T2DM	REALM, DNT, WRAT-3
Osborn C, Cavanaugh K, Wallston K, & Rothman R, 2010 [54] USA	Cross-sectional	N = 383, Mean age (SD): 54.4 (13) T1DM & T2DM	REALM, WRAT-3
Osborn CY, Bains SS, & Egede LE, 2010 [55] USA	Cross-sectional	N = 130, Mean age (SD): 60.7 (11.8) T2DM	REALM-R
Powell CK, et al., 2007 [4] USA	Cross-sectional	N = 68, Median age (IQR): 55 (51-60) T2DM	REALM
Rees C et al., 2011 [56] USA	Cross-sectional	N = 17,795, Age: NR T1DM & T2DM	3-brief SQ
Rothman R, DeWalt D, et al., 2004 [24] USA	Randomized controlled trial	N = 217, Control group Low HL mean age (SD) = 59(10.4), Control group High HL mean age (SD) = 56(10.9), Intervention group Low HL mean age (SD) = 57(10.5), Intervention group High HL mean age (SD)= 51(13.1) T2DM	REALM
Rothman R, Malone, et al., 2004 [57] USA	Randomized controlled trial	N = 159, Control low HL group mean age (SD): 59 (10.4), Control high HL group mean	REALM

		age (SD): 56 (10.9), Intervention low HL group mean age (SD): 57 (10.5), Intervention low HL group mean age (SD): 51 (13.1) T2DM	
Sarkar U, Fisher L, & Schillinger D, 2006 [58] USA	Cross-sectional	N = 408, Mean age (SD): 58.1 (11.4) T2DM	s-TOFHLA
Sarkar U et al., 2008 [59] USA	Cross-sectional	N = 796, Mean age (SD): 58 (12) T1DM & T2DM	3-brief SQ
Sarkar U, Karter AJ, et al., 2010 [60] USA	Cross-sectional	N = 14,102, Mean age: 59 T1DM & T2DM	3-brief SQ
Sarkar U, Karter A, et al., 2010 [5] USA	Cross-sectional	N = 14,357, Mean age (SD): 58 (10) T2DM	3-brief SQ
Sarkar U, Schillinger D, López A, & Sudore R, 2011 [61] USA	Cross-sectional	N = 296, Mean age (SD): 54.9 (12.1) T2DM	3-brief SQ s-TOFHLA
Schillinger D, et al., 2002 [7] USA	Cross-sectional	N = 408, Mean age (SD): 58.1 (11.4) T2DM	s-TOFHLA
Schillinger D et al., 2003 [62] USA	Cross-sectional	N = 408, Age: 55.7% ≥ 65 & 44.3% < 65 T2DM	s-TOFHLA
Schillinger D, et al., 2004 [6] USA	Cross-sectional	N = 408, Mean age: 58.1 T2DM	s-TOFHLA
Schillinger D, Barton LR, Karter AJ, Wang F, & Adler N, 2006 [63] USA	Cross-sectional	N = 395, Mean age (SD): 57.9 (11.4) T2DM	s-TOFHLA
Schillinger D, Handley M, Wang F, & Hammer H, 2009 [64] USA	Randomized controlled trial	N = 339, Mean age (SD): 56.1 (12) T2DM	s-TOFHLA
Seligman HK et al., 2005 [65] USA	Cross-sectional	N = 182, Intervention group mean age (SD): 62.3 (11.3), Control group mean age (SD): 63.4 (9.5) T2DM	s-TOFHLA
Shigaki C et al., 2010 [66] USA	Cross-sectional	N = 77, Mean age (SD): 63 (13) T2DM	REALM, NVS
Tang YH, Pang SM, Chan MF, Yeung GS, & Yeung VT, 2008 [67]	Cross-sectional	N = 149, Mean age (range): 59.8 (27-90) T2DM	s-TOFHLA

China			
Thabit H et al., 2009 [68] Ireland	Cross-sectional	N = 100, Mean age (SD): 45.8 (11.8) T2DM	REALM
Wallace A, Seligman H, Davis T, & et al, 2009 [69] USA	Cross-sectional	N = 208, Mean age (range): 56 (29-93) T2DM	s-TOFHLA
Wallace AS, Carlson JR, Malone RM, Joyner J, & DeWalt DA, 2010 [70] USA	Interventional: one-group pre-post	N = 250, Mean age (range): 58 (23-85) T2DM	s-TOFHLA
White RO, Osborn CY, Gebretsadik T, Kripalani S, & Rothman RL, 2011 [71] USA	Cross-sectional	N = 144, Mean age (SD): 47.8 (12.1) T1DM & T2DM	s-TOFHLA, DNT-15, WRAT-4
Whitten P, Buis L, Love B, & Mackert M, 2008 [72] USA	Interventional: one-group pre-post	N = 50, Mean age (SD): 40.3 (13.1) Diabetes type not reported	s-TOFHLA REALM
Williams M, Baker D, Parkes R, & Nurss R, 1998 [73] USA	Cross-sectional	N = 114, Low HL group mean age (SD): 57.4 (10.2), marginal HL group mean age (SD): 53.2 (8.8), high HL group mean age (SD): 49.2 (10.3) T1DM & T2DM	TOFHLA

HL: Health Literacy; NR: Not Reported; IQR: Interquartile Range; T1DM: Type 1 Diabetes Mellitus; T2DM: Type 2 Diabetes Mellitus; REALM: Rapid Estimate of Adult Literacy in Medicine; REALM-R: Rapid Estimate of Adult Literacy in Medicine-Revised; TOFHLA: Test of Functional Health Literacy in Adults; s-TOFHLA: Test of Functional Health Literacy in Adults-Short Form; 3-brief SQ; 3 brief screening questions; SILS: Single Item Literacy Screener; LAD: Literacy assessment in Diabetes; NVS: Newest Vital Sign; DNT: Diabetes Numeracy Test; WRAT: Wide Range Achievement Test; SNS: Subjective Numeracy Scale

Chapter 4

Measurement Properties and Comparative Performance of Health Literacy Screening Questions in a Predominantly Low Income African American Population With Diabetes †

Abstract

Background: The evidence on the utility of the 16 screening questions (16-SQ) of inadequate health literacy (HL) and their briefer version (3-SQ) is inconsistent and limited to studies that validated these questions among predominantly white, English speaking populations drawn from academic practices. Additionally, no investigation of measurement model of these questions and their scoring has been undertaken. The objective of this study was to examine the measurement properties of the16-SQ of inadequate HL and their briefer version (3-SQ) in greater detail and identify the best screen for inadequate HL in lower income non-white populations.

Methods: We used cross-sectional data from a study of 378 predominantly African-American individuals with type-2 diabetes. We computed sensitivity, specificity, positive and negative likelihood ratios, and C-indices. We also conducted factor analysis to examine the measurement model of these questions, and used structural equation modeling (SEM) for confirmatory purposes. The s-TOFHLA was used as a reference measure.

---

† A version of this chapter has been submitted for publication. Co-authors: Sumit R Majumdar[1], MD MPH; Leonard E Egede[2], MD MS; Jeffrey A Johnson[3], PhD ([1] Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada; [2] Department of Medicine, Medical University of South Carolina, Charleston, South Carolina, US; [3] School of Public Health, University of Alberta, Edmonton, Alberta, Canada).

Results: Mean age was 56.1 years (SD=12.4), 69% were females, and 83% were African-American. The prevalence of limited HL (s-TOHFLA scores <23) was 10%. Six questions (6-SQ) were included in the final item-reduced factor analysis and produced two factors explaining 58% of the variance. In confirmatory SEM, this 2-factor model had a good fit (chi-square = 9.5; P = 0.305; RMSEA = 0.023). Weighted summative score of the 6-SQ and the item "problems learning" performed better than the 3-SQ in identifying patients with inadequate HL with C-indices of 0.67 and 0.75 respectively.

Conclusion: The weighted summative score of the 6-SQ and the item "problems learning" performed better than the other items or combinations of these items in identifying individuals with inadequate HL in this sample of predominantly low income African-Americans. Further exploration of the measurement properties of these questions needs to be undertaken before widely recommending brief HL screening instruments.

4.1. Introduction

Health literacy (HL) is “the ability to obtain, process, or understand basic health information needed to make appropriate health care decisions” [1]. Inadequate HL has been found to adversely influence health outcomes, especially in low-income patients with chronic diseases [2-5]. Although routine screening for inadequate HL in clinical settings is still controversial, its high (46%) prevalence in the US population [6], and its association with poor health outcomes [7, 8], has led to an increased interest in HL assessments. However, common measures of HL, such as the Test of Functional Health Literacy in Adults – Short form (s-TOFHLA) and the Rapid Estimate of Adult Literacy in Medicine (REALM), are time consuming, require face-to face interviews, might introduce discomfort and embarrassment, especially for those with inadequate HL, cannot be administered by telephone, and they are not feasible in large surveys [9-12].

Chew and colleagues developed 16 self-reported HL screening questions (16-SQ) [13], then identified a briefer version of three questions (3-SQ) [14]. They subsequently reported that out of the 3-SQ, a single item about “confidence with completing forms” with a response cut-point of “somewhat,” may be sufficient to detect patients with inadequate HL; this item did not, however, perform as well in identifying patients with limited (i.e. inadequate plus marginal) HL [13, 14]. This single item was also reported by others to perform best in identifying patients with inadequate HL at a university-based primary clinic [15]. Chew and colleagues also found that a scale combining the three questions offered no additional benefit to the one question about confidence with forms [14]. The question about needing help to read hospital materials was predictive of inadequate HL in sample of patients at a university based vascular surgery clinic [16]. This same question, which is known as the Single Item Literacy Screener (SILS), performed reasonably well in ruling out inadequate HL in adults [17]. Sarkar & colleagues later evaluated the performance of the 3-SQ among Spanish and English-speaking individuals with type 2 diabetes, and found that one of the 3 items “confidence with forms” or a summative score of the three items were both useful in identifying inadequate HL in this population [18].

Overall, the evidence on the utility of these screening questions in identifying inadequate HL, whether as single items or a combination of these items, is inconsistent. This evidence is based on studies that validated these questions among predominantly white, English speaking populations drawn from academic practices. Additionally, the original identification of the 3-SQ from the 16-SQ was based only on Receiver Operating Characteristic (ROC) analysis in one patient population [13], while no investigation of the factor structure of the 16-SQ has been undertaken. Furthermore, in all of the studies that used a summative score of these items [13-15, 18, 19], it was done by simple summation of the item scores assuming that all these items equally contribute to the total score. Therefore, we sought to examine the factor structure and the measurement properties of the 16-SQ and the 3-SQ in greater detail, and identify the best set of items to screen for inadequate HL in a predominantly lower income non-white population.

4.2. Methods

4.2.1. Design and Data Source

This validation study used cross-sectional data from a study conducted in

South Carolina, USA that has been previously described in detail [20]. Briefly, patients were recruited at two adult primary care clinics, and were included if they were 18 years or older with a diagnosis of type 2 diabetes in their medical record and a clinic appointment between June and August 2010. Patients were ineligible if they did not speak English or if the research assistants determined that they were too ill or cognitively impaired to participate. Ethics approval of this study was obtained from the Institutional Review Board (IRB) at the University of South Carolina.

4.2.2. Self-reported Health Literacy Measure

Participants completed the 16-SQ assessing HL, including the 3-SQ on "problems in learning" (HL12), "confidence with forms" (HL14), and "needing help in reading hospital materials" (HL16). All questions were scored on a five-point Likert scale (always=1, often=2, sometimes=3, occasionally=4, never=5) with higher scores indicating lower HL. The scores of items HL5 – HL13 and item HL16 were reversed so that higher scores indicate lower self-reported HL.

4.2.3. Reference Health Literacy Measure

The s-TOFHLA was administered to all subjects; we considered it our reference measure as it is the most frequently used HL measure in the literature [9]. The s-TOFHLA is the short form of the TOFHLA, which was developed in the United States to measure "functional health literacy", defined as assessing reading, writing, and numeracy skills in relation to health [21]. The s-TOFHLA includes 36 reading comprehension and four numeracy items, and uses the modified Cloze procedure, where every fifth to seventh word in passage is omitted, and the respondent selects a response from four options [10]. The s-TOFHLA scores range from 0 – 36, with higher scores indicating better reading comprehension, and thus higher functional HL. We used standard cut-offs where scores from 0-16 represent inadequate HL, 17-22 marginal HL, and 23-36 adequate HL [10]. S-TOFHLA scores of 0-22 are collectively referred to as limited HL. We assessed the performance of the self-reported questions and their summative score compared to the s-TOFHLA categories of inadequate (scores 0-16) and limited HL (scores 0-22).

4.2.4. Other Measures

Data on self-reported age, sex, race/ethnicity (black; white), years of formal education, employment status (employed; unemployed), health insurance (insured; uninsured), and annual household income (<$10,000; < $25,000; > $25,000) were collected. Diabetes knowledge was assessed with the Diabetes Knowledge Questionnaire (DKQ) [22], and diabetes-specific self-efficacy with the Perceived Diabetes Self-Management Scale (PDSMS) [23]. Most recent A1c results were obtained from patients' medical records.

4.2.5. Statistical Analysis

Reliability and Validity: Descriptive statistics and estimates of reliability and validity were computed. Internal consistency reliability was assessed using Cronbach's alpha coefficient. Criterion validity was assessed by examining correlations of each of the 16-SQ with the s-TOFHLA. Construct validity was assessed using the hypothesis-testing approach. Specifically, we established *a priori* hypotheses about the direction and magnitude of correlations of the 16-SQ with related traits and constructs based on evidence on the relationships between HL and these constructs [24, 25]. We hypothesized that the 16-SQ would be positive and moderately correlated with education, positive and strongly correlated with diabetes knowledge, and positive and moderately correlated with self-efficacy. We used the following criteria for the strength of correlation: < 0.3 "weak"; 0.3 – 0.5 "moderate"; and > 0.5 "strong" [26].

ROC Analysis: We calculated C-Indices (the area under the ROC curve) for each question for the HL categories of inadequate (comparing s-TOFHLA scores of 0-16 versus 17-36) and limited (comparing s-TOFHLA scores of 0-22 versus 23-36). We considered a C-index greater than 0.6 to be useful; this is higher than the 0.5 cut-off that reflects discrimination no better than chance. We also calculated sensitivity, specificity, and positive (LR+) and negative (LR-) likelihood ratios for each question.

Factor Analysis: We conducted exploratory factor analysis of the 16-SQ using the principal factor analysis method and oblique Promax rotation. The eigenvalues from the factor analysis were used to determine the number of factors in the optimum solution. The 16 questions were assigned to the factor on which they loaded most heavily in the rotated solution. Next, to minimize redundancy and simplify the 16-item model, we eliminated redundant items. To do so, we grouped the items based on their content and measurement scope, and selected the best item within each set based on factor loadings and C-indices. Among the 16 questions, four items assess "ease of reading" (set 1: HL1 – HL4), four assess "difficulty in understanding" (set 2: HL5 – HL8), five assess "problems due to difficulty in understanding" (set 3: HL9 – HL13), and the remaining items (HL14-HL16) were miscellaneous, and were all included in our final selection. The items within each set were moderately to strongly correlated with each other (Pearson r: set 1: 0.55 – 0.66; set 2: 0.43 – 0.73; set 3: 0.32 – 0.68). The item with the largest factor loading and C-index within each set was selected. The 3 selected items along with the 3 remaining miscellaneous items formed the simplified model with 6 items (6-SQs) and was tested in a second factor model.

Based on the factor models structures, weighted summative scores of the items within a fit factor model were computed. We used the factor models to predict factor-scoring coefficients, which were used to generate weights for each question. Then a weighted score for each set was generated using a linear combination of these items, each with its corresponding weight. Additionally, we used the net reclassification index (NRI) [27] to compare these weighted scores.

Structural Equation Modeling (SEM): To further validate these factor models, we tested them as measurement models using SEM. The two models were estimated with a covariance matrix generated by the 378 cases, and model specification was based on the factor structures obtained in exploratory factor analysis. Subsequently, we examined the validity of the final factor model by testing it in a structural model, which also included age, sex, educational level and income as exogenous variables, and diabetes knowledge and self-efficacy as endogenous variables. Paths from HL variables to diabetes knowledge and self-efficacy, and from diabetes knowledge to self-efficacy were specified. Additionally, model specification included correlations among exogenous variables. Next, A1c was added to the model as an endogenous variable, with paths from HL and the other endogenous variables. For all SEM models, we considered adequate model fit statistics to be represented by an insignificant Chi-square goodness-of-fit test result at 0.05 threshold, and a root mean square error of approximation (RMSEA) less than 0.05 [28]. All analyses were performed using STATA 11.1 [29] and LISREL 8.0 [30].

4.3. Results

4.3.1. Participants' Characteristics

Characteristics of this sample of 378 adults with type 2 diabetes are shown in Table 4-1. Mean (SD) age was 56.1 (12.4) years, 69% were females, 83% were African-American, 39% unemployed, 39% uninsured, and 80% had annual income <$25,000. Diabetes knowledge (mean=15.8; SD=3.9) and self-efficacy (mean=28.8; SD=5.5) were generally low in this population. Based on the s-TOFHLA, the prevalence of limited HL was 10%.

4.3.2. Reliability and Validity

The Cronbach's alpha coefficient was 0.87 for the 16-SQ and 0.53 for the 3-SQ, which indicates good internal consistency of the former and poor of the latter. There was a fair to good evidence of criterion and construct validity of the 16-SQ. The associations between all of the 16 items and the s-TOFHLA were all statistically significant, however, with weak-moderate Pearson correlations ranging from 0.12 to 0.34 (Table 4-2). Out of the 16-SQ, items HL12, HL14, HL16 (i.e. the 3-SQ) were, in general, weakly correlated with the s-TOFHLA. As for the relationship between the 16-SQ and education, fourteen items had a significant but small correlation with years of education, and two items (HL3 and HL13) on medication use were not correlated with education. The 16 items were all weakly correlated with diabetes knowledge. As for the relationship between the 16-SQ and self-efficacy, twelve items had significant, but weak, correlations with self-efficacy (0.11 – 0.29); only one item of the 3-SQs (HL12) was significantly correlated with self-efficacy (r = 0.24). The s-TOFHLA was weakly correlated with years of education (r = 0.27) and self-efficacy (r = 0.11) but strongly correlated with diabetes knowledge (r = 0.51).

4.3.3. ROC Analysis

Overall, participants who reported worse HL based on the 16 items were consistently more likely to have inadequate or limited HL on the s-TOFHLA (Table 4-3). Six of 16-SQ each had a C-index > 0.6 for inadequate HL and 10 questions did so for limited HL compared to the s-TOFHLA (Table 4-3). Two items of the 3-SQ (HL12 “problems learning” and HL16 “needing help in reading”) successfully differentiated those with inadequate HL (s-TOFHLA 0 – 17) compared to those with marginal plus adequate HL (s-TOFHLA 18 – 36) with C-indices of 0.71 (95% CI: 0.58 – 0.85) and 0.66 (95%CI: 0.51 – 0.81) respectively. The same items were also successful in differentiating between those with limited HL (s-TOFHLA 0 – 22) compared to those with adequate HL (s-TOFHLA 23 – 36) with a C-index of 0.71 (95% CI: 0.61 – 0.81) for HL12 and 0.71 (95%CI: 0.61 – 0.80) for HL16. The third item of the 3-SQ “HL14” (confidence with forms) failed to differentiate those with inadequate HL (s-TOFHLA 0 – 17) compared to those with marginal plus adequate HL (s-TOFHLA 18 – 36) and those with limited HL (s-TOFHLA 0 – 22) compared to those with adequate HL (s-TOFHLA 23 – 36) with C-indices of 0.55 (95% CI: 0.40 – 0.70) and 0.60 (95%CI: 0.49 – 0.70) respectively.

Sensitivities, specificities, LR+ and LR- for the 16-SQs for detecting inadequate HL and limited HL based on the S-TOFHLA at each threshold are shown in Tables S4-1 and S4-2. The screening threshold that optimized both sensitivity and specificity for the majority of the questions was at the response of **"**Sometimes**"** or more for detecting both inadequate HL and limited HL. The performance of the 3-SQ in detecting inadequate HL varied, with "confidence with forms" (HL14) having the lowest sensitivity (53.3%) and specificity (60.6%) compared to "problems learning" (HL12) (sensitivity= 66.7%; specificity= 80.9%), and "help read" (HL16) (sensitivity= 66.7%; specificity= 69.2%). As for limited HL, "confidence with forms" (HL14) also had the lowest sensitivity (58.1%) and specificity (61.9%) compared to "problems learning" (HL12) (sensitivity= 58.1%; specificity= 82.6%; Diagnostic Odds Ratio (DOR)= 6.53), and "help read" (HL16) (sensitivity 71%; specificity=71.7%; DOR=6.25). Overall, "problems learning" had the highest sensitivity and specificity for detecting both inadequate HL and limited HL compared to the s-TOFHLA.

4.3.4. Factor Analysis and Structural Equation Modeling

Factor analysis of the 16-SQ produced three factors explaining 60% of the variance, all with eigenvalues greater than 1.0, and thus were all retained in the factor model (Table 4-4). This 3-factor model was tested in SEM with each of the factors as a latent variable with its corresponding indicators as identified in factor analysis, and had poor fit indices (Chi-square (d.f. 101) = 423.02; $P < 0.0001$, and RMSEA = 0.098). The second simplified factor model included 6 items (6-SQ): HL1, HL5 and HL12 were selected from three sets based on their factor loadings and C-indices which were the highest for both HL thresholds, along with the three miscellaneous items (HL14, HL15, HL16). Factor analysis of this simplified model produced two factors explaining 58% of the variance, both with eigenvalues grater than 1.0, and thus were both retained in the factor model (Table 4-4). In this model, items HL12, HL14, HL16 (i.e. the 3-SQ) did not all load under one factor. Cronbach's alpha was 0.64 for factor 1, 0.58 for factor 2, and 0.65 for the overall scale. In confirmatory SEM, this 2-factor model (6-SQ), which was specified based on its factor structure, had a good fit (Chi-square (d.f. 8) = 9.46; P = 0.305; RMSEA = 0.023). The previously validated 3-SQ had the largest path coefficients; 0.72 for "confidence with forms" (HL14), 0.61 for "problems learning" (HL12), and 0.62 for "help read" (HL16) in this measurement model.

In further validation of the 6-SQ model, the structural model which included diabetes knowledge and self-efficacy as endogenous variables well fit the data with a Chi-square (d.f. 43) = 55.89; P = 0.09, and RMSEA = 0.032. Adding A1C as an endogenous variable to this model, which is an extension to the structural model validation, also fit the data well with a Chi-square (d.f. 52) = 63.45; P = 0.133, and RMSEA = 0.028.

4.3.5. *Summative Scores*

Based on the weighted summative scores, a cut-off point of 17 for the 6-SQ maximized sensitivity and specificity and had the highest C-indices for both HL thresholds; 0.67 (95%CI: 0.53 - 0.82) for inadequate HL, and 0.69 (95%CI: 0.60 - 0.79) for limited HL (Table 4-5). We also computed a weighted summative score for the 3-SQ to compare it with that of the 6-SQ. A cut-off point of 9 for the

3-SQ maximized sensitivity and specificity and had the highest C-indices for both inadequate (0.65; 95%CI: 0.52 - 0.79) and limited HL (0.69; 95%CI: 0.60 - 0.78).

The 6-SQ and 3-SQ weighted summative scores had similar performance in identifying patients with limited HL with C-indices of 0.69 for the former and 0.67 for the latter; however, the 6-SQ was better than the 3-SQ in identifying patients with inadequate HL with C-indices of 0.67 and 0.62 respectively. The summative scores were also compared to each item within the 3-SQ. "Problems learning" (HL12) was superior to the 6-SQ and 3-SQ summative scores and to the other items in identifying patients with inadequate HL (C-index = 0.75), followed by the summative score of the 6-SQ (C-index = 0.67) (Figure 4-1). "Problems learning" (HL12), "help read" (HL16) and the summative score of the 6-SQ were the best in identifying patients with limited HL, and had similar performance with C-indices of 0.70, 0.70 and 0.69 respectively (Figure 4-2). Net proportion of patients reclassified correctly by the 6-SQ summative score compared to the 3-SQ summative score was 4% for inadequate HL; however, it was only 0.2% for limited HL.

4.4. Discussion

To our knowledge, this is the first study to explore the measurement properties and performance of the previously reported 16-SQ and their brief version, the 3-SQ, in a predominantly African American population, and the first to explore the factor structure of these questions. We found that the performance of these questions in identifying patients with inadequate HL in this population was not as good as their previously reported performance in other populations [13, 14, 18]. We also found that the test characteristics (sensitivity, specificity, C-indices) of the 16-SQ were poorer in this patient population than in other populations [13, 18]. Overall, the 16-SQ were better in discriminating between patients with inadequate HL vs. marginal plus adequate HL than those with limited HL versus adequate HL, which is consistent with the literature [13, 14, 18]. In contrast to the literature [13-15], the item “confidence with forms” had the poorest performance in identifying individuals with inadequate or limited HL compared to the other 2 items of the 3-SQs and the summative scores. In this sample, “problems learning” had the highest correlations with relevant constructs, largest C-indices, and performed better than both “confidence with forms” and “help read” (SILS) in discriminating across different levels of HL.

Based on the factor analysis, we identified a set of 6 items that had a valid factor structure, and performed better than the 3-SQ in discriminating between patients with inadequate HL and those with marginal plus adequate HL. Additionally, the factor structures of the 3-SQ and 6-SQ indicated that the items within each set do not have equal contributions to the total score, and thus should be weighted before they are summed. The lack of weighting of these items in prior studies could be a reason for unsuccessful attempts to improve the performance of these screening questions by summing a score from a combination of items [14, 15]. Although none of the brief items or combinations was very good compared with the reference standard in this population, the item “problems in learning” and the weighted summative score of the 6-SQ were the best screeners of inadequate HL in this population. Further, given the inconsistency of evidence of the utility of each of the HL items separately, the 6-SQs present a better self-report approach of screening for inadequate HL in this population.

This study has several limitations. First, we used the s-TOFHLA as our one and only reference measure and all analyses and interpretation are contingent upon the validity of this choice. The s-TOFHLA mainly addresses reading comprehension skills, while the 16-SQ focus on identifying problems in understanding and applying health information. Although, the s-TOFHLA has been the most commonly used measure by others in the field [4, 9, 31], the differences in what these instruments measure should be taken into consideration in interpreting the findings of this study. Second, because we aimed to study the SQ performance in diabetic patients with low socioeconomic status, the generalizability of the findings may be limited, but this reflects the importance of further validation of these questions in more diverse samples of patients. Third, although the parent study was not about health literacy per se, patients with low HL could have avoided participation because the study involves reading and completing a self-administered survey. This would have underestimated the prevalence of limited HL, and could have resulted in a biased assessment of screening performance of these questions. To the degree that this bias exists, we believe it would be even greater for studies whose purpose was to study HL. Fourth, and in relation to the last point, the low prevalence of inadequate and limited HL in this population, and thus the small numbers, might have played a role in the differences in performance of these questions compared to prior literature; however, this has implications on the utility of these questions in populations with low prevalence of inadequate HL. Finally, due to the cross-sectional nature of the data, we were not able to explore test-retest reliability and the longitudinal validity of these questions. Future studies should explore these measurement properties.

4.5. Conclusion

We identified a new set of six items “6-SQ” and a novel scoring system that performed better than standard approaches to screening for inadequate HL in this sample. The weighted summative score of the 6-SQ and the item “problems learning” performed better than the other items or combinations of items in identifying individuals with inadequate HL. We believe that further exploration of the measurement properties of these questions in other populations, and their relative weighting, is needed before recommending any of the brief HL screening instruments for use in clinical practice.

Tables and Figures

Table 4-1: General Characteristics of the Sample

Variable	Overall (N=378) mean or N (SD or %)
Age - years	56.12 (12.35)
Gender- Female	257 (69.09)
Race/ethnicity – African American	303 (83.24)
Education years	12.14 (2.72)
Employment status – unemployed	147 (38.89)
Annual income level	
<$10,000	164 (46.59)
$10,000-$25,000	119 (33.81)
>=$25,000	69 (19.60)
Health insurance – uninsured	141 (39.06)
Diabetes knowledge (DKQ)	15.85 (3.94)
Self-efficacy (PDSMS)	28.78 (5.46)
s-TOFHLA	
Inadequate health literacy	15 (4.72)
Marginal health literacy	16 (5.03)
Adequate health literacy	287 (90.25)
A1C	8.35 (2.44)

DKQ: Diabetes Knowledge Questionnaire; PDSMS: Perceived Diabetes Self-Management Scale; s-TOFHLA: Short form – Test of Functional Health Literacy in Adults; A1c: Glycated hemoglobin

Table 4-2: Correlations of 16-SQ items with the s-TOFHLA, education years, diabetes knowledge, and self-efficacy

Item	s-TOFHLA	Education years	Diabetes Knowledge	Self-Efficacy
HL1	0.25**	0.13*	0.18*	0.12*
HL2	0.16*	0.13*	0.18*	0.20*
HL3	0.18*	0.10	0.26**	0.17*
HL4	0.12*	0.13*	0.16*	0.17*
HL5	0.24**	0.15*	0.13*	0.15*
HL6	0.27**	0.16*	0.16*	0.11*
HL7	0.24**	0.15*	0.12*	0.22*
HL8	0.29**	0.13*	0.12*	0.17*
HL9	0.21**	0.23**	0.17*	0.21*
HL10	0.34**	0.14*	0.19*	0.17*
HL11	0.31**	0.20*	0.26**	0.29**
HL12	0.34**	0.21**	0.26**	0.24**
HL13	0.16*	-0.01	0.12*	0.07
HL14	0.22**	0.17*	0.25**	0.09
HL15	0.24**	0.19*	0.22*	0.09
HL16	0.22**	0.27**	0.20*	0.10
s-TOFHLA	1.00	0.27**	0.51**	0.11
Education years	---	1.00	0.27**	0.09
Diabetes knowledge	---	---	1.00	0.12*
Self-efficacy	---	---	---	1.00

Numbers in the table are Pearson Correlations
* P-value < 0.05
** P-value <0.001

Table 4-3: C-indices (95%CI) for 16-SQ items compared to the s-TOFHLA scores for inadequate and limited HL

Item	**Inadequate HL** C-index (95%CI) Ref: s-TOFHLA	**Limited HL** C-index (95%CI) Ref: s-TOFHLA
HL1*	**0.66** [0.52, 0.81]	**0.64** [0.54, 0.74]
HL2	0.55 [0.39, 0.70]	0.58 [0.47, 0.69]
HL3	0.58 [0.43, 0.74]	0.60 [0.50, 0.71]
HL4	0.58 [0.43, 0.73]	0.59 [0.48, 0.69]
HL5	**0.74** [0.59, 0.88]	0.61 [0.50, 0.71]
HL6	**0.67** [0.53, 0.82]	**0.64** [0.54, 0.73]
HL7*	0.59 [0.43, 0.76]	**0.64** [0.53, 0.74]
HL8	0.71 [0.56, 0.85]	**0.63** [0.53, 0.73]
HL9*	0.58 [0.42, 0.73]	**0.63** [0.53, 0.73]
HL10	**0.68** [0.54, 0.83]	**0.67** [0.57, 0.77]
HL11	0.65 [0.50, 0.80]	**0.68** [0.58, 0.78]
HL12*	**0.71** [0.58, 0.85]	**0.71** [0.61, 0.81]
HL13	0.55 [0.42, 0.68]	0.58 [0.48, 0.67]
HL14*	0.55 [0.40, 0.70]	0.60 [0.49, 0.70]
HL15*	0.62 [0.48, 0.77]	**0.63** [0.53, 0.72]
HL16*	**0.66** [0.51, 0.81]	**0.71** [0.61, 0.80]

***** items that had the highest AUROCs in Chew et al. study (2004); Items HL12, HL14 & HL16 are the 3-SQs

Table 4-4: Rotated factor loadings of the 16-SQ and 6-SQ factor models

		16-SQ 3-factor model			6-SQ 2- factor model	
Item	**Description**	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 1**	**Factor 2**
HL1	How often are appointment slips written in a way that is easy to read and understand?		0.771		0.692	
HL2	How often are medical forms written in a way that is easy to read and understand?		0.827			
HL3	How often are medication labels written in a way that is easy to read and understand?		0.848			
HL4	How often are patient educational materials written in a way that is easy to read and understand?		0.821			
HL5	How often are hospital or clinic signs difficult to understand?	0.689				0.784
HL6	How often are appointment slips difficult to understand?	0.883				
HL7	How often are medical forms difficult to understand and fill out?	0.754				
HL8	How often are directions on medical bottles difficult to understand?	0.870				
HL9	How often do you have difficulty understanding written information your healthcare provider (like a doctor, nurse, nurse practitioner) gives you?	0.778				
HL10	How often do you have problems getting to your clinic appointments at the right time because of difficulty understanding written instructions?	0.652				
HL11	How often do you have problems completing medical forms because of difficulty understanding the instructions?	0.625				
HL12	How often do you have problems learning about your medical condition because of difficulty understanding the written information?	0.618				0.709
HL13	How often are you unsure on how to take your medication(s) correctly because of problems understanding the written instructions on the bottle label?			0.631		
HL14	How confident are you filling out medical forms by yourself?			0.850	0.766	
HL15	How confident do you feel you are able to follow the instructions on the label of a medication bottle?			0.550	0.820	
HL16	How often do you have someone (like family members, friend, hospital/clinic worker, or caregiver) help you read hospital materials?			0.417		0.705

Items HL12, HL14 & HL16 are the 3-SQs

Table 4-5: Test characteristics for the summative scales of the 6-SQ and the 3-SQ for identifying inadequate and limited HL compared to the s-TOFHLA

	Summative score for the 6-SQ					
	Cut-point	**Sensitivity**	**Specificity**	**LR+**	**LR-**	**C-index [95%CI]**
	Reference: s-TOFHLA (Limited HL)					
Summative scale-weighted	>=16	58.62	79.65	2.88	0.52	0.69 (0.60, 0.78)
	>=17	55.17	83.16	3.27	0.54	0.69 (0.60, 0.79)
	>=18	41.38	85.26	2.81	0.69	0.63 (0.54, 0.73)
	>=19	41.38	89.47	3.93	0.65	0.65 (0.56, 0.75)
	>=20	27.59	93.33	4.14	0.77	0.60 (0.52, 0.69)
	>=21	24.14	94.39	4.3	0.8	0.59 (0.51, 0.67)
	Reference: s-TOFHLA (Inadequate HL)					
Summative scale-weighted	>=16	53.85	77.41	2.38	0.6	0.66 (0.51, 0.80)
	>=17	53.85	81.06	2.84	0.57	0.67 (0.53, 0.82)
	>=18	46.15	84.05	2.89	0.64	0.65 (0.51, 0.79)
	>=19	46.15	88.04	3.86	0.61	0.67 (0.53, 0.81)
	>=20	38.46	92.69	5.26	0.66	0.65 (0.52, 0.79)
	>=21	30.77	93.69	4.87	0.74	0.62 (0.49, 0.75)
	Summative score for the 3-SQ					
	Cut-point	**Sensitivity**	**Specificity**	**LR+**	**LR-**	**C-index [95%CI]**
	Reference: s-TOFHLA (Limited HL)					
Summative scale-weighted	>=9	58.06	80.07	2.91	0.52	0.69 (0.60, 0.78)
	>=10	48.39	83.92	3.01	0.61	0.66 (0.57, 0.75)
	>=11	29.03	88.11	2.44	0.8	0.58 (0.50, 0.67)
	>=12	22.58	90.21	2.31	0.86	0.56 (0.49, 0.64)
	Reference: s-TOFHLA (Inadequate HL)					
Summative scale-weighted	>=9	53.33	77.81	2.4	0.6	0.65 (0.52, 0.79)
	>=10	46.67	82.12	2.61	0.65	0.64 (0.51, 0.78)
	>=11	20	86.75	1.51	0.92	0.53 (0.43, 0.64)
	>=12	20	89.4	1.89	0.89	0.55 (0.44, 0.65)

Figures

Figure 4-1: ROC curves of the 3-SQ, and the summative scales of the 3-SQ and 6-SQ compared to the s-TOFHLA for identifying inadequate HL

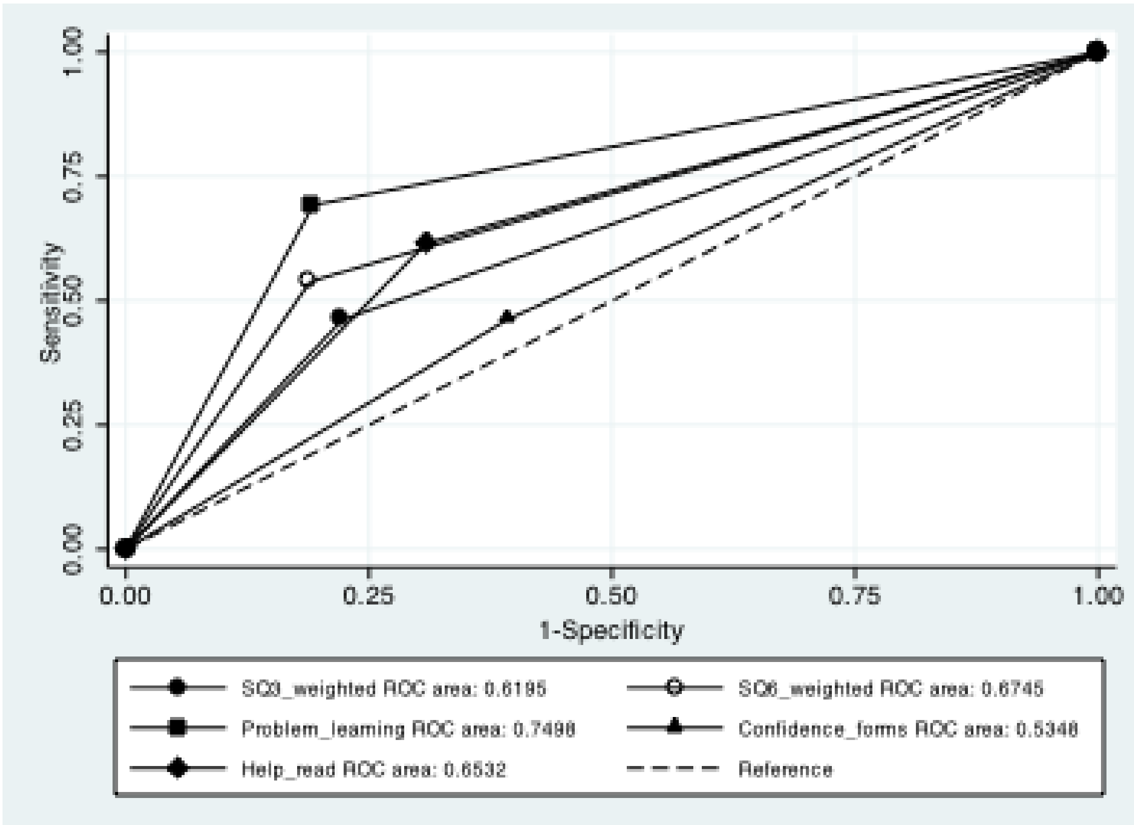
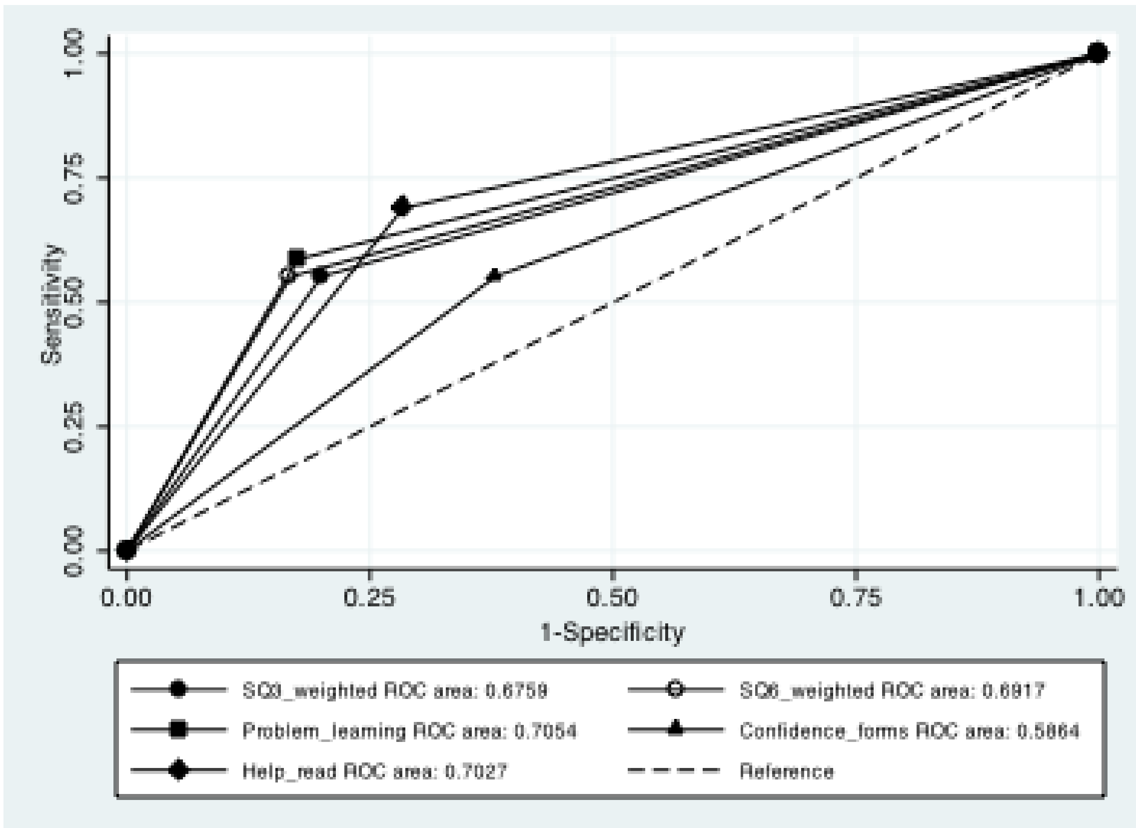


Figure 4-2: ROC curves of the 3-SQ, and the summative scales of the 3-SQ and 6-SQ compared to the s-TOFHLA for identifying limited HL



References

1. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, A.M.A., *Health literacy: Report on the Council of Scientific Affairs.* JAMA, 1999. **282**: p. 526-527.

2. Al Sayah F, et al., *Health Literacy and Health Outcomes in Diabetes: A Systematic Review* JGIM 2012. **28**(3): p. 444-52.

3. Gazmararian JA, et al., *Health literacy and knowledge of chronic disease.* Patient Education and Counseling 2003. **51**: p. 267-275.

4. Schillinger D, et al., *Association of health literacy with diabetes outcomes.* JAMA, 2002. **288**: p. 475-482.

5. Powell CK, Hill EG, and Clancy DE, *The relationship between health literacy and diabetes knowledge and readiness to take health actions.* The Diabetes Educator 2007. **33**: p. 144-151.

6. Paasche-Orlow MK, et al., *The prevalence of limited health literacy.* J Gen Intern. Med. , 2005. **20**(2): p. 175–84.

7. Berkman ND, et al., *Low health literacy and health outcomes: an updated systematic review* Annals of Internal Medicine 2011. **155**(2): p. 97-107.

8. DeWalt DA, et al., *Literacy and Health Outcomes: A Systematic Review of the Literature.* Journal of General Internal Medicine, 2004 **19**(12): p. 1228–1239.

9. Al Sayah F, Williams B, and Johnson JA, *Measuring Health Literacy in Individuals With Diabetes: A Systematic Review and Evaluation of Available Measures* Health Education and Behavior 2012(April).

10. Baker D, et al., *Development of a brief test to measure functional health literacy.* Patient Education and Counseling 1999. **38**: p. 33-42.

11. Weiss BD, et al., *Quick assessment of literacy in primary care: the newest vital sign.* Annals of Family Medicine 2005. **3**: p. 514–522.

12. Davis TC, et al., *Rapid assessment of literacy levels of adult primary care patients* Family Medicine 1991. **23**: p. 433-435.

13. Chew LD, Bradley KA, and Boyko EJ, *Brief questions to identify patients with inadequate health literacy.* Family Medicine, 2004. **36**(8): p. 588-594.

14. Chew LD, et al., *Validation of screening questions for limited health literacy in a large VA outpatient population.* J Gen Intern Med., 2008. **23**(5): p. 561-566.

15. Wallace LS, et al., *Brief report: screening items to identify patients with limited health literacy skills* J Gen Intern Med. , 2006. **21**(8): p. 874-877.

16. Wallace LS, et al., *Can screening items identify surgery patients at risk of limited health literacy?* J Surg Res., 2007. **140**(2): p. 208-13.

17. Morris NS, et al., *The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability.* BMC Family Practice, 2006. **7**(21).

18. Sarkar U, et al., *Validation of self-reported health literacy questions among diverse English and Spanish-speaking populations* J Gen Intern Med., 2011. **26**(3): p. 265-71.

19. Peterson PN, et al., *Health literacy and outcomes among patients with heart failure.* JAMA, 2011. **305**(16): p. 1695-701.

20. Smalls BL, et al., *Associations between coping, diabetes knowledge, medication adherence and self-care behaviors in adults with type 2 diabetes* Gen Hosp Psychiatry 2012. **34**(4): p. 385-9.

21. Parker R, et al., *The test of functional health literacy in adults (TOFHLA): a new instrument for measuring patients' literacy skills* Journal of General Internal Medicine 1995. **10**: p. 537-545.

22. Garcia AA, et al., *The Starr County Diabetes Education Study.* Diabetes Care, 2001. **24**: p. 16–21.

23. Wallston KA, Rothman RL, and Cherrington A, *Psychometric properties of the Perceived Diabetes Self-Management Scale (PDSMS)* J Behav Med. , 2007. **30**(5): p. 395-401.

24. Al Sayah F, et al., *Health Literacy and Health Outcomes in Diabetes: A Systematic Review* JGIM 2012.

25. Berkman, N.D., et al., *Low health literacy and health outcomes: an updated systematic review.* Ann Intern Med, 2011. **155**(2): p. 97-107.

26. Cohen J, *Statistical power analysis for the behavioral sciences*. Second ed1988, Hillsdale, NJ Lawrence Erlbaum Associates Inc. .

27. Steyerberg EW, et al., *Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures.* Epidemiology, 2010. **21**(1): p. 128-38.

28. Hooper D, Coughlan J, and Mullen M R, *Structural Equation Modelling: Guidelines for Determining Model Fit* The Electronic Journal of Buisiness Research Methods 2008 **6**(1): p. 53-60

29. StataCorp, *Stata Statistical Software: Release 12*, 2011, StataCorp LP: College Station, TX.

30. Joreskog K and Sorbom D, *LISREL* 2001, Scientific Software International Lincolnwood, IL.

31. Osborn CY, Bains SS, and Egede LE, *Health literacy, diabetes self-care, and glycemic control in adults with type 2 diabetes* Diabetes Technology and Therapeutics, 2010. **12**: p. 913-919.

Chapter 4 Supplementary Tables

Table S4-1: Test characteristics for the 16-SQ for detecting inadequate HL compared to the s-TOFHLA

Item	Cut-point	Description	Sensitivity	Specificity	LR+	LR-
HL1	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	61.54	68.54	1.96	0.56
	>=3 Sometimes	low=3,4,5	61.54	82.12	3.44	0.47
	>=4 Occasionally	low=4,5	15.38	87.42	1.22	0.97
	>=5 Never	low=5	7.69	90.40	0.80	1.02
HL2	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	46.67	54.79	1.03	0.97
	>=3 Sometimes	low=3,4,5	46.67	72.61	1.70	0.73
	>=4 Occasionally	low=4,5	20.00	87.46	1.59	0.91
	>=5 Never	low=5	13.33	91.42	1.55	0.95
HL3	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	46.67	63.04	1.26	0.85
	>=3 Sometimes	low=3,4,5	40.00	82.51	2.29	0.73
	>=4 Occasionally	low=4,5	20.00	92.74	2.75	0.86
	>=5 Never	low=5	13.33	94.39	2.38	0.92
HL4	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	53.33	58.09	1.27	0.80
	>=3 Sometimes	low=3,4,5	40.00	79.87	1.99	0.75
	>=4 Occasionally	low=4,5	13.33	91.75	1.62	0.94
	>=5 Never	low=5	6.67	94.39	1.19	0.99
HL5	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__

	>=2 Occasionally	low=2,3,4,5	73.33	57.43	1.72	0.46
	>=3 Sometimes	low=3,4,5	73.33	72.61	2.68	0.37
	>=4 Often	low=4,5	46.67	88.45	4.04	0.60
	>=5 Always	low=5	40.00	92.41	5.27	0.65
HL6	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	53.33	76.24	2.24	0.61
	>=3 Sometimes	low=3,4,5	53.33	85.48	3.67	0.54
	>=4 Often	low=4,5	26.67	93.40	4.04	0.78
	>=5 Always	low=5	20.00	95.38	4.33	0.84
HL7	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	53.33	54.79	1.18	0.85
	>=3 Sometimes	low=3,4,5	53.33	70.96	1.84	0.66
	>=4 Often	low=4,5	26.67	88.78	2.38	0.83
	>=5 Always	low=5	20.00	93.73	3.19	0.85
HL8	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	60.00	73.60	2.27	0.54
	>=3 Sometimes	low=3,4,5	60.00	83.50	3.64	0.48
	>=4 Often	low=4,5	33.33	93.73	5.31	0.71
	>=5 Always	low=5	20.00	96.04	5.05	0.83
HL9	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	46.67	62.25	1.24	0.86
	>=3 Sometimes	low=3,4,5	46.67	76.82	2.01	0.69
	>=4 Often	low=4,5	13.33	93.05	1.92	0.93
	>=5 Always	low=5	6.67	95.03	1.34	0.98
HL10	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__

	>=2 Occasionally	low=2,3,4,5	53.33	80.53	2.74	0.58
	>=3 Sometimes	low=3,4,5	53.33	85.81	3.76	0.54
	>=4 Often	low=4,5	26.67	94.39	4.75	0.78
	>=5 Always	low=5	13.33	96.04	3.67	0.90
HL11	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	60.00	61.59	1.56	0.65
	>=3 Sometimes	low=3,4,5	60.00	75.17	2.42	0.53
	>=4 Often	low=4,5	20.00	91.72	2.42	0.87
	>=5 Always	low=5	13.33	95.03	2.68	0.91
HL12	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	66.67	70.96	2.29	0.47
	>=3 Sometimes	low=3,4,5	66.67	80.86	3.48	0.41
	>=4 Often	low=4,5	13.33	93.73	3.13	0.92
	>=5 Always	low=5	13.33	96.04	3.37	0.90
HL13	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	33.33	75.42	1.35	0.88
	>=3 Sometimes	low=3,4,5	33.33	82.39	1.89	0.81
	>=4 Often	low=4,5	13.33	90.70	1.43	0.95
	>=5 Always	low=5	6.67	93.36	1.00	0.99
HL14	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	53.33	52.98	1.13	0.88
	>=3 Sometimes	low=3,4,5	53.33	60.60	1.35	0.77
	>=4 Occasionally	low=4,5	33.33	70.20	1.12	0.95
	>=5 Never	low=5	33.33	79.80	1.65	0.83
HL15	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__

	>=2 Often	low=2,3,4,5	46.67	74.83	1.85	0.71
	>=3 Sometimes	low=3,4,5	46.67	82.45	2.66	0.65
	>=4 Occasionally	low=4,5	26.67	88.41	2.30	0.83
	>=5 Never	low=5	26.67	90.40	2.78	0.81
HL16	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	—
	>=2 Occasionally	low=2,3,4,5	66.67	56.62	1.54	0.59
	>=3 Sometimes	low=3,4,5	66.67	69.21	2.16	0.48
	>=4 Often	low=4,5	33.33	82.12	1.86	0.81
	>=5 Always	low=5	33.33	87.75	2.72	0.76

Table S4-2: Test characteristics for the 16-SQ for detecting limited HL compared to the s-TOFHLA

Item	Cut-point	Description	Sensitivity	Specificity	LR+	LR-
HL1	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	55.17	69.58	1.81	0.64
	>=3 Sometimes	low=3,4,5	48.28	83.22	2.88	0.62
	>=4 Occasionally	low=4,5	24.14	88.46	2.09	0.86
	>=5 Never	low=5	13.79	90.91	1.52	0.95
HL2	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	51.61	55.40	1.16	0.87
	>=3 Sometimes	low=3,4,5	45.16	73.52	1.70	0.74
	>=4 Occasionally	low=4,5	29.03	88.85	2.60	0.80
	>=5 Never	low=5	19.35	92.33	2.52	0.87
HL3	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	51.61	64.11	1.44	0.75
	>=3 Sometimes	low=3,4,5	38.71	83.62	2.36	0.73
	>=4 Occasionally	low=4,5	16.13	93.03	2.31	0.90
	>=5 Never	low=5	9.68	94.43	1.73	0.96
HL4	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	51.61	58.54	1.24	0.83
	>=3 Sometimes	low=3,4,5	41.94	81.18	2.23	0.71
	>=4 Occasionally	low=4,5	16.13	92.33	2.10	0.91
	>=5 Never	low=5	9.68	94.77	1.85	0.95
HL5	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	58.06	57.49	1.36	0.73

	>=3 Sometimes	low=3,4,5	45.16	72.13	1.62	0.76
	>=4 Often	low=4,5	29.03	88.50	2.54	0.80
	>=5 Always	low=5	22.58	92.33	2.94	0.84
HL6	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	48.39	77.35	2.14	0.67
	>=3 Sometimes	low=3,4,5	38.71	86.06	2.78	0.71
	>=4 Often	low=4,5	19.35	93.73	3.09	0.86
	>=5 Always	low=5	9.68	95.12	1.98	0.95
HL7	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	61.29	56.10	1.40	0.69
	>=3 Sometimes	low=3,4,5	54.84	72.47	1.99	0.62
	>=4 Often	low=4,5	29.03	89.90	2.87	0.79
	>=5 Always	low=5	19.35	94.43	3.47	0.85
HL8	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	48.39	74.22	1.88	0.69
	>=3 Sometimes	low=3,4,5	41.94	83.97	2.62	0.69
	>=4 Often	low=4,5	22.58	94.08	3.81	0.82
	>=5 Always	low=5	12.90	96.17	3.37	0.90
HL9	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	58.06	63.99	1.61	0.65
	>=3 Sometimes	low=3,4,5	41.94	77.62	1.87	0.75
	>=4 Often	low=4,5	19.35	94.06	3.26	0.86
	>=5 Always	low=5	9.68	95.45	2.13	0.95
HL10	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	48.39	81.88	2.67	0.63

	>=3 Sometimes	low=3,4,5	48.39	87.46	3.86	0.59
	>=4 Often	low=4,5	25.81	95.47	5.70	0.78
	>=5 Always	low=5	12.90	96.52	3.70	0.9
HL11	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	64.52	63.29	1.76	0.56
	>=3 Sometimes	low=3,4,5	58.06	76.92	2.52	0.54
	>=4 Often	low=4,5	25.81	93.01	3.69	0.80
	>=5 Always	low=5	16.13	95.80	3.84	0.87
HL12	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	64.52	72.82	2.37	0.49
	>=3 Sometimes	low=3,4,5	58.06	82.58	3.33	0.51
	>=4 Often	low=4,5	19.35	94.77	3.70	0.85
	>=5 Always	low=5	16.13	96.86	5.14	0.86
HL13	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	38.71	76.49	1.65	0.80
	>=3 Sometimes	low=3,4,5	32.26	83.16	1.91	0.81
	>=4 Often	low=4,5	16.13	91.23	1.84	0.92
	>=5 Always	low=5	6.45	93.33	0.97	1.00
HL14	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	61.29	54.20	1.34	0.71
	>=3 Sometimes	low=3,4,5	58.06	61.89	1.52	0.68
	>=4 Occasionally	low=4,5	38.71	70.98	1.33	0.86
	>=5 Never	low=5	35.48	80.77	1.84	0.80
HL15	>=1 Always	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Often	low=2,3,4,5	48.39	76.22	2.03	0.68

	>=3 Sometimes	low=3,4,5	45.16	83.92	2.81	0.65
	>=4 Occasionally	low=4,5	19.35	88.46	1.68	0.91
	>=5 Never	low=5	16.13	90.21	1.65	0.93
HL16	>=1 Never	low=1,2,3,4,5	100.00	0.00	1.00	__
	>=2 Occasionally	low=2,3,4,5	74.19	58.74	1.80	0.44
	>=3 Sometimes	low=3,4,5	70.97	71.68	2.50	0.40
	>=4 Often	low=4,5	35.48	83.22	2.11	0.77
	>=5 Always	low=5	35.48	89.16	3.27	0.72

Chapter 5

Lack of Association Between Inadequate Health Literacy and Health Outcomes in Patients with Type 2 Diabetes and Depression: Secondary Analysis of a Controlled Trial ‡

Abstract

Background: Inadequate health literacy (HL) and depression have each been associated with poor diabetes-related outcomes, but cross-sectional studies have provided conflicting evidence of the association between inadequate HL and depressive symptoms and other outcomes in individuals with type 2 diabetes (T2DM).

Objective: To examine the influence of inadequate HL on changes in depressive symptoms, health-related quality of life, and cardio-metabolic outcomes among patients with T2DM recently screened positive for depression.

Design, Setting and Patients: Secondary analysis of data from a neutral trial (N=154) that compared a collaborative team care model and enhanced usual care for primary care T2DM patients and depression.

Outcome Measures: Exposure of interest was inadequate HL, defined as a total summative score of 9 or more on the 3 brief screening questions. Outcomes of interest were differences in the changes in depressive symptoms (Patient Health Questionnaire-9 (PHQ-9) at 12-months, health-related quality of life (SF-12 and EQ-5D), glycemic control (A1c), low-density lipoprotein cholesterol (LDL), and systolic blood pressure (SBP).

---

‡ A version of this chapter has been submitted for publication. Co-authors: Sumit R Majumdar[1], MD MPH; Jeffrey A Johnson[2], PhD ([1] Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada; [2] School of Public Health, University of Alberta, Edmonton, Alberta, Canada).

Results: Average age of patients was 58 years, 56% women, and predominantly white. Only a small proportion (N=24; 16%) had inadequate HL. In adjusted random effects models comparing patients with inadequate and adequate HL, all outcomes were neither statistically significant nor clinically important. The between group differences were -0.52 points for PHQ-9, -0.66 points for physical and -0.53 points for mental summaries of the SF-12, -0.001 points for EQ-5D, -0.14% for A1c, -0.02 mmol/L for LDL, and -0.33 mmHg for SBP.

Conclusion: Among primary care patients with T2DM who recently screened positive for depression, inadequate health literacy is not associated with worse outcomes over 1 year.

5.1. Introduction

Health literacy (HL) is the degree to which individuals can obtain, process, understand, and communicate health-related information needed to make informed health care decisions [1]. A growing body of research suggests that inadequate HL is associated with adverse effects on health, such as lack of use of preventive services [2], delayed diagnoses [3], misunderstandings about medical conditions [4], less compliance with medical instructions, decreased self-management skills [5], poor self-rated health [6], higher health care costs [7], and increased mortality risk [8, 9].

Inadequate HL is common among individuals with type 2 diabetes (T2DM) and has been associated with diabetes outcomes such as worse glycemic control [10, 11] and increased risk of hypoglycemia [12]. Depression, another complex chronic condition, affects 10-30% of individuals with T2DM [13], making it one of the most common comorbidities in this patient population. Comorbid depression is known to limit self-care, lead to worse health outcomes and greater health care costs [14-16]. With this dual complexity, the demands for self-management in individuals with T2DM and depression are increased, and the ability to execute required tasks is challenged [15]. Given the increased self-management demands on these patients, in the face of a complex and dynamic healthcare environments, exploring the influence of inadequate HL on health outcomes in this population is crucial.

Despite growing evidence of the effects of inadequate HL on health outcomes, the evidence is limited and inconsistent in the diabetes population [17]. Cross-sectional studies have provided conflicting evidence of the association of inadequate HL with diabetes-related outcomes [17]. Given that both inadequate HL and depression are potentially modifiable, it is surprising that there are no rigorous longitudinal data to examine the effect of inadequate HL on depressive symptoms and other health outcomes in this patient population. Our objective was to examine the longitudinal associations of inadequate HL with depression-related and other health outcomes in patients with T2DM who recently screened positive for depression in a clinical trial.

5.2. Methods

5.2.1. Study Population and Recruitment

This longitudinal observational study used data from the TeamCare-PCN trial [18], conducted to evaluate the effectiveness of a collaborative team care model for primary care patients with T2DM [18]. Briefly, the goal of the intervention was to reduce depressive symptoms, achieve targets for cardio-metabolic measures and improve lifestyle behaviours. The intervention was coordinated by a nurse care manager, working in collaboration with family physicians and specialists (i.e., psychiatrists and internists). Participants for the TeamCare-PCN study were recruited from four Primary Care Networks (PCN) in rural Alberta. Comparing this model to an enhanced usual care (i.e., screening for depressive symptoms and general practitioner notification), the results of the trial were robustly negative, permitting us to examine the natural history of newly identified depressive symptoms in T2DM by treating our sample as a single longitudinal cohort [19]. All study procedures were approved by the University of Alberta Health Research Ethics Board (PRO#00012663), and all participants provided written informed consent.

5.2.2. Measurements

The TeamCare-PCN trial involved three data collection points: baseline, 6-, and 12 months. Anthropometric and clinical measurements data were collected by the care managers at the PCNs. Self-reported data was collected through a self-administered survey mailed to all participants at each data collection point. Data on the following measures were used in this study:

Health literacy was assessed using a previously validated set of three brief questions screening for inadequate HL [20]. The questions were: 1) How often do you have problems learning about your medical condition because of difficulty understanding written information?; 2) How confident are you filling out medical forms by yourself?; and 3) How often do you have someone like a family member, friend, hospital or clinic worker or caregiver help you read health plan materials (such as written information about your health or care you are offered)? These questions were each scored on a five-point Likert scale with higher scores indicating lower HL. We used a weighted summative score of the three HL items, and stratified summative scores such that < 9.0 was identified as adequate HL and >= 9.0 as inadequate HL [21].

Depressive symptoms were assessed using the Patient Health Questionnaire-9 items (PHQ-9) [22, 23]. An overall score is computed by adding scores of all items, ranging from 0 to 27, with higher scores indicating more severe depressive symptoms [23]. Patients enrolled in the trial had at least moderate depressive symptoms at baseline, based on a score of 10 or more [22, 23]. Improvements of 5 points are considered clinically important [24]. Remission of depressive symptoms was defined as a PHQ-9 score < 10 at 12 months.

HRQL was assessed using SF-12 version 2 and EQ-5D-5L. The SF-12 contains 12 items from the eight scales of the SF-36 health status inventory [25, 26]. These 12 items were used to generate oblique physical (PCS) and mental (MCS) component scores [27]; 5 points difference is considered important on these scores. The EQ-5D-5L is a preference-based index of HRQL, based on five dimensions (mobility, self-care, usual activity, pain/discomfort, and anxiety/ depression), each with five levels (no problem, slight problem, moderate problem, severe problem, and unable/extreme problem) [28]. A difference of 0.03 is considered important on the EQ-5D index [28].

Cardio-metabolic outcomes: Included A1c, LDL cholesterol, and systolic blood pressure (SBP). Capillary blood samples were assessed for A1c and lipid profiles using point-of-care devices; blood pressure readings were obtained using the BPTru automated system [18]. Clinically important improvements in A1c, LDL and SBP were defined as 10% reduction or more at 12-months [29].

Socio-demographic covariates collected were age, sex, self-identified race/ethnicity (categorized as white or non-white), level of education (categorized as less than high school or high school or more), employment status (categorized as employed or unemployed), and annual household income (categorized as <$40,000, $40,000 - $60,000, and >=$80,000). Other covariates included self-reported co-morbidities and diabetes complications. Body mass index (BMI) was calculated using measured height and weight.

5.2.3. Statistical Analysis

Descriptive analyses were performed on all baseline variables, comparing patients by HL status using t tests for continuous variables and chi square test for categorical variables. Mixed-effect repeated measures generalized linear models with random effects for each subject were used to examine the independent association of HL with the change in each of the outcomes. Each model was set up with change in the respective outcome as the dependent variable and HL status as the main explanatory variable. We also used logistic regression models to examine the independent association between HL status and remission of depressive symptoms and improvement in HRQL indicators and cardio-metabolic outcomes at 12 months.

In all models (i.e., linear and logistic), we included as covariates those that were significantly different between the HL groups at baseline (i.e., sex, MCS), and variables known to be related to HL or health outcomes (age, educational level, number of comorbidities, number of diabetes complications), as well as allocation status and time. Percentage of missing data was less than 5% in most variables. Mean imputation was used for missing data at baseline for A1c, LDL, SBP, and HL score. To identify values to impute for missing data in EQ-5D and SF-12 items, a “crosswalk approach” between these two measures was used, i.e. imputing missing data on one measure based on available responses for similar items in the other. Missing data at follow-up assessments were handled using last observation carried forward (LOCF) method [18]. For each analysis, the null hypothesis was evaluated at a 2-sided significance level of .05. All analyses were performed using STATA 11.1 [30].

5.3. Results

5.3.1. General Characteristics

The TeamCare-PCN trial enrolled 157 patients; 62 allocated to the intervention arm, and 95 to the enhanced usual care arm. Follow-up data were available for 112 of the 157 participants (71%) at 12 months. Three participants did not complete the survey (including the HL items) at any time point, and were therefore excluded from this analysis. Thus, the total number of participants included in this study was 154.

The average age of participants was 58.1 years (SD=9.4), more than half females (56%), predominantly white (94%), the majority had high school education or more (86%), around half (48%) unemployed, and fairly distributed across income categories (Table 5-1). At baseline, the participants generally had moderate depressive symptoms (mean PHQ-9 = 14.5; SD=3.7), were obese (mean BMI = 36.8; SD=8.1), but had relatively good cardio-metabolic control. Additionally, participants had low self-reported HRQL at baseline (mean EQ-5D = 0.7; SD=0.2; mean PCS=35.5; SD=10; mean MCS=35.1; SD=7.7).

The mean weighted summative score of the three brief screening questions of inadequate HL was 5.9 (SD 2.6; median 5.6; interquartile range 3.4 – 7.5). Around 16% of participants had inadequate HL (weighted summative score >= 9.0), and were equally distributed between the trial groups. Participants with inadequate HL were more likely to be male and had higher MCS scores at baseline than participants with adequate HL. Patients with inadequate HL had lower education and income compared to those with adequate HL, although these differences were not statistically significant.

5.3.2. Change in Depressive Symptoms

Participants in both HL groups had significant and important improvements in depressive symptoms from baseline to 12-months, with the majority of the improvement occurring in the first 6 months (Figure 5-1a). There were no differences in the 12-month change of PHQ-9 score between the HL groups. In a random effects model, the adjusted difference in the average change of PHQ-9 scores between the HL groups (-0.52) was neither statistically significant (P = 0.652) nor clinically important (Table 5-2). Additionally, there was no difference in the odds of remission of depressive symptoms at 12 months between the HL groups (Table 5-3).

5.3.3. Changes in Health Related Quality of Life

Participants in both HL groups had modest improvements in all HRQL indicators from baseline to 12-months (Figure 5-1b, 5-1c, 5-1d). In adjusted random effects models, the difference between the HL groups in the average change of PCS (-0.66), MCS (-0.53), and EQ-5D (-0.001) were neither statistically significant nor of important magnitude (Table 5-2). Additionally, there were no differences in the odds of improvement in PCS and MCS at 12 months between the HL groups (Table 5-3).

5.3.4. Changes in Cardio-metabolic Outcomes

For cardio-metabolic outcomes, participants in both HL groups had minimal improvements in A1c (Figure 5-1e), LDL (Figure 5-1f), and SBP (Figure 5-1g) from baseline to 12-months. In adjusted random effects models, the difference between the HL groups in the average change of A1c (-0.14), LDL (-0.02), and

SBP (-0.03) were statistically non-significant and clinically unimportant (Table 5-2). Additionally, there were no differences in the odds of improvement in A1c, LDL and SBP at 12 months between the HL groups (Table 5-3).

5.4. Discussion

In this study, inadequate HL, as assessed by the 3-brief screening questions, was present in 1 in 6 primary care patients with T2DM who recently screened positive for depression. The estimated prevalence of inadequate HL in this population is higher than that in the general population or in any of these conditions separately, using the same measure of HL [31]. Nonetheless, our study demonstrates that inadequate HL did not negatively influence changes in their health outcomes. In particular, we did not find any evidence of independent associations of inadequate HL with changes or improvements in depressive symptoms, HRQL, or cardio-metabolic outcomes.

Research linking HL with health outcomes has been mixed. We previously reported a systematic review demonstrating that evidence on the association between HL and health outcomes in the diabetes population was insufficient and inconsistent across studies for most outcomes [17]. Although it was previously suggested that that people with low HL are more likely to have symptoms of depression [32-34], we found no association between HL and change in depressive symptoms in this study. This is consistent with what has been reported earlier for diabetes patients [35]. Our results, which showed no association between HL and clinical outcomes like A1c and SBP, are in contrast to those prior studies that showed that higher levels of HL were associated with better glycemic control [11, 36, 37], but consistent with others that did not find an association [35, 38]. Our finding of lack of association between HL and HRQL in this study is also consistent with previous reports for diabetic patients [39].

Discrepancies between this and previous studies, especially for glycemic control, may be related to study design, instruments used to measure HL, populations involved, or other factors. For the most part, the previous literature has been limited to cross-sectional studies, while ours is one of the few to examine longitudinal associations between HL and these health outcomes in the diabetes population, thus providing more robust evidence. Despite their limitations, the 3-brief HL screening questions have been shown to provide valid estimates of inadequate HL in different chronic disease patients including those with diabetes [40-42]. We previously examined the utility of these questions compared to the s-TOFHLA in examining the cross-sectional associations between HL and health outcomes in a similar population with T2DM and found similar results, regardless of the measure used [43].

Although inadequate HL might be associated with worse health outcomes in the general population or in some patient groups, as suggested by some cross-sectional studies, it does not appear to have an impact in “sicker” populations [43]. Previous research suggested that low HL is significantly associated with higher levels of social support in T2DM patients [44] as well as higher rates of healthcare utilization [7, 31]. It might be that sicker patients with inadequate HL, such as this group of obese T2DM patients with depressive symptoms (and at least two other comorbidities and diabetes complications), seek and receive appropriate care, services and support from their healthcare providers and possibly significant others (e.g., providers, family, friends). Thus, it is possible that inadequate HL may not have had an adverse impact on health management and ultimately outcomes in these individuals.

Several factors should be considered in the interpretation of our findings. First, participants in the TeamCare-PCN trial were recruited through an invitation letter and brief screening survey that were mailed to them by PCNs. By design, this recruitment approach could have excluded individuals with very low HL levels, such as those unable to read the invitation letter and survey, or who read them but did not understand them, and therefore they did not participate. This bias could have lead to recruiting fewer individuals with low levels of HL, and thus the reported prevalence could be an underestimation of the true prevalence among this population. Second, this study was not powered to detect differences by HL status. Based on a 2-sided significance level of .05 and 80% power, we estimated a range of sample sizes from 55 to 63 patients per group to detect a difference corresponding to a moderate effect size of 0.5 in the average change of the outcomes between the HL groups. Although this study was underpowered, the magnitude of the differences between the HL groups was small and clinically unimportant for all outcomes. Third, the small number of patients with inadequate HL and the unbalanced group size limited our ability to draw definitive conclusions. Fourth, we did not consider changes in HL over the course of the follow-up, which, if HL is modifiable and amenable to intervention, would be a worthwhile area for future study. Finally, this study was conducted among predominantly Caucasian English-speaking patients and may not be generalizable to non-white non-English-speaking individuals.

5.5. Conclusion

This study demonstrates, through a rigorous longitudinal design, that inadequate HL does not play a role in improving depressive symptoms, HRQL, or cardio-metabolic outcomes among T2DM patients who screened positive for depression. The presumption that the impact of HL on health outcomes might vary by the severity of the health condition of populations studied is an important one and warrants further exploration.

Tables and Figures

Table 5-1: General Characteristics of the overall sample and by health literacy (HL) level at baseline

Characteristic	**Overall (N = 154)** N or mean (% or SD)	**Adequate HL (N =130; 84.4%)** N or mean (% or SD)	**Inadequate HL (N =24; 15.6%)** N or mean (% or SD)	**P-value**
Age – years	58.1 (9.4)	58.6 (9.2)	55.2 (10.2)	0.102
Female sex	86 (55.8)	78 (60.0)	8 (33.3)	**0.016**
Ethnicity/Race-White	130 (94.2)	108 (93.1)	22 (100.0)	0.204
Education - less than high school	21 (13.7)	15 (11.6)	6 (25.0)	0.080
Employment status - Unemployed	73 (48.0)	61 (47.3)	12 (52.2)	0.666
Income				
<$40,000	44 (36.1)	38 (36.5)	6 (33.3)	
$40,000-$80,000	39 (32.0)	30 (28.9)	9 (50.0)	0.154
>= $80,000	39 (32.0)	36 (34.6)	3 (16.7)	
Number of diabetes complications	2.4 (1.9)	2.4 (1.9)	2.6 (1.8)	0.620
Number of comorbidities	2.9 (1.5)	3.0 (1.5)	2.5 (1.2)	0.117
BMI	36.8 (8.1)	36.6 (7.9)	37.8 (9.1)	0.503
Health-related Quality of Life				
PCS	35.5 (10.0)	35.4 (9.7)	36.9 (10.7)	0.477
MCS	35.2 (7.6)	34.5 (7.3)	38.7 (8.3)	**0.012**
EQ-5D index score	0.70 (0.16)	0.70 (0.16)	0.68 (0.17)	0.631
PHQ-9 score	14.5 (3.7)	14.7 (3.7)	13.8 (3.5)	0.295
Depression status				
Moderate (10 - 14)	90 (58.4)	75 (57.7)	15 (64.5)	
Moderately severe (15 - 19)	44 (28.6)	37 (28.5)	7 (29.2)	0.757
Severe (>= 20)	20 (13.0)	18 (13.8)	2 (8.3)	
A1c	7.6 (1.8)	7.6 (1.8)	7.6 (1.6)	0.938
LDL	2.2 (0.8)	2.2 (0.8)	2.0 (0.7)	0.246
SBP	125.4 (15.9)	125.6 (15.7)	124.2 (17.6)	0.686

BMI: body mass index; PHQ-9: Patient Health Questionnaire – 9 items; PCS: physical component score; MCS: mental component score; LDL: Low-density lipoprotein; SBP: systolic blood pressure

Table 5-2: Adjusted* random effects models for change in health outcomes by health literacy level

	Coefficient	SE	95%CI	P-value
PHQ-9	-0.52	1.16	[-2.79, 1.75]	0.652
PCS	-0.66	1.46	[-3.53, 2.20]	0.651
MCS	-0.53	1.77	[-4.00, 2.95]	0.767
EQ-5D	-0.001	0.02	[-0.04, 0.04]	0.978
A1c	-0.14	0.24	[-0.60, 0.32]	0.555
LDL	-0.02	0.14	[-0.29, 0.25]	0.868
SBP	-0.33	3.42	[-7.02, 6.37]	0.924

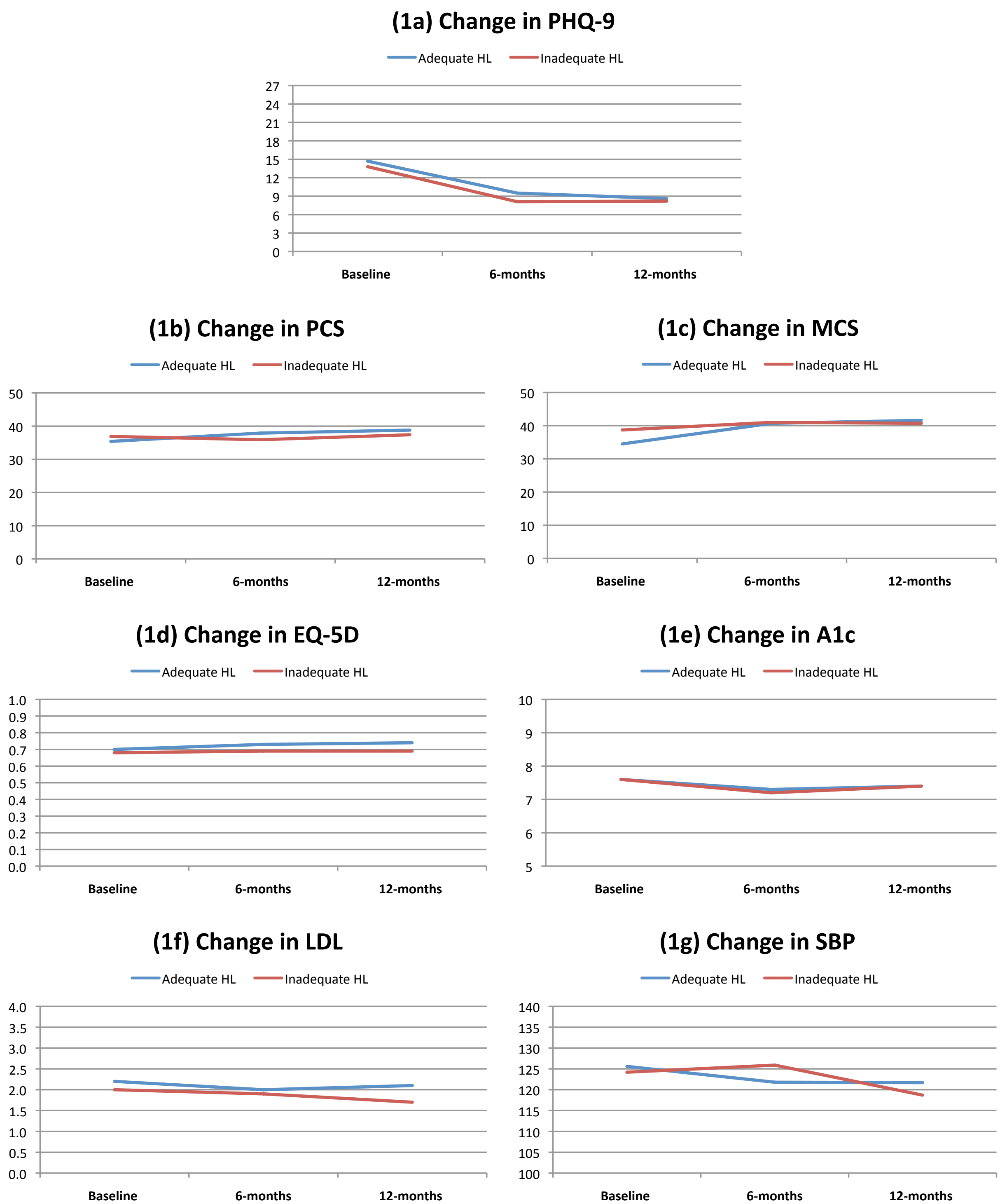
* All models were adjusted for age, sex, education, allocation status, time, baseline MCS score, and number of comorbidities and number of complications at baseline

Table 5-3: Adjusted* logistic regression models for improvement in health outcomes at 12 months by health literacy level

	OR	95%CI	P-value
PHQ-9 Remission	0.92	[0.34, 2.45]	0.861
PCS Improvement	0.81	[0.28, 2.38]	0.706
MCS Improvement	1.44	[0.54, 3.84]	0.464
EQ-5D Improvement	0.74	[0.28, 1.91]	0.533
A1c Improvement	0.41	[0.09, 1.99]	0.272
LDL Improvement	1.56	[0.55, 4.39]	0.400
SBP Improvement	1.13	[0.25, 5.15]	0.876

* All models were adjusted for age, sex, education, allocation status, baseline MCS score, and number of comorbidities and number of complications at baseline

Figure 5-1: 12-months change in health outcomes by health literacy (HL) level



References

1. Berkman N, Terry D, and McCormack L, *Health literacy: What is it?* Journal of Health Communication, 2010. **15**: p. 9-19.

2. Scott T, et al., *Health literacy and preventive health care use among Medicare enrollees in a managed care organization.* Medical Care, 2002. **40**: p. 395-404.

3. Wolf MS, et al., *Associations of race, literacy, and income with PSA level among men with newly diagnosed prostate cancer.* Urology, 2006. **68**: p. 89-93.

4. Gazmararian JA, et al., *Health literacy and knowledge of chronic disease.* Patient Education and Counseling 2003. **51**: p. 267-275.

5. Williams M, et al., *Relationship of functional health literacy to patients' knowledge of chronic disease: A study of patients with hypertension and diabetes* Archives of Internal Medicine 1998. **158**: p. 166-172.

6. Bennett IM, et al., *The contribution of health literacy to disparities in self-rated health status and preventive health behaviors in older adults.* Ann Fam Med, 2009. **7**(3): p. 204-11.

7. Howard D, Gazmararian JA, and Parker R, *The impact of low health literacy on the medical costs of Medicare managed care enrollees.* American Journal of Medicine, 2005. **118**: p. 371-377.

8. Bostock S and Steptoe A, *Association between low functional health literacy and mortality in older adults: longitudinal cohort study* BMJ 2012 (March).

9. Baker D, et al., *Health literacy and mortality among elderly persons.* Archives of Internal Medicine, 2007. **167**: p. 1503-1509.

10. Powell CK, Hill EG, and Clancy DE, *The relationship between health literacy and diabetes knowledge and readiness to take health actions.* The Diabetes Educator 2007. **33**: p. 144-151.

11. Schillinger D, et al., *Association of health literacy with diabetes outcomes.* JAMA, 2002. **288**: p. 475-482.

12. Sarkar U, et al., *Hypoglycemia is more common among type 2 diabetes patients with limited health literacy: The diabetes Study of Northern California (DISTANCE).* Journal of General Internal Medicine 2010. **25**: p. 962-967.

13. Ali S, et al., *The prevalence of co-morbid depression in adults with type 2 diabetes: a systematic review and meta-analysis.* Diabetic Medicine, 2006. **23**(11): p. 1165-1173.

14. Black SA, Markides KS, and Ray LA, *Depression predicts increased incidence of adverse health outcomes in older Mexican Americans with type 2 diabetes.* Diabetes Care, 2003. **26**(10): p. 2822-8.

15. Lin EH, et al., *Relationship of depression and diabetes self-care, medication adherence, and preventive care.* Diabetes Care, 2004. **27**(9): p. 2154-60.

16. Ciechanowski PS, Katon WJ, and Russo JE, *Depression and diabetes: impact of depressive symptoms on adherence, function, and costs.* Arch Intern Med, 2000. **160**(21): p. 3278-85.

17. Al Sayah F, et al., *Health Literacy and Health Outcomes in Diabetes: A Systematic Review* JGIM 2012. **28**(3): p. 444-52.

18. Johnson JA, et al., *Controlled trial of a collaborative primary care team model for patients with diabetes and depression: Rationale and design for comprehensive evaluation.* BMC Health Serv Res, 2012. **12**(258).

19. Johnson JA, et al. *Limited Effectiveness of Collaborative Primary Care Model for Diabetes and Depression: An Implementation Trial*. in *73rd Annual American Diabetes Association Scientific Sessions*. 2013. Chicago, IL.

20. Chew LD, et al., *Validation of screening questions for limited health literacy in a large VA outpatient population.* Journal of General Internal Medicine 2008. **23**: p. 561-566.

21. Al Sayah F, et al., *Measurement Properties and Comparative Performance of Health Literacy Screening Questions in a Predominantly Low Income African American Population With Diabetes* BMC Health Services Research, Under review.

22. Arroll B, Goodyear-Smith F, and Crengle S, *Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population.* Ann Fam Med, 2010. **8**(4): p. 348-53.

23. Kroenke K, Spitzer RL, and Williams JB, *The PHQ-9: validity of a brief depression severity measure.* Journal of General Internal Medicine, 2001. **16**(9): p. 606-13.

24. Löwe B, et al., *Monitoring Depression Treatment Outcomes With the Patient Health Questionnaire-9.* Med Care, 2004. **42**(12): p. 1194-201.

25. Johnson JA and Maddigan SL, *Performance of the RAND-12 and SF-12 summary scores in type 2 diabetes.* Quality of Life Research, 2004. **13**: p. 449-456.

26. Hays RD, *RAND-36 Health Status Inventory.* , 1998, The Psychological Corporation San Antonio: Texas

27. Fleishman JA, Selim AJ, and Kazis LE, *Deriving SF-12v2 physical and mental health summary scores: a comparison of different scoring algorithms.* Quality of Life Research, 2010. **19**(2): p. 231-241.

28. Herdman M, et al., *Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L).* Qual Life Res., 2011. **20**(10): p. 1727-36

29. Majumdar SR, et al., *A Canadian consensus for the standardized evaluation of quality improvement interventions in type-2 diabetes.* Can J Diabetes, 2005. **29**: p. 220-229.

30. StataCorp, *Stata Statistical Software: Release 12*, 2011, StataCorp LP: College Station, TX.

31. Al Sayah F, Johnson JA, and Majumdar SR, *Health Literacy in Alberta: Initial Results from the 2012 HQCA "Satisfaction & Experience with Health Care Services" Survey*, In Progress.

32. Coffman M and Norton C, *Demands of Immigration, Health Literacy, and Depression in Recent Latino Immigrants.* Home Health Care Management and Practice 2010. **22**: p. 116-122.

33. Nokes KM, et al., *Health literacy and health outcomes in HIV seropositive persons.* Research in Nursing and Health 2007. **30**(6): p. 620-627.

34. Lincoln A, et al., *Impact of Health Literacy on Depressive Symptoms and Mental Health-related: Quality of Life Among Adults with Addiction.* Journal of General Internal Medicine, 2006. **21**(8): p. 818-822.

35. Mancuso JM, *Impact of health literacy and patient trust on glycemic control in an urban USA population.* Nursing and Health Sciences, 2010. **12**: p. 94-104.

36. Ishikawa H and Yano E., *The relationship of patient participation and diabetes outcomes for patients with high vs. low health literacy.* Patient Education and Counseling 2011. **in press**.

37. Tang YH, et al., *Health literacy, complication awareness, and diabetic control in patients with type 2 diabetes mellitus* AJAN, 2008. **62**: p. 74-83.

38. Bains SS and Egede LE, *Associations between health literacy, diabetes knowledge, self-care behaviors, and glycemic control in a low income population with type 2 diabetes.* Diabetes Technology and Therapeutics 2011. **13**: p. 335-341.

39. DeWalt DA, Boone RS, and Pignone MP, *Literacy and Its Relationship With Self-efficacy, Trust, and Participation in Medical Decision Making.* American Journal of Health Behavior, 2007 **31**(Suppl 1): p. S27-S35.

40. Peterson PN, et al., *Health literacy and outcomes among patients with heart failure.* JAMA, 2011. **305**(16): p. 1695-701.

41. Omachi TA, et al., *Lower health literacy is associated with poorer health status and outcomes in chronic obstructive pulmonary disease.* Journal of General Internal Medicine, 2013. **28**(1): p. 74-81.

42. Sarkar U, et al., *Validation of self-reported health literacy questions among diverse english and spanish-speaking populations.* Journal of general Internal Medicine 2011. **26**: p. 265-271.

43. Al Sayah F, et al., *Inadequate health literacy and health outcomes in a predominantly low-income African-American population with type 2 diabetes* Journal of Health Communication, Under review

44. Osborn CY, Bains SS, and Egede LE, *Health literacy, diabetes self-care, and glycemic control in adults with type 2 diabetes* Diabetes Technology and Therapeutics, 2010. **12**: p. 913-919.

Chapter 6

Nurses' Interaction with Type-2 Diabetes Patients in Primary Care Settings: Insights into Communication Loops, Jargon, and Health Literacy §

Abstract

Objective: To examine the application of interactive communication loop, use of jargon, and the impact of health literacy (HL) when nurses and type-2 diabetes (T2DM) patients discuss self-care management.

Methods: Audio-recording of 36 encounters between nurses and T2DM patients, and administration of a patient survey including a HL measure.

Results: A complete communication loop was noted in only 11% of the encounters. Clarifying health information was the most commonly applied component (58% often used), followed by repeating health information (33% often used). Checking for understanding was the least applied (81% never used), followed by asking for understanding (42% never used). Medical jargon and mismatched language were often used in 17% and 25% of the encounters respectively. Patients' HL did not materially affect patterns of communication in terms of using communication loops, however nurses used less jargon and mismatched words with patients with inadequate HL.

Conclusion: Nurses rarely used a full communication loop but frequently used jargon in providing self-care education for T2DM patients; however the latter was

---

§ A version of this chapter has been submitted for publication. Co-authors: Beverly Williams[1], PhD; Jenelle L Pederson[2], BA; Sumit R Majumdar[2], MD MPH; Jeffrey A Johnson[3], PhD ([1] Faculty of Nursing, University of Alberta, Edmonton, Alberta, Canada; [2] Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada; [3] School of Public Health, University of Alberta, Edmonton, Alberta, Canada).

less common in patients with inadequate HL. Nurses and other healthcare professionals should elicit patient perceptions and apply communication strategies to enhance understanding and recall of health information.

6.1. Introduction

Healthcare providers are key components of every healthcare system, and medical encounters are a key aspect of patient care. Verifying and evaluating patients' understanding during these encounters is one of the most critical elements of effective communication. In particular, the "teach-back" technique is an effective method of verifying understanding in patients, especially those with learning challenges [1-3]. Teach-back involves asking patients to explain the intended message in their own words or demonstrate the target skill being taught [4]. Studies have shown that teach-back methods improve both patient comprehension and retention of information, is associated with better outcomes (at least in type 2 diabetes) [1], and does not result in longer patients' encounters [1, 5, 6]. This approach, also known as the interactive communication loop [1], is useful to assess and enhance patient's comprehension and recall, reveal health beliefs, reinforce and tailor health messages, and activate patients by opening a dialogue [1].

Clear and plain communication is as important as applying communication loops to ensure patient understanding. Previous studies have shown that the use of medical terminology by care providers during medical encounters affects patient's understanding and hinders effective communication [7, 8]. Clear communication and ensuring recall and comprehension are essential in all patient-provider interactions, but may be especially important for patients with chronic conditions, such as diabetes, where self-management education and counseling is a cornerstone in managing these conditions. Further, these approaches could be more critical for patients who have challenges in obtaining, processing, understanding, and communicating about health –related information that are necessary to make informed health care decisions and effectively manage their health conditions i.e. those with inadequate health literacy (HL) [9].

According to the International Adult Literacy and Skills Survey, 60% of adult Canadians do not have the necessary to manage their health effectively [10], and thus are more likely to have adverse health effects including poor self-care management, improper use of medications and healthcare services, and poor health outcomes [10-12]. Patient-provider communication has been suggested as a potential pathway through which HL might impact health outcomes, particularly in individuals with chronic diseases [13, 14].

Providing patient education is an important nursing role and a core competency of nursing practice, particularly in diabetes self-management education [15]. However, the majority of the available literature examining patient-provider communication and HL involves only physicians. Therefore, our objective was to examine whether nurses addressed all components of the interactive communication loop, particularly with respect to assessing recall and comprehension and avoiding jargon, while providing self-management education and counseling to individuals with type 2 diabetes. A secondary objective was to explore whether these aspects of nurses' communication are influenced by patient's HL level.

6.2. Methods

6.2.1. Setting

A primary care network (PCN) is a network of doctors and other health providers (nurses, dieticians, pharmacists, social workers, nutritionists, etc) working together to provide health services to a defined patient population; akin to what is referred to as the “medical home” in the USA [16]. At the time of the study, there were 40 PCNs operating in the province of Alberta, with approximately 2500 (75%) of family physicians working in PCNs. This study was conducted at three PCNs; each serving a different target population with respect to socio-demographic characteristics, and ultimately health needs. At the time of the study, site 1 had 15 primary care nurses and 102 physicians, serving approximately 250,000 patients, site 2 had two primary care nurses and 43 physicians, serving approximately 52,000 patients, and site 3 had 3 nurses and 150 physicians, serving 167,00 patients.

6.2.2. Participant Recruitment

This study involved two stages of recruitment: first nurses and then patients. Nurses were approached at their monthly PCN meeting, and following an explanation of the study purpose and procedures, were provided with invitation letters. All nurses working at the PCNs involved in chronic disease management were eligible. Patients were approached by research assistants (RAs) who explained the purpose and procedures of the study. Patients 18 years or older, who had type 2 diabetes, were able to communicate in English, and did not have any severe mental or physical illness were eligible. Patients who agreed to participate provided written informed consent, and nurses provided written consent for each recorded encounter. Ethical approval for this study was obtained from the Health Research Ethics Board at the University of Alberta (PRO00029211).

6.2.3. Measures and Data Collection

After consent was obtained, data collection involved audio-recording patient encounters followed by administration of a patient survey to obtain demographic information (age, sex, marital status, educational level, first language, ethnicity, and income), diabetes related information (diabetes duration, medication regimen, comorbidities), health behaviors (smoking, alcohol consumption), self-care behaviors (Summary of Diabetes Self Care Activities [17]), and self-efficacy (Stanford Self-Efficacy for Managing Chronic Disease 6-Item Scale [18]). Visit length (in minutes) and type (new or follow-up) were recorded.

HL was assessed after the patient encounter (and thus, nurses were effectively masked to HL status) using a previously validated set of three brief screening questions [19]. The questions were: 1) How often do you have problems learning about your medical condition because of difficulty understanding written information?, 2) How confident are you filling out medical forms by yourself?, and 3) How often do you have someone like a family member, friend, hospital or clinic worker or caregiver help you read health materials (such as written information about your health or care you are offered)? These questions were scored using a weighted summative score of the three HL items, which we stratified such that scores < 9.0 were identified as adequate HL and scores >= 9.0 as inadequate HL [20].

Clinical and anthropometric measures including A1c, low-density lipoprotein (LDL), systolic blood pressure (SBP), and body mass index (BMI) were obtained from the patient's medical record using values within the last 3 months of the recorded patient-nurse encounter. At least two encounters per nurse were recorded, and data collection continued until data saturation was reached.

6.2.4. Coding and Analysis

Two professional transcriptionists, who were masked to the aim of the study and to the identity of participants, transcribed the recorded encounters. A coding manual (available upon request) was adapted from a previous study addressing physician-patient communication . Coding was done independently by two authors (FAS, JP). Since the focus of this study was the interactive [1] communication loop and use of jargon, only scripts of interactions pertaining to these components were coded. Additionally, as we chose to focus on diabetes self-management education, we only coded the interactions where any topic related to diabetes self-management activities (diet, exercise, self-monitoring of blood glucose (SMBG), foot care) and other related topics (such as hypoglycemia management and weight management) were discussed. Coding started while data was still being collected, and data saturation was determined as the point in which new information produced little or no change to the codebook and to the emerging themes [21].

A "topic" was coded as "sufficient" (and thus was included in the total count of topics per encounter) if the nurse provided health information (e.g. information about the nature of diabetes, or complications), taught a self-management skill (e.g. diet, exercise, foot care), or discussed a change in self-management (e.g., monitoring blood glucose, changing medication timing or dosing). Other topics were often briefly discussed during the encounters. For instance, the nurse might briefly ask the patient about their diet plan without providing any education or counseling. A topic was coded as "brief" (and thus not included in the total count of topics per encounter) if the nurse provided very brief information about a diabetes self-management activity, answered a patient's question by providing brief information, or only collected information about a self-management activity as a form of follow-up that did not involve any teaching or counseling. Since a complete communication loop is unlikely in these brief topics where no education or counseling was provided, we first evaluated and coded the topics as briefly or sufficiently discussed, and only coded “sufficiently discussed” topics.

Coding topics within each interaction was based on the five key components of the communication loop [1]: 1) repeating health information and instructions, 2) clarifying health information and instructions, 3) asking for understanding, 4) checking for understanding, and 5) seeking patient’s perceptions; as well as the two categories of jargon: 1) medical jargon, and 2) mismatched language. Jargon that was immediately clarified by the nurse was not coded as jargon. A detailed description and examples of each of the communication loop components and jargon are provided in Table 6-1.

After coding the interactions, the frequency [i.e. 0 (never), 1 (seldom), 2 (sometimes) or 3 (often)] of applying each of the communication loop components and use of jargon was rated. Rating was done per topic within each encounter, and then an average of the total ratings of each component was computed. Rating was done by the two coders, and inter-rater agreement was assessed using percentage agreement and Kappa. Regular meetings were held to compare the codes and ratings and ensure consistency in coding and interpretation.

Descriptive and comparative statistics were reported using appropriate statistical tests. We also examined the application of the communication loop components and use of jargon between HL groups. To simplify this comparison and the interpretation, the frequency of each component was collapsed into two categories, such that “never and seldom” were combined into one category “rarely”, and “sometimes and often” into another category “frequently”.

6.3. Results

6.3.1. General Characteristics of Participants

Nine nurses (representing 45% of the nurses working at the three PCNs) and 36 patients participated in this study. Two to four encounters were recorded per nurse, the average encounter length was 46 minutes (SD=19), and 18 encounters (56%) were first visits with the nurse. Nurses were all females with an average age of 50 years (SD=9.3), an average of 21.5 years (SD= 14.0) of various nursing experience, and 4.2 years (SD=1.8) of experience at a PCN. Six out of 9 nurses indicated that they received minimal training on health communication mainly in the form of lectures or seminars.

The average age of patients was 58.5 years (SD=14.1), more than half were female (58%), predominantly white (78%), the majority had at least high school education (86%), and English as a first language (78%) (Table 6-2). About two thirds of participants had high self-efficacy, and the mean scores of self-care behaviors were relatively good, except for exercise (mean= 2.7 days/week; SD=2.0). The average diabetes duration was 7.9 years (SD=7.1), while 11% were newly diagnosed, and most participants had at least one chronic condition in addition to diabetes, with hypertension being the most common comorbidity (58%). Overall, the participants had adequate metabolic control and the majority (75%) were obese (BMI >= 30).

The mean weighted summative score of the three brief screening questions of inadequate HL was 7.0 (SD 3.3; median 6.3; interquartile range 3.9 – 9.9). Thirteen of 36 (36%) participants had inadequate HL (i.e., weighted summative score >= 9.0); these patients were more likely to not have English as their first language, and had lower self-efficacy than participants with adequate HL (Table 6-2). Further, participants with inadequate HL were less likely to be white, had lower educational level, and had a greater burden of comorbidity compared to those with adequate HL; however, these difference were not statistically significant (Table 6-2).

6.3.2. Number and Types of Topics Discussed

Inter-rater agreement was moderate for all codes (ranging from 63% to 75% agreement between two raters, with kappa ranging from 0.59 to 0.71). The total number of topics discussed in the 36 encounters was 90, with an average of approximately 3 topics per encounter (SD=1.1; range: 1 – 5). The most commonly discussed diabetes-self-management topic was SMBG (30 out of 36 encounters), followed by diet (23 encounters), medications and insulin (8 encounters), and physical activity and foot care (4 encounters each). Other diabetes self-management related topics discussed included diabetes-specific blood tests such as A1c (4 encounters), blood pressure control (3 encounters), mental health (3 encounters), hypoglycemia management (2 encounters), and smoking cessation and vaccinations (1 encounter each).

6.3.3. The Communication Loop

Nurses used a complete communication loop in only four of the 36 encounters (11%); three with adequate HL patients and one with an inadequate HL patient. Of the five key components of the communication loop, clarifying health information/instructions was the most commonly used by nurses across all encounters (58% often used), followed by repeating health information/instructions (33% often used) (Figure 6-1). Examples of clarification and repetition are presented in Text Box 6-1. On the other hand, checking for understanding was the least applied across all encounters (81% never used), followed by asking for understanding (42% never used) (Figure 6-1). Examples of checking and asking for understanding are presented in Text Box 6-2. Seeking patient’s perceptions was not evident in the majority (58%) of the encounters, and only sometimes used in 11% of these encounters (Figure 6-1). Examples of phrases used to seek out patient perception are presented in Text Box 6-2.

Application of the communication loop components varied little between the HL groups (Figure 6-2). Repeating health information and instructions was more frequently used with patients with adequate HL (57% of encounters) compared to those with inadequate HL (46%). The frequency of clarifying health information and instructions was similar between the HL groups. Asking for understanding was more frequent in interactions with individuals with inadequate HL (31%) compared to those with adequate HL patients (17%). Checking for understanding and seeking patient’s perception were the least used in both groups, being rarely used in all interactions with individuals with inadequate HL, and only frequently used in 4% and 17% respectively of interactions with adequate HL patients.

6.3.4. Medical Jargon and Mismatched language

Although nurses did not use medical jargon in 31% of interactions, they still did use it sometimes in 19%, and often in 17%, of the encounters (Figure 6-1). Nurses used mismatched language in slightly more than half of the interactions; sometimes in 33% of the interactions and often in 25% of the interactions (Figure 6-1). Examples of the most commonly used medical jargon and mismatched language are presented in Text Box 6-3. The use of medical jargon and mismatched language was less frequent in interactions with inadequate HL patients compared to those with adequate HL (Figure 6-2).

6.4. Discussion

We found that primary care nurses rarely used a complete communication loop while providing self-management education and counseling to patients with type 2 diabetes, particularly with deficits observed in checking understanding. Nurses frequently used medical jargon and mismatched language in their communication with these patients. Health literacy did not seem to materially influence patterns of nurses' communication in this study, although nurses did tend to use less jargon and mismatched words with patients who had inadequate HL.

There is very limited literature on nurse-patient communication within the primary care setting, with little attention to the application of the communication loop in encounters for diabetes self-management education. Nonetheless, our finding of limited application of the communication loop in this study is consistent with a similar study involving family physicians. In that study, where direct observation was used to measure the extent to which primary care physicians assessed recall and comprehension of new concepts during encounters with type 2 diabetes patients with inadequate HL, recall and comprehension of adherence-related health instructions was only assessed in 20% of patient encounters [1].

The finding that nurses' use of medical jargon and mismatched language was frequent in patient encounters is also consistent with prior research that involved physicians. In a study that involved audio-taping physician-patient interactions, it was reported that physicians used unclarified jargon at least once in 81% of all recorded visits, with a mean of four per visit [22]. Other studies have also indicated that nurses use medical terminology in their interactions with patients within the hospital setting [23]; however, to our knowledge, there are no studies that address nurses' use of medical jargon in interactions with patients in the primary care setting.

The underuse of the communication loop components by the nurses might occur for several reasons. First, nurses might not have discussed patient education and counseling using the communication loop components in their education programs. Second, clinicians, including nurses, are usually rushed and therefore make patients feel rushed and reluctant to ask questions. Patients often prefer to be quiet than to admit that they do not understand their care providers' instructions [14, 24], which ultimately influences how healthcare providers communicate with them. Third, nurses might not be prepared to identify patients with learning challenges or difficulties, including inadequate HL [25], and tailor their communication approaches to address patient's learning needs.

Several factors should be considered in the interpretation of our findings. First, the practices of the nurses who participated in this study might not be representative of all nurses' practices across all PCNs, which might be due to a potential self-selection bias in which nurses who believed they provided adequate education and counseling volunteered to participate in this study. Second, this study involved only patients with type 2 diabetes, and thus the generalizability of the findings is limited to the diabetes population, and perhaps to other chronic disease populations with similar self-care management education. Third, the unit of analysis in this study was the interaction (N=36), and thus, due to low statistical power, we could not examine whether a complete communication loop or the frequency of applying any of its components is associated with behavioral or cardio-metabolic outcomes. Last, we purposefully sampled three PCNs in one Canadian province, and some might be concerned about generalizability to other primary care environments in Canada and elsewhere.

6.5. Conclusions and Implications

In summary, we found that nurses caring for individuals with type 2 diabetes within primary care settings rarely used a full communication loop in providing diabetes self-management education and counseling, particularly with respect to assessing recall and comprehension. Overlooking of this important feature in the communication process with patients reflects a missed opportunity to enhance patient education and counseling. We also found that nurses frequently used medical jargon and mismatched language in patient encounters. This overuse of medical jargon and mismatched language, accompanied with underuse of communication loop components jeopardizes patients' comprehension and retention of what they need to know to properly self manage their diabetes. Last, HL did not considerably affect patterns of communication by nurses in terms of using communication loops, however nurses used less jargon and mismatched words with patients with inadequate HL. The reasons about why nurses do not frequently apply certain communication strategies to enhance patient learning and understanding warrants further exploration.

These findings have a few implications. Nurses and other healthcare professionals should explicitly elicit patient perceptions about their illness and healthcare needs, and should apply communication strategies such as the teach back technique to enhance patient understanding and recall of health information. Additionally, given the high rates of use of diabetes-related medical jargon and mismatched language in nurse-patient interactions, nurses need to develop more effective ways to communicate concepts critical to chronic disease self-care education and management. Such communication strategies should be incorporated into professional development strategies and continuous education of nurses.

Tables, Text Boxes and Figures

Table 6-1: Description and examples of the communication loop components and types of jargon

Component	Description	Examples
Repeating health information and instructions	Repetitions are word for word repeats, or different words with the same meaning. Repetitions provide no new information. A unit of repetition can be as small as one word but as large as an idea spanning a few sentences.	"Last time we increased the metformin: the dosing of your diabetes medication"
Clarifying health information and instructions	Clarifications require a higher threshold; they represent information that illustrates or gives meaning to what was just said. Clarifications answer how or why the statement just made is true, or explain the meaning of what was just stated. A unit of clarification is usually a sentence or sentence fragment.	"Exercise is good for your heart: makes it pump, makes it work, makes it strong"; "the last time we got your A1c, it was 7.0 which is very good"
Asking for understanding	This is a direct way of assessing if the patient understands what has been discussed or explained. This is usually used after explaining health information or concepts.	"Is this clear now?"; "do you think I gave you enough information about ..?"; "Does this make sense?"
Checking for understanding	This is an indirect way of assessing if the patient understands what has been discussed or explained. This is usually used after explaining certain tasks such as diet management, exercise, self-testing of blood glucose levels. The "show me" technique is used when the patient is asked to demonstrate the task that was explained.	The "teach back" technique is when the patient is asked to teach back the nurse (or care provider) what s/he has been taught. The "show me" technique is usually used when teaching the patient how to perform a certain skill such as using a blood glucose meter to test blood sugar level, to do insulin self-injection, or to assess feet for ulcers or sores. The "teach back" technique is usually used when teaching the patients other concepts or health information such as diet management, exercise, or medication use.
Seeking patient's perceptions	This allows the nurse and patient to arrive at a common understanding about the nature of the illness and treatment plan. Moreover, it helps the nurse assess how to tailor information for the patient	"Do you know the symptoms of low blood sugar? Tell me." "Do you know why exercise is good for you?; "Tell me what the problem with drinking soda is"; "why do you think eating lots of bread is not good for you?"
Medical Jargon	Encompasses the terms used only in reference to medical concepts and	Hemoglobin A1c, biopsy or prognosis

	settings (i.e. medical terminology)	
Mismatched language	Consists of words common to everyday language but used in the medical environment with different or specifically modified meanings	Blood count, negative results, or stool card.

Table 6-2: General characteristics of patients

Characteristic	Overall (N=36) Mean or N (SD or %)	Inadequate HL (N = 13) Mean or N (SD or %)	Adequate HL (N = 23) Mean or N (SD or %)	P-value
Age	58.5 (14.1)	61.3 (16.9)	57.0 (12.4)	0.479
Sex - Female	21 (58.3)	10 (76.9)	11 (47.8)	0.087
Ethnicity - White	28 (77.8)	8 (61.5)	20 (87.0)	0.051
Educational level				
Less than high school	5 (13.9)	4 (30.8)	1 (4.4)	
Completed high school	15 (41.7)	3 (23.1)	12 (52.2)	0.063
More than high school	16 (44.4)	6 (46.2)	10 (43.5)	
Employment status - unemployed	16 (44.4)	8 (61.5)	8 (34.8)	0.115
Income				
< $40,000	5 (13.9)	4 (30.8)	1 (4.5)	
$40,000 - $80,000	12 (33.3)	4 (30.8)	8 (34.8)	0.081
>= $80,000	19 (52.8)	5 (38.4)	14 (60.9)	
English - 1st language	28 (77.8)	7 (53.9)	21 (91.3)	**0.016**
Self-efficacy				
Low	1 (2.8)	1 (7.6)	0 (0.0)	
Medium	10 (27.8)	6 (46.2)	4 (17.4)	**0.034**
High	25 (69.4)	6 (46.2)	19 (82.6)	
Self-care activities				
Diet	4.5 (2.1)	4.4 (2.2)	4.6 (2.1)	0.790
Exercise	2.7 (2.0)	2.3 (1.9)	3.0 (2.1)	0.345
SMBG	5.6 (2.1)	5.5 (2.4)	5.7 (2.0)	0.727
Medications	6.8 (0.9)	6.6 (1.4)	6.9 (0.4)	0.648
Diabetes duration (years)	6.8 (7.0)	7.2 (2.0)	6.6 (1.4)	0.947
Medical History				
Cardiac problems	5 (13.9)	3 (23.1)	2 (8.7)	0.576
Hypertension	21 (58.3)	8 (61.5)	13 (56.5)	0.526
Hyperlipidemia	17 (47.2)	6 (46.2)	11 (47.8)	0.599
Renal problems	3 (8.3)	1 (7.7)	2 (8.7)	0.709
Respiratory problems	8 (22.2)	4 (30.8)	4 (17.4)	0.638
Joint problems	9 (25.0)	6 (46.2)	3 (13.0)	**0.046**
Thyroid problems	8 (22.2)	5 (38.5)	3 (13.0)	0.070
Vision problems	8 (22.2)	4 (30.8)	4 (17.4)	
Mental or psychological illness	6 (16.7)	3 (23.1)	3 (13.0)	
A1c	8.1 (1.8)	7.9 (0.5)	8.1 (0.4)	0.586
LDL	2.3 (0.7)	2.5 (0.2)	2.2 (0.1)	0.128
SBP	135.0 (11.2)	132.2 (3.2)	136.6 (2.3)	0.454
BMI - Obese	27.0 (75.0%)	17 (73.9)	10 (76.9)	0.619

SMBG: Self Monitoring of Blood Glucose; SBP: Systolic Blood Pressure; BMI: Body Mass Index

Text Box 6-1: Examples of repetition and clarification used in the encounters

Examples of *repetition*:

"… pre- and post-mean, so you take it before you sit down and have lunch, and then 2 hours after you eat…"
"… the fiber of most fruits comes in the skin, so if you are eating an apple, when you eat the apple with the skin then you are getting more of the fiber …"
"… the dietitian would tell you that it is very important to eat breakfast; it is the most important meal …"

Examples of *clarification*:

"… we know your blood sugar drops below 4 mmol, , then you might feel a little weak, shaky, maybe a little disoriented …"
"… they recommend that we do ½ an hour, so 30 minutes of exercise 5 times a week. So something that would get your heart rate.. like a brisk walk …"
"… whenever you are choosing a product, try and choose something that is high in fiber, because fiber slows the rise of blood sugar…"

Text Box 6-2: Examples of phrases used in the encounters to ask or check for understanding and to seek patient's perceptions

Examples of phrases used to ask for understanding included:

"… does that make sense?..."
"… do you more questions about this?..."
"… do you have any questions at this point, with the change in insulin? …"
"… does that make it a little clear about the A1c meaning? …"

Examples of phrases used to check for understanding included:

"… what have you learned from that? What are the things that you can eat? …"
" … so what is going on with the pancreas? …" [after explaining the role of pancreas in diabetes]
"… so are you now comfortable adjusting your insulin at home when you are getting the higher readings? Do you know which one to increase a little bit?..." [after explaining how to adjust insulin based on blood glucose readings]

Examples of phrases used to seek out patient perception included:

"… tell me what you know about diabetes? …"
"… have you heard of glycemic index? …Tell me what you know about it…?"
"… what you think are the good things about the diet that you follow? ..."
"… what does carbohydrates mean to you? ..."

Text Box 6-3: Examples of the most commonly used medical jargon and mismatched language in the encounters

Examples of medical jargon:

neuropathy, glycemic control, insulin producing organ, autoimmune attack, gestational, dilated eye exam, ribcage, triglycerides, vessel disease, coronary artery disease, insulin insensitivity, incidence, adjuvant therapy, bone marrow, gastric distress

Examples of mismatched language:

blood work/requisition, fiber scoreboard, food journal, fasting sugar, simple sugars, quick random, starvation mode, honeymoon period, reassessment timetable

Figure 6-1: Application of the communication loop and its components and use of jargon in all participants

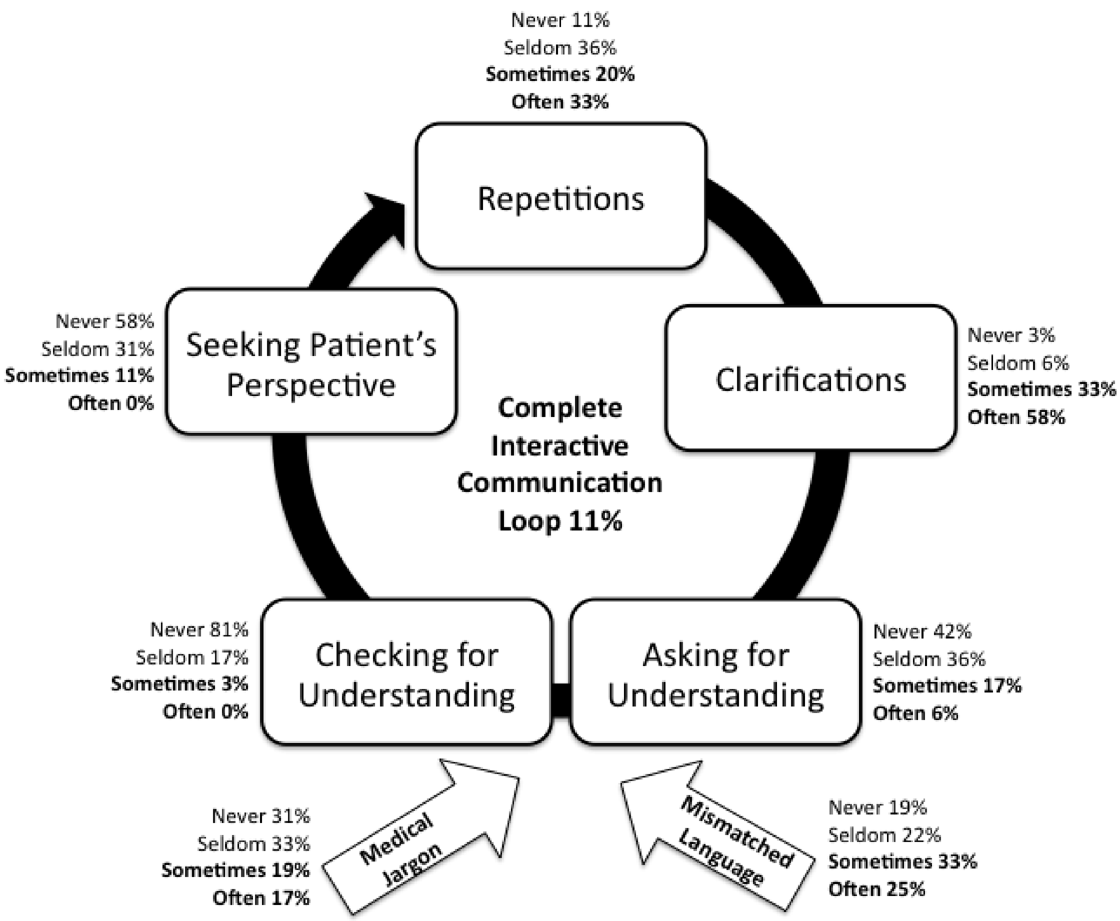
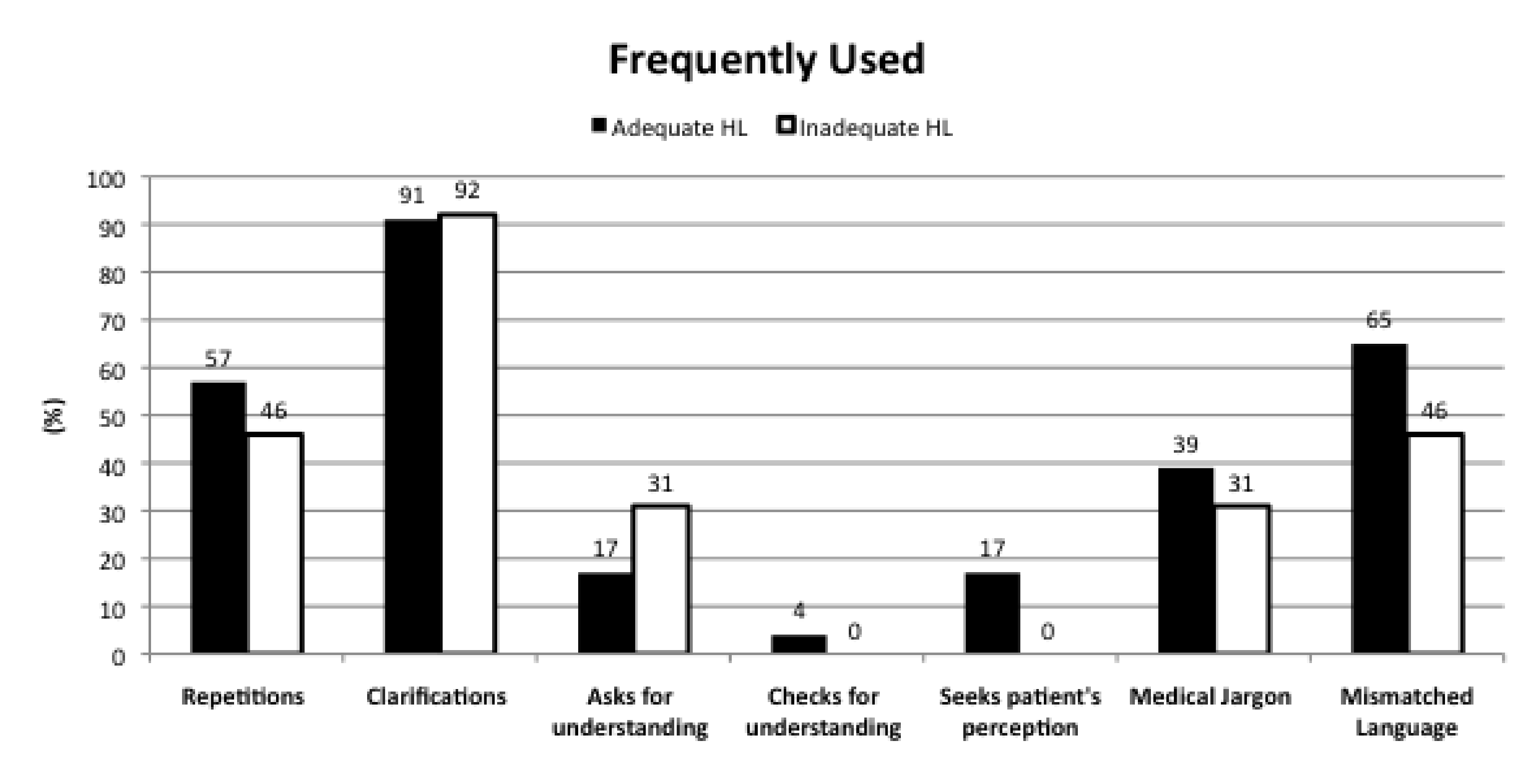


Figure 6-2: Application of the communication loop components and use of jargon by health literacy (HL)



References

1. Schillinger D, et al., *Closing the loop: Physician communication with diabetic patients who have low health literacy.* Archives of Internal Medicine 2003. **163**: p. 83-90.
2. Kripalani S and Weiss BD, *Teaching about health literacy and clear communication.* J Gen Intern Med, 2006. **21**: p. 888–90.
3. Williams MV, et al., *The role of health literacy in patient-physician communication.* Fam Med, 2002. **34**(5): p. 383-9.
4. Xu P, *Using teach-back for patient education and self-management.* American Nurse Today, 2012. **7**(3).
5. Schillinger D, et al., *Functional health literacy and the quality of physician-patient communication among diabetes patients.* patient Education and Counseling 2004. **52**: p. 315-323.
6. Weiss BD, *Health literacy: a manual for clinicians*. 2nd editon ed2007, Chicago: American Medical Association Foundation.
7. Hadlow J and Pitts M, *The understanding of common health terms by doctors, nurses and patients.* Soc Sci Med. , 1991. **32**(2): p. 193-6.
8. Lerner EB, et al., *Medical communication: do our patients understand?* Am J Emerg Med. , 2000. **18**(7): p. 764-6.
9. Berkman N, Terry D, and McCormack L, *Health literacy: What is it?* Journal of Health Communication, 2010. **15**: p. 9-19.
10. Learning, C.C.o., *Health Literacy in Canada: Initial results from the International Adult Literacy and Skills Survey*, 2007: Ottawa.
11. Berkman ND, et al., *Health literacy interventions and outcomes: an update of the literacy and health outcomes systematic review of the literature. Evidence Report/Technology Assessment no.199.* , 2011, Agency for Healcare Research and Quality: Rockville, MD.
12. DeWalt DA, et al., *Literacy and Health Outcomes: A Systematic Review of the Literature.* Journal of General Internal Medicine, 2004 **19**(12): p. 1228–1239.
13. Hironaka LK and Passche-Orlow MK, *The implications of health literacy on patient-provider communication.* Arch Dis Child, 2008. **93**: p. 428-432.
14. Al Sayah F and Williams B, *An Integrated Model of Health Literacy using Diabetes as an Exemplar* Canadian Journal of Diabetes 2012 **36**: p. 27-31.
15. Haas L, et al., *National standards for diabetes self-management education and support.* Diabetes Care, 2013 **36**(Suppl 1): p. S100-8.
16. Rittenhouse DR and Shortell SM, *The patient-centered medical home: will it stand the test of health reform?* JAMA, 2009. **301**(19): p. 2038-2040.

17. Toobert DJ, Hampton SE, and Glasgow RE, *The Summary of Diabetes Self-Care Activities Measure: results from 7 studies and a revised scale.* Diabetes Care, 2000. **23**: p. 943-950.

18. Lorig KR, et al., *Effect of a self-management program for patients with chronic disease.* Effective Clinical Practice, 2001. **4**: p. 256-262.

19. Chew LD, et al., *Validation of screening questions for limited health literacy in a large VA outpatient population.* Journal of General Internal Medicine 2008. **23**: p. 561-566.

20. Al Sayah F, et al., *Measurement Properties and Comparative Performance of Health Literacy Screening Questions in a Predominantly Low Income African American Population With Diabetes* BMC Health Services Research, Under review.

21. Kerr C, Nixon A, and Wild D, *Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research.* Expert Rev Pharmacoecon Outcomes Res, 2010. **10**(3): p. 269-81.

22. Castro C, et al., *Babel Babble: Physicians' use of unclarified medical jargon with patients.* American Journal of Health Behavior 2007. **31**(Suppl 1).

23. Bourhis RY, Roth S, and MacQueen G, *Communication in the hospital setting: A survey of medical and everyday language use amongst patients, nurses and doctors* Soc Sci Med, 1989. **28**(4): p. 339-46.

24. Parikh NS, et al., *Shame and health literacy: the unspoken connection.* Patient Educ Couns., 1996. **27**: p. 33–39.

25. Horowitz CR, et al., *Low health literacy is common and unhealthy: do we recognize it in our own patients?* J Gen Intern Med., 2004. **19**(suppl 1).

Chapter 7

Summary and Implications of Findings

7.1. Summary of Research

Our ability to find, understand, evaluate and communicate health-related information – our health literacy (HL) – is thought to be critical for maintaining a healthy lifestyle, managing our own or our family's health-care needs, making informed decisions about our health and navigating the health-care system. These skills would seem to be especially important for people with chronic disease, such as diabetes, due to the high demands of self-care and the complexity of the health environments that they continuously interact with. The relationship between HL and health outcomes in the diabetes population has been studied and evidence has been accumulating. However, there are several discrepancies and gaps in the literature, particularly with respect to measurement of HL, and the impact of inadequate HL on health outcomes. To address these gaps, several sequential studies were undertaken to enhance our knowledge and understanding in this area. This research is timely and important for both researchers and clinicians as the interest in HL research and interventions, particularly for chronic disease patients, is escalating in Canada and other countries.

The overall objective of this program of research was to enhance our understanding about the impact of inadequate HL on the health of individuals with type 2 diabetes (T2DM). After developing a theoretical model that conceptualizes the potential pathways through which HL might impact health outcomes in this population, we undertook several projects to examine different aspects of this model. First, and in order to attain a comprehensive understanding of the available evidence on the associations between HL and health outcomes, we conducted a systematic literature review in which we identified, appraised, and synthesized research evidence on the relationships between HL or numeracy and health outcomes (i.e., knowledge, behavioral and clinical) in people with diabetes (Chapter 2). We identified and included 34 eligible publications reporting data from 24 studies. We found that low HL was consistently associated with poorer diabetes knowledge. However, there was insufficient or inconsistent evidence suggesting that low HL is independently associated with processes or outcomes of diabetes-related care. Given the importance of the topic, we were surprised by the paucity of high quality evidence. The majority of the studies were from US primary care setting (87.5 %), and none were conducted in Canadian settings. This review indicates that the current understanding of the effect of low HL on health outcomes in people with diabetes is limited. Additionally, since the existing evidence is mainly based on cross-sectional explorations, it highlighted the need for more research to examine these relationships, and particularly the need for longitudinal studies.

Our next step was to investigate how HL has been measured, and identify the best available measure of HL to include in our studies. This was accomplished through another systematic review in which we identified instruments used to measure HL in people with diabetes, evaluated their use, measurement scope and psychometric properties, and their strengths and weaknesses, and identified the most useful, reliable, and applicable measure for use in research and practice settings (Chapter 3). We identified one diabetes-specific (LAD) and eight generic measures of HL (REALM, REALM-R, TOFHLA, s-TOFHLA, NVS, 3-brief SQ, 3-level HL Scale, SILS) and one diabetes-specific (DNT) and two generic measures of numeracy (SNS, WRAT).

We evaluated the measurement scope and use of these instruments based on Nutbeam's domains of HL [1] and a diabetes HL skill set. We found that the most commonly used instruments assess selective domains of HL, focus mainly on reading and writing skills, and do not address other important skills such as verbal communication, health care system navigation, health-related decision making, and numeracy. Additionally, the structure, mode, and length of administration and measurement properties were found to affect the applicability of these instruments in clinical and research settings. We also found that indirect self- or clinician-administered measures are the most useful in both clinical and research settings. One of these measures was the 3 screening questions of inadequate HL, which is a brief version of 16 screening questions [2, 3]. Although these questions, similar to other measures, assess only functional aspects of HL, they are the most congruent with the current definition and conceptualization of HL, and the most applicable in clinical and research settings. We therefore chose to use these questions as a measure of HL in our subsequent research projects. Furthermore, we found in this review that there is limited evidence on the measurement properties of almost all of the available HL instruments, including the 3-brief screening questions. This led to our next project, a psychometric validation study.

In this validation study (Chapter 4), we used cross-sectional data from a study of 378 predominantly African-American individuals with T2DM to examine the measurement properties of the 3 brief screening questions of inadequate HL. We computed sensitivity, specificity, positive and negative likelihood ratios, and

C-indices. We also conducted factor analysis to examine the measurement model of these questions, and used structural equation modeling (SEM) for confirmatory purposes. We used the s-TOFHLA as a reference measure. Six questions (6-SQ) were included in the final item-reduced factor analysis and produced two factors explaining 58% of the variance. In confirmatory SEM, this 2-factor model had a good fit (chi-square = 9.5; P = 0.305; RMSEA = 0.023). We also found that weighted summative score of the 6-SQ and the item "problems learning" performed better than the 3-brief screening questions in identifying patients with inadequate HL with C-indices of 0.67 and 0.75 respectively. Nonetheless, despite the fact that the 6-SQ performed better than the 3-SQ in those psychometric tests, the latter still had satisfactory sensitivity, specificity and AUROC. We therefore included the 3-SQ in three key studies that are part of the "Alberta's Caring for Diabetes" (ABCD) project, which also allowed us to ensure comparability of our findings to previous literature.

After examining the state of literature on HL measures and impact of HL on health outcomes in the diabetes population, we used data from one of the ABCD projects "The TeamCare-PCN" trial, and examined through a rigorous longitudinal research design the associations of inadequate HL with health outcomes including depressive symptoms, A1c, LDL, SBP, and health-related quality of life (Chapter 5). TeamCare-PCN was a controlled clinical trial that found no difference between a collaborative team care model and enhanced usual care for primary care T2DM patients who recently screened positive for depression. We combined the two study arms to provide a single cohort of patients, and evaluated changes in outcome measures over the 12-month study period, controlling for the intervention arm among other baseline characteristics. We found that among primary care patients with T2DM who recently screened positive for depression, inadequate HL was not associated with worse depression-related or other outcomes over 1-year. In contrast to many cross sectional studies, our findings suggest that inadequate HL in this population would be unlikely to have an impact on improvement of health outcomes. The discrepancy between our study and previous research might be related to methodological differences between the studies in terms of study design, HL measures used, or populations studied. It could also be related to the fact that HL is a distal factor in relation to these outcomes, which are dependent on a large number and variety of other factors, and this makes finding more than a weak association with these outcomes implausible. We also proposed the presumption that the impact of inadequate HL on health outcomes might vary by the severity of the health condition of populations studied; we believe this is an observation that warrants further investigation.

Last, given our view of HL being an interaction between individual's needs and health system demands, we felt it was important to assess the complexity of the system with which individuals with T2DM interact. One of the key components of any healthcare system is interaction with healthcare providers. Healthcare providers, particularly nurses, play a key role in providing education, counseling and support for patients with T2DM. For that, we examined different aspects of nurses' interaction with T2DM patients (Chapter 6). In specific, we explored the use of interactive communication loops and medical jargon in relation to HL, using coded transcriptions of 36 audio-recorded interactions of T2DM patients with 9 primary care nurses. We found that nurses rarely used a full communication loop and frequently used jargon in providing self-care education for T2DM patients; however the latter was less common in patients with inadequate HL. This study showed that the level of communication might be complex for all patients receiving self-care education within primary care setting; an important observation that needs to be further explored.

Collectively, the results of these studies suggest that 1) evidence on the impact of HL on health outcomes in the diabetes population is limited and inconclusive; 2) measures of HL are not comprehensive enough to capture the different components of this complex construct, and the measurement properties of existing measures need to be further explored particularly with respect to longitudinal validity; 3) the 3-brief screening questions are potentially a useful measure for screening for inadequate functional HL, however, their use needs to be further examined in different population groups and compared to different reference measures; 4) inadequate HL does not appear to have an impact on health outcomes in “sicker” populations such as those with T2DM who screened positive for depression; and 5) healthcare providers may place high demands on patients through their communication and interaction with them; this highlights the need to explore the level of complexity of different components of this system and the demands it places on patients, perhaps regardless of their HL.

7.2. Practice and Policy Implications

Given that HL research is still in its early stages, and with the inconsistent and inconclusive evidence hitherto generated, the findings of this dissertation have more implications for research than practice and/or policy. We found that the existing evidence and, indeed, the evidence that we generated does not support the premise that inadequate HL negatively affects health outcomes in the diabetes population. These findings have potential explanations that warrant further exploration.

Although there are a number of studies that have suggested that addressing HL needs of individuals with diabetes leads to better outcomes [4-6], there isn't sufficient evidence to support that improving HL would improve these outcomes, or even whether or not an individual's HL skills could be improved. It is crucial to emphasize that addressing the needs of individuals with inadequate HL skills may be achieved through decreasing the demands and complexity of the healthcare in terms of making systems easier to navigate, health information easier to access, read and understand, health communication clearer and simpler, and facilitating patient engagement and ensuring an environment conducive to learning. These strategies are essential especially given the lack of evidence on interventions or strategies to improve individual HL skills. Nonetheless, until better evidence is available, we believe that, outside of the study setting, it might be premature to routinely screen for HL or to try to improve HL for the purposes of improving patient-related outcomes in diabetes—although there may be other reasons to do so.

Given the emerging trends of self-management and patient-centered medicine, the increasingly complex health care, scientific advancements, and new technologies that have intensified the reading, writing, numeracy and problem-solving skill demands on health service consumers [7], the health care system needs to do a better job of managing all patients, perhaps regardless of the level of HL. Until, or unless, stronger evidence is generated, perhaps through better measures, it would appear best to recommend that healthcare professionals follow and implement “universal precautions” with all patients in all healthcare facilities [8, 9].

Universal precautions means that specific actions should be taken to minimize the risk for everyone when it is unclear which patients may be affected with the problem [10]. Healthcare providers should apply communication strategies to enhance patient understanding and recall of health information [11-14]. These communication strategies include using plain language and avoiding medical jargon, limiting the number of items discussed per encounter, repeating information, eliciting patients’ perceptions to identify learning gaps, using multiple forms of communication (e.g., oral, written, visual), confirming comprehension using techniques such as “teach back” and “show me”, and providing an environment conductive to learning and asking questions.

From a broader perspective, HL may be as much a public health issue as it is the challenge of individual patients or the health care system [15, 16]. The health sector may be able to substantially influence health knowledge, or it could be modified to be less complex and service those with less adequate HL skills; however, it is less able to influence HL. An individual’s personal, cognitive and social skills play a crucial role in HL but are subject to influences well outside of the control of health professionals and the health care system. Therefore, pursuing the goals of improved HL in the population will also require more overt alliances between health, education, and many other sectors as well as individuals, community-based groups and government agencies. Improving HL in a population involves more than the transmission of health information, although that remains a fundamental task. Helping people to develop confidence to act on that knowledge and the ability to work with and support others will best be achieved through more personal forms of communication, and through community-based educational outreach. Furthermore, given that low levels of HL stem in part from Canada's aging population, shrinking youth cohort and growing immigrant population [17], HL is more likely to improve if it is dealt with at the policy level. Although some observers suggest that advancing HL may become a global challenge for the 21st century [18], several research questions need to be answered before any initiatives or resources are deployed to screen for and enhance HL.

7.3. Research Implications

Overall, while there is an increasingly large body of literature on the topic of HL, with numerous presumptions on the importance and role of this concept as a risk factor for negative health outcomes, we identified substantial limitations in the current research literature surrounding the measurement and relevance of HL in improving health outcomes in patients with chronic diseases, or population health, for that matter. Our findings suggest that it might be premature to embark on randomized trials or controlled interventions to improve HL in those with diabetes, and perhaps other chronic diseases, given how little we actually know. Variations in research questions, study foci, methodologies, population groups, age, interventions and other indicators have contributed to disparate research findings. Lack of consensus about definitions and measurement of HL is one of the main contributors to the discrepancies that we observed, and makes it difficult to answer the question "Does HL matter?"

Nonetheless, there is a considerable research agenda that should be attended to first. For example, there is a pressing need to address some basic questions about HL (e.g., what is it and how can we best measure it?) before we can address other complex questions (e.g., what difference does it make to health outcomes? Is HL a modifiable factor? If so, how can we improve it among various populations? Is it cost-effective to do so? How do pay-offs from investing in HL compare with other investments we might make to improve people's health?).

While the limitations of various HL assessment tools has been compared and contrasted, collectively these tools may be better markers of health-related knowledge and health outcomes than is self-reported educational attainment. However, none of the currently available measures could be considered a gold standard, and current HL assessment tools will need to be redesigned or new ones developed to reflect revisions in HL definitions and accompanying conceptual models. Of particular importance, the level of complexity of healthcare systems and the demands of that system on individuals need to be incorporated into the measurement of HL, in order to reflect the most comprehensive conceptualization of this concept.

In addition to the limited measurement scope of existing HL measures, a major gap in the research on the measurement properties of these instruments is their ability to detect and measure change over time, that is, longitudinal validity (also known as responsiveness). The lack of information on this measurement property limits the utility of these instruments in studying whether HL is a fixed or dynamic concept. Viewing HL as static was primarily an artifact of its origins in prose literacy and related to limitations in existing measurement tools. We believe that conceptualizing HL as dynamic is inevitable. This would then lead to the consideration of whether HL is modifiable, and if so, to what extent, to what benefit, and at what cost?

Furthermore, there is a need for translation and cultural adaptation of HL measures to be used with multicultural populations and across different countries, which would permit conducting national and international comparisons of HL prevalence and potential impact. Moreover, given the attention to health status and HL in immigrant populations in North America, valid measurement of HL in recent immigrant populations may be enhanced through adequately adapted and validated measures.

In conclusion, building a new comprehensive approach to the measurement of HL may well be the next significant and necessary task and priority for HL researchers. Several questions need to be examined and answered in different populations and patient groups including those with diabetes before ascertaining that HL does affect the health of people; these include and not limited to the following:

- Is HL a socio-medical determinant of health? Or is it merely an indicator of other social determinants of health, such as educational level, income, or social class?
- What is the minimum level of HL skills that individuals need to have to manage their health effectively? How could we define this threshold in relation to the existing complex healthcare environments?
- Is there a causal relationship between HL and health outcomes? If so, what are the causal pathways of how HL influences these outcomes? And do these effects vary by populations studied?

- Is HL associated with and/or possibly influences health disparities? How does a person's HL influence access to needed medical care and uptake of preventive health measures?
- Would addressing HL needs of individuals by lowering the demands of the healthcare system in terms of complexity of system navigation, patient education and management without improving individual HL skills be sufficient to solve the issue of HL?
- How could we develop interventions and strategies to improve HL of individuals and that of populations?
- How could the public health sector and the education sector work together to enhance HL of populations?
- How could healthcare organizations, including healthcare providers, provide less complex health information and services that are more adaptable for everyone including those with inadequate HL?
- Could we estimate the costs of health care delivery/healthcare expenditures related to the direct and indirect impacts of inadequate HL? And if so, what are these costs?

7.4. Conclusions

We have addressed key gaps in the literature on the measurement of HL and its impact on health outcomes in individuals with diabetes. Overall, we found that the available tools do not comprehensively measure HL, and that inadequate HL does not appear to have an impact on health outcomes in this population. Despite the use of rigorous research methods and the robust evidence generated, the available overall evidence on these relationships is still inconsistent and thus inconclusive, and many questions need to be examined before making any recommendations to clinical practice or health policies. Our work highlights two crucial questions that need to be examined “how to comprehensively measure HL?” and “whether HL is modifiable?” Until, these questions – and others – are answered and conclusive evidence is available, we believe that, outside of the study setting, it might be premature to invest in routinely screening for HL or to trying to improve HL for the purposes of improving patient-related outcomes in diabetes.

References

1. Nutbeam, D., *The evolving concept of health literacy.* Soc Sci Med, 2008. **67**(12): p. 2072-8.

2. Chew LD, Bradley KA, and Boyko EJ, *Brief questions to identify patients with inadequate health literacy.* Family Medicine 2004. **36**: p. 588–594.

3. Chew LD, et al., *Validation of screening questions for limited health literacy in a large VA outpatient population.* Journal of General Internal Medicine 2008. **23**: p. 561-566.

4. Cavanaugh K, et al., *Addressing literacy and numeracy to improve diabetes care: Two randomized controlled trials.* Diabetes Care, 2009. **32**: p. 2149-2155.

5. Rothman R, et al., *Influence of patient literacy in the effectiveness of a primary care-based diabetes disease management program.* JAMA, 2004. **292**: p. 1711-1716.

6. Rothman R, et al., *The relationship between literacy and glycemic control in a diabetes disease-management program.* The Diabetes Educator, 2004. **30**: p. 263-273.

7. Nielsen-Bohlman L, Panzer AM, and Kindig DA, *Health Literacy: A Prescription to End Confusion*, 2004: Washington, DC.

8. US Department of Health and Human Services, O.o.D.P.a. and H. Promotion, *National action plan to improve health literacy*, 2010: Washington, DC.

9. Wilson-Stronks A, et al., *One size does not fit all: meeting the health care needs of diverse populations*, 2008: Oakbrook Terrace (IL).

10. Quality, A.f.H.R.a., *Health literacy universal precautions toolkit.* Vol. 10-0046-EF. 2010, Rockville (MD): Agency for Healthcare Research and Quality.

11. Al Sayah F and Williams B, *An Integrated Model of Health Literacy using Diabetes as an Exemplar* Canadian Journal of Diabetes 2012 **36**: p. 27-31.

12. Hironaka LK and Passche-Orlow MK, *The implications of health literacy on patient-provider communication.* Arch Dis Child, 2008. **93**: p. 428-432.

13. Ishikawa H and Takahiro K, *Health literacy and health communication.* BioPsychoSocial Medicine 2010. **4**(18).

14. Williams MV, et al., *The role of health literacy in patient-physician communication.* Fam Med, 2002. **34**(5): p. 383-9.

15. Nutbeam D, *Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century* Health Promotion International, 2000. **15**(3): p. 259-267.

16. Rudd RE, *Improving Americans' health literacy.* N Engl J Med, 2010. **363**(24): p. 2283-2285.

17. Rootman I and Gordon-El-Bihbety D, *A vision for health literate Canada: Report of the Expert Panel on Health Literacy*, 2008, Canadian Public Health Association Ottawa.

18. Nutbeam D and Kickbusch I, *Advancing health literacy: A global challenge for the 21st century.* Health Promotion International, 2000. **15**(3): p. 183–184.