

Research Article

Identifying At-Risk Subpopulations of Canadians with Limited Health Literacy

Takashi Yamashita,¹ A. John Bailer,^{2,3} and Douglas A. Noe²

¹ Department of Sociology, University of Nevada, Las Vegas, Las Vegas, NV 89154, USA

² Department of Statistics, Miami University, Oxford, OH 45056, USA

³ Scripps Gerontology Center, Miami University, Oxford, OH 45056, USA

Correspondence should be addressed to Takashi Yamashita; takashi.yamashita@unlv.edu

Received 9 October 2012; Revised 3 March 2013; Accepted 25 March 2013

Academic Editor: Wolfgang Ahrens

Copyright © 2013 Takashi Yamashita et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Health literacy, the set of skills for locating, understanding, and using health-related information, is associated with various health outcomes through health behaviors and health care service use. While health literacy has great potential for addressing health disparities stemming from the differing educational attainment in diverse populations, knowledge about subpopulations that share the same risk factors is useful. **Objective.** This study employed a logistic regression tree algorithm to identify subpopulations at risk of limited health literacy in Canadian adults. **Design.** The nationally representative data were derived from the International Adult Literacy and Skills Survey ($n = 20,059$). The logistic regression tree algorithm splits the samples into subgroups and fits logistic regressions. **Results.** Results showed that the subpopulation comprised of individuals 56 years and older, with household income less than \$50,000, no participation in adult education programs, and lack of reading activities (i.e., newspaper, books) was at the greatest risk (82%) of limited health literacy. Other identified subgroups were displayed in an easily interpreted tree diagram. **Conclusions.** Identified subpopulations organized in tree diagrams according to the risk of limited health literacy inform not only intervention programs targeting unique subpopulations but also future health literacy research.

1. Identifying At-Risk Subpopulations of Canadians with Limited Health Literacy

Health literacy (HL) describes the set of abilities to locate, understand, and use health-related information to make better health decisions throughout one's life [1]. Along with rapidly advancing medical/health sciences and changing health care systems, being able to obtain timely health-related information is critical for health maintenance/promotion [2]. Limited HL is prevalent in many adult populations and is directly and indirectly associated with lack of knowledge about health/diseases, miscommunication in healthcare settings, and adverse health behaviors (e.g., infrequent preventive healthcare service use and sedentary lifestyle) and, in turn, poorer health outcome [2–4]. Nearly half of the adult populations in developed nations like the US and Canada have insufficient HL skills [5, 6]. Improving HL is one strategy suggested for addressing health disparities [7].

The Rootman and Ronson HL conceptual model [8] depicts predictors of HL and theoretical pathways between HL and health. In this model, community-/societal-level (e.g., community development and public policy) and individual-level (e.g., aging and primary language) HL determinants are called "Actions" and "Determinants," respectively. HL may directly influence one's health, for instance, due to misunderstanding/misuse of medications [9]. At the same time, HL may influence health through various well-known health determinants including healthcare service use, health and disease management behaviors (e.g., smoking), economic resource access, and living environments [10–14]. Additionally, previous studies found that social factors such as access to quality education early in life and continuous opportunities for continuous literacy improvement (e.g., work-related literacy tasks) and health literacy proficiency are interrelated [15, 16]. In short, the Rootman and Ronson HL conceptual model suggests the complex associations

of demographic, socioeconomic and geographic factors on health and, in turn, emphasizes the diversity of limited HL risk factors as well as theoretical pathways to health.

The health disparities research model suggests that identification of populations at risk of adverse health should be the first step in public health research [17]. Generally, public health intervention programs can be categorized as either a homogeneous intervention program for a large target population (i.e., population approach) or as a tailored program for each individual at risk (i.e., individual-at-high-risk approach) [18, 19]. However, a review of recent public health interventions suggests that both approaches have not yet made significant impacts on health disparities [20, 21]. Arguably, lack of attention to the intrinsic variability in populations in the population approach and resource-/labor-intensiveness in the individual-at-high-risk approach partially explain some unsuccessful interventions. A reasonable compromise between the population and individual intervention targeting options would be to target at-risk subpopulations identified by a series of risk factors.

Guided by the Rootman and Ronson HL conceptual model (2005) and the health disparities research framework [17], this study identifies subpopulations at risk of limited HL in Canadian adults using the International Adult Literacy and Life Skills Survey data (described later). Specifically, the following research question is addressed: what characteristics define subpopulations at high risk of limited HL in the Canadian adult population? While a combination of bivariate analyses may identify some reasonable subpopulations, more insight is likely to be provided by simultaneously considering a large collection of potentially relevant predictors of limited HL.

2. Methods

2.1. Data. The data come from the International Adult Literacy and Skills Survey (IALSS). IALSS data include a nationally representative sample of Canadian adults ($n = 20,059$) and provide measures in literacy skills; the data also include information regarding demographic, socioeconomic, lifestyle, and health conditions [22]. The definition of each literacy domain is described elsewhere [23]. HL measure is derived from a subset of health-related prose literacy, document literacy, and numeracy test items. The score ranges from zero to 500. The respondents were categorized according to their HL scores into 5 levels (level 1 (very poor skills) = 0–225; level 2 (poor skills) = 226–275; level 3 (adequate skills) = 276–325; level 4 (Strong skills) = 326–375; level 5 (strongest skills) = 376–500) [6, 24]. The literacy measurement tool was jointly developed by the Educational Testing Service (ETS) in the US and Statistics Canada and is comparable to other major national-level adult literacy assessments such as the National Assessment of Adult Literacy (NAAL) and the International Adult Literacy Survey (IALS), assessment tools also developed by the same organization (ETS) [25].

2.2. Outcome Variable. The outcome variable is a dichotomous variable coded as [1] if individuals have the lowest category of HL, limited HL (*level 1*) defined in the IALSS, and

(0) if individuals have the basic, intermediate or proficient HL skills (*levels 2–5*). Assuming that limited HL translates into higher risk of adverse health outcomes we believe is an appropriate emphasis. In our preliminary analyses (results now shown) which are to compare the respondents with limited HL *level 1* to those with all other levels (each level separately), those with limited HL *level 1* had significantly lower scores in all widely used physical health indicators and most of mental health indicators except two emotional problems indicators (i.e., feeling depressed or anxious; did not do work or other activity as carefully as usual) than those with upper level HL skills (i.e., SF-12; 26). In this regard, focusing on *level 1* as the highest risk group is empirically supported by a variety of health indicators.

2.3. Predictor Variables. The model includes predictor variables that reflect “determinants” factors from the Rootman and Ronson model (2005). We focus on the “determinants” factors due to unavailability of “actions” or community/society-level factors in the IALSS. In this model, the determinants include “Education,” “Early child development,” “Aging,” “Personal capacity,” “Living/working condition,” “Gender,” and “Culture.” Although the variables from the IALSS may not always reflect each component, this study preserves the terminologies in the Rootman and Ronson model.

Education. *College education* is an indicator variable recording whether the respondents completed a college degree. *Adult education* records whether the respondents attended any postsecondary education or training including any educational/training program, course, private lessons, workshops, and so forth in last 12 months. The following three variables are used to reflect the respondents’ everyday literacy-related activities. *Reading Newspapers* records whether a respondent reads newspapers at least once a week. *Watching TV* records the hours per day a respondent watches TV, video, or DVD: categories (1) 1 hour or less, (2) over 1 hour to 2 hours, (3) more than 2 but less than 5 hours, and (4) 5 or more hours. *Reading Books* records whether a respondent reads books at least once a week.

Early Child Development. *English as first language learned* records whether the first language a respondent learned was English. *English as a primary language* records whether the primary language spoken at home was English when a respondent grew up.

Aging. *Age group* records the six age groups (age 16–25; 26–35; 36–45; 46–55; 56–65; 66 and older) according to the respondents’ age at the time of the survey in 2003. Age in years was not available for all IALSS respondents.

Personal Capacity. *Self-rated health* records whether a respondent reported his/her health is (5) excellent, (4) very good, (3) good, (2) fair, or (1) poor. While health status is not a common indicator of personal capacity including literacy skills, several studies found that health status indicators can serve as a surrogate measure of chronic conditions (e.g., diabetes

and disease-caused visual impairment) that are negatively associated with one's HL [26, 27]. *Emotional problem* records whether the respondent reported any emotional problems (e.g., depressed feeling and feeling anxious) during the past 4 weeks. *Life satisfaction* records whether a respondent reported his/her life satisfaction as (5) extremely satisfied, (4) satisfied, (3) neither satisfied nor unsatisfied, (2) unsatisfied, or (1) extremely unsatisfied.

Living/Working Condition. *Household Income* records the respondent's annual household income (converted to the US dollar unit) as (1) \$25,000 or less, (2) \$25,001 to \$50,000, (3) \$50,001 to 90,000, and (4) more than \$90,000. *Employment* records whether a respondent was working at least part-time at the time of the study. *Urban area* indicates whether a respondent lived in an urban area. *Region* records one of the five regions where the respondents reside and includes Atlantic, Quebec, Ontario, West, and North. Health insurance was not taken into account due to the universal health care coverage in Canada. Also, marital status information was not available in IALSS. *Student* records whether a respondent was a student (including work programs).

Gender. *Female* records if a respondent is female or not.

Culture. *Immigrants* records whether a respondent was an immigrant. *Born in Canada* records whether a respondent reported that they were born in Canada. *Immigrants* indicates whether a respondent was born outside of Canada and moved to Canada at some point of their lives. A value = 1 for *Immigrants* also identifies individuals who become Canadian citizens by naturalization.

2.4. Statistical Methods. This study employs a logistic regression tree approach, which has several advantages over commonly used traditional methods such as bivariate testing and logistic regression in epidemiologic research [28]. Compared to standard logistic regression, the logistic regression tree approach allows a more flexible model structure, including the incorporation of complex interactions, for describing how predictor variables impact the odds of limited HL. This approach implements a method that identifies subpopulations with similar HL and then allows for the fitting of a logistic regression model with potentially different predictor variable for each subpopulation. In addition, this method does not require imputation or case-wise deletion to deal with missing values.

2.5. Logistic Tree with Unbiased Selection—LOTUS. This study employs the logistic regression tree algorithm named Logistic Tree with Unbiased Selection (LOTUS) [29]. LOTUS is designed specifically to deal with binary responses (limited HL versus not limited HL in this study) and has advantages (e.g., unbiased split variable selection) over other classification and regression tree algorithms such as CART [30] and CHAID [31]. LOTUS can be considered a combination of classification trees and logistic regression. LOTUS merges the desirable capabilities of these two methods in such a way that logistic regressions are fit in recursively split subdata (i.e., subpopulations).

To determine the final size (i.e., number of subpopulations) of the logistic regression tree, LOTUS employs an overfit-and-prune approach. The recursive process of splitting nodes continues until no further splits are possible due to either homogeneity (e.g., a subgroup has 100% limited HL) or a lack of sufficient numbers of cases. Obviously, such a complete, unpruned tree provides an excellent fit to the provided data but may not generalize well to new observations. To prune back the overgrown tree in this study, 10-fold cross-validation using the 1-SE rule is used as described by Breiman et al. [30]. More technically, detailed descriptions of LOTUS are provided in Chan and Loh [29] and Loh [32]. The LOTUS algorithm version 2.3 was applied for the IALSS respondents to identify subgroups at high risk of limited HL.

3. Results

Table 1 includes descriptive summaries of candidate predictor variables for the individuals with and without limited HL. Approximately 29% of the respondents had limited HL, the lowest, Level 1 category of HL. As expected, a series of bivariate comparisons showed that higher education, learning/reading activities (e.g., adult education, reading newspapers), English as the first language, younger age, higher income, employment, and nonimmigrants and being born in Canada are more common in individuals with greater HL skills than in individuals with limited HL.

The results of the LOTUS analysis are summarized in a tabular form listing all identified subpopulations (Table 2) ordered from highest percentage to lowest percentage of limited HL. In addition, a graphical display of these subpopulations is presented in tree diagrams (Figure 1(a) (left half of the diagram) and Figure 1(b) (right half of the diagram)). The pathway to the highest risk groups is in bold in Figures 1(a) and 1(b). The respondents aged 56 and older with household income less than \$50,000 who read books less than once a week, read newspapers less than once a week, and report no adult education program participation last 12 months were at the highest risk (81.8%) of limited HL (see Figure 1(b)).

Among each subpopulation (i.e., terminal nodes or rectangular boxes in the tree diagram) identified according to the risk of limited HL, further within-subpopulation differences were detected by simple logistic regression models; 5 out of 15 subpopulations had household income as the predictor variable in the simple logistic regression models. Also, a noteworthy interaction effect between reading newspapers and having a college education on the odds of limited HL was detected in the subgroup of people aged 56 and older with household income less than \$50,000 who read books less than once a week and report no participation in adult education programs in the last 12 months. In this particular subpopulation, the respondents who read a newspaper more than once a week had 0.18 times odds of having limited HL when they had college education compared to those who did not. On the other hand, respondents who read a newspaper less than once a week had 0.29 times the odds of having limited HL when they had college education compared to those who did not. In other words, college education was

TABLE 1: Descriptive summary statistics, 95% confidence interval, and bivariate tests by the IALSS respondents with and without limited health literacy.

Variables	All respondents (<i>n</i> = 23,038)			Difference (95% CL: LL, UL.)	<i>X</i> ² (DF, <i>P</i> value)
	Total	Limited HL	Not limited HL		
Limited HL ^d	29.3%	—	—		
College education ^c	40.0%	18.6%	48.5%	30.0 (28.6, 31.2)	1780.2 (1, <i>P</i> < .0001)
Adult education ^c	45.1%	20.7%	55.1%	34.3 (33.0, 35.7)	2182.8 (1, <i>P</i> < .0001)
Read newspapers ^c	75.2%	65.9%	79.0%	13.1 (11.9, 14.3)	421.6 (1, <i>P</i> < .0001)
Read books ^d	41.7%	25.2%	48.4%	23.3 (21.9, 24.7)	1020.5 (1, <i>P</i> < .0001)
Watch TV				NA	908.9 (3, <i>P</i> < .0001)
1 hour or less per day	25.0%	18.8%	27.6%	—	—
Over 1 hour to 2 hours per day	33.0%	27.9%	35.1%	—	—
More than 2 hours but less than 5	33.5%	37.1%	32.1%	—	—
5 or more hours per day	8.5%	16.3%	5.3%	—	—
English was the first language learned ^c	50.2%	40.5%	54.1%	13.6 (12.2, 15.1)	341.6 (1, <i>P</i> < .0001)
English is the primary language ^c	62.0%	55.2%	64.7%	9.6 (8.1, 11.0)	163.9 (1, <i>P</i> < .0001)
Age group				NA	2256.8 (5, <i>P</i> < .001)
16–25	16.8%	10.9%	19.2%	—	—
26–35	15.5%	10.3%	17.7%	—	—
36–45	21.3%	16.5%	23.3%	—	—
46–55	20.2%	17.4%	21.3%	—	—
56–65	13.3%	17.1%	11.7%	—	—
66 and older	12.9%	27.9%	6.7%	—	—
Self-rated health				NA	1322.6 (4, <i>P</i> < .0001)
Poor	3.9%	7.8%	2.3%	—	—
Fair	11.5%	19.2%	8.4%	—	—
Good	29.0%	33.7%	27.1%	—	—
Very good	33.3%	25.0%	36.7%	—	—
Excellent	22.2%	14.3%	25.4%	—	—
Emotional problem ^c	8.9%	12.8%	7.3%	-5.6 (-6.4, -4.7)	175.1 (1, <i>P</i> < .0001)
Life satisfaction				NA	396.9 (4, <i>P</i> < .0001)
Extremely unsatisfied	1.8%	2.7%	1.4%	—	—
Unsatisfied	6.5%	9.9%	5.1%	—	—
Neither	9.6%	9.8%	9.5%	—	—
Satisfied	61.3%	62.7%	60.7%	—	—
Extremely satisfied	21.0%	15.0%	23.4%	—	—
Household income				NA	1739.6 (3, <i>P</i> < .0001)
\$25,000 or less (bottom quintile)	25.3%	43.4%	18.2%	—	—
\$25,000.01 to \$50,000	30.7%	33.7%	29.5%	—	—
\$50,000.01 to \$90,000	28.5%	16.6%	33.1%	—	—

TABLE 1: Continued.

Variables	All respondents (<i>n</i> = 23,038)			Difference (95% CL: LL, UL,)	<i>X</i> ² (DF, <i>P</i> value)
	Total	Limited HL	Not limited HL		
More than \$90,000.01 (top quintile)	15.6%	6.3%	19.2%	—	—
Employed ^c	61.9%	41.5%	70.4%	29.0 (27.6, 30.3)	1696.4 (1, <i>P</i> < .0001)
Student ^c	5.7%	3.2%	6.7%	3.6 (2.9, 4.2)	113.3 (1, <i>P</i> < .0001)
Urban areas ^c	74.4%	72.3%	75.3%	3.0 (1.7, 4.4)	18.5 (1, <i>P</i> < .0001)
Regions				NA	
Atlantic	20.3%	23.0%	19.2%	-3.8 (-5.0, -2.7)	43.0 (1, <i>P</i> < .0001)
Quebec	18.1%	14.5%	20.0%	5.1 (4.0, 6.1)	82.3 (1, <i>P</i> < .0001)
Ontario	21.5%	23.8%	20.5%	3.0 (1.7, 4.4)	31.0 (1, <i>P</i> < .0001)
West	28.9%	26.4%	30.0%	3.6 (2.3, 4.9)	30.2 (1, <i>P</i> < .0001)
North	11.2%	12.3%	10.8%	-1.5 (-2.4, -0.6)	11.2 (1, <i>P</i> < .0008)
Gender (female) ^c	54.5%	53.4%	55.0%	1.7 (0.3, 3.1)	5.2 (1, <i>P</i> < .022)
Immigrants ^c	18.2%	25.7%	15.1%	-10.6 (-11.7, -9.4)	315.5 (1, <i>P</i> < .0001)
Born in Canada ^c	71.8%	64.4%	74.9%	10.6 (9.3, 11.8)	263.1 (1, <i>P</i> < .0001)

^aData source: International Adult Literacy and Skills Survey (IALSS) 2003.

^bAll descriptive summary statistics are unweighted

^cDummy variables (1 versus 0).

^dLimited health literacy: health literacy level 1; nonlimited health literacy: health literacy level 2,3, and 4.

NA: not applicable; HL: health literacy; CL: confidence interval; UL: upper limit; LL: lower limit; DF: degrees of freedom; IALSS = International Adult Literacy and Life Skills Survey.

associated with lower chance of limited health literacy in both of these subpopulations although the effect of college education varied between individuals who read a newspaper more than once a week and those who did not.

4. Discussion

Building on the HL conceptual model [8] and the public health disparities research model [17], this study employs a logistic regression tree analysis and identifies the subpopulations at risk of having limited HL in the Canadian adult population aged 16 years and older. The proportions of limited HL in the identified subpopulations range from 12% to 82% (see Figures 1(a) and 1(b)). Approximately 82% of the subpopulation of Canadian adults aged 56 years and older with household income less than \$50,000 who read books and newspapers less than once a week and report no participation in adult education last 12 months exhibited limited HL, the highest percentage among all 15 subpopulations identified.

Combinations of demographic and other relevant factors clearly identified subpopulations at risk of limited HL. The

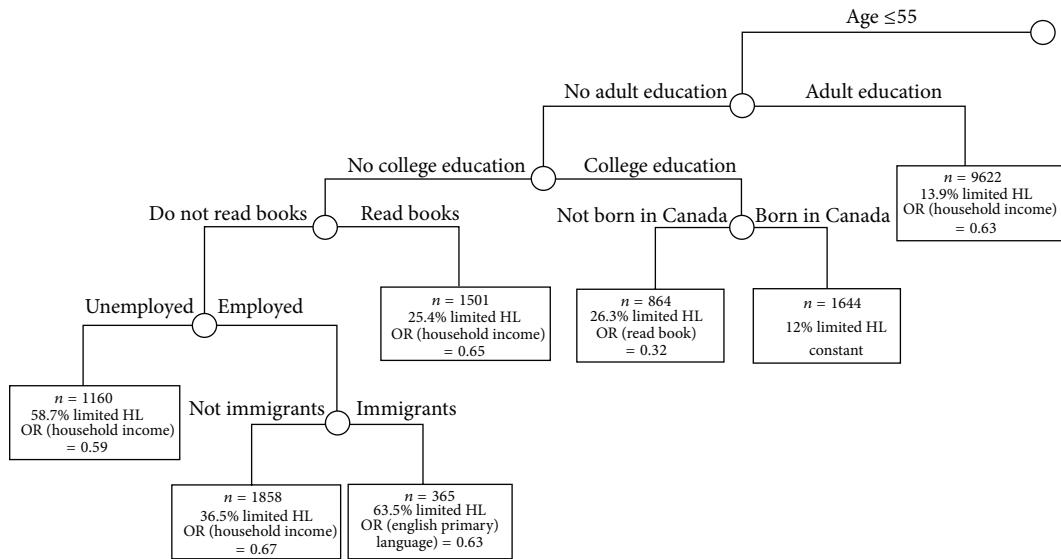
identified risk factors can be roughly classified into two groups: (1) the demographic/socioeconomic factors of age and household income; (2) everyday learning activities: reading books, participation in adult education programs, and reading newspapers. These risk factors are also reflected in the “Determinants” domain of Rootman and Ronson’s model. Although the observational data in this study do not allow causal inference, a lack of regular learning activities may further increase the risk of limited HL in particular subgroups (e.g., older age and low income) with individuals who are already at greater risk [12]. At the same time, it is also possible that some aging-related physical or cognitive declines and/or a lack of resources (e.g., time for reading and money to buy new books) might influence opportunities for learning activities and, in turn, resulted in limited HL [33, 34].

In addition, earlier educational attainment (lack of college education was the best predictor of limited HL in the highest risk subpopulation) is likely to determine one’s socioeconomic status (e.g., income and occupation) and put those with lower educational attainment are put at risk of limited HL [35–38]. Arguably, from a long-term perspective, lack of

TABLE 2: Identified subpopulations and predictor variables by the risks of limited health literacy.

Subpopulations	% Limited HL (n)	Age	Household income	Read books	Adult education	Read newspaper	College education	Employed	Immigrants	Born in Canada
1	81.8% (817)	≥ 56	$\leq \$50,000$	No	No	No	—	—	—	—
2	64.9% (1,775)	≥ 56	$\leq \$50,000$	No	No	Yes	—	—	—	—
3	63.5% (365)	≤ 55	—	No	No	—	Yes	Yes	Yes	Yes
4	58.7% (1,160)	≤ 55	—	No	No	—	No	No	No	No
5	52.3% (779)	≥ 66	$\leq \$50,000$	Yes	No	—	—	—	—	—
6	37.7% (245)	≥ 56	$\geq \$50,000$	—	—	No	—	—	—	—
7	36.5% (679)	≤ 55	—	No	No	Yes	—	—	—	—
8	35.1% (262)	≥ 56	$\leq \$50,000$	No	Yes	—	—	—	—	—
9	34.0% (586)	56–65	$\leq \$50,000$	Yes	No	—	—	—	—	—
10	33.5% (477)	≥ 56	$\leq \$50,000$	Yes	Yes	—	—	—	—	—
11	26.3% (864)	≤ 55	—	—	No	Yes	—	—	—	—
12	25.4% (1,501)	≤ 55	—	—	No	Yes	—	—	—	—
13	16.6% (692)	≥ 56	$\geq \$50,000$	—	—	Yes	—	—	—	—
14	13.9% (9,622)	≤ 55	—	—	Yes	—	—	—	—	—
15	12.0% (197)	≤ 55	—	—	No	Yes	—	—	—	Yes

Note: see Figures 1(a) and 1(b) for simple logistic regression results for terminal nodes (i.e., identified sub-populations).
 HL: health literacy.

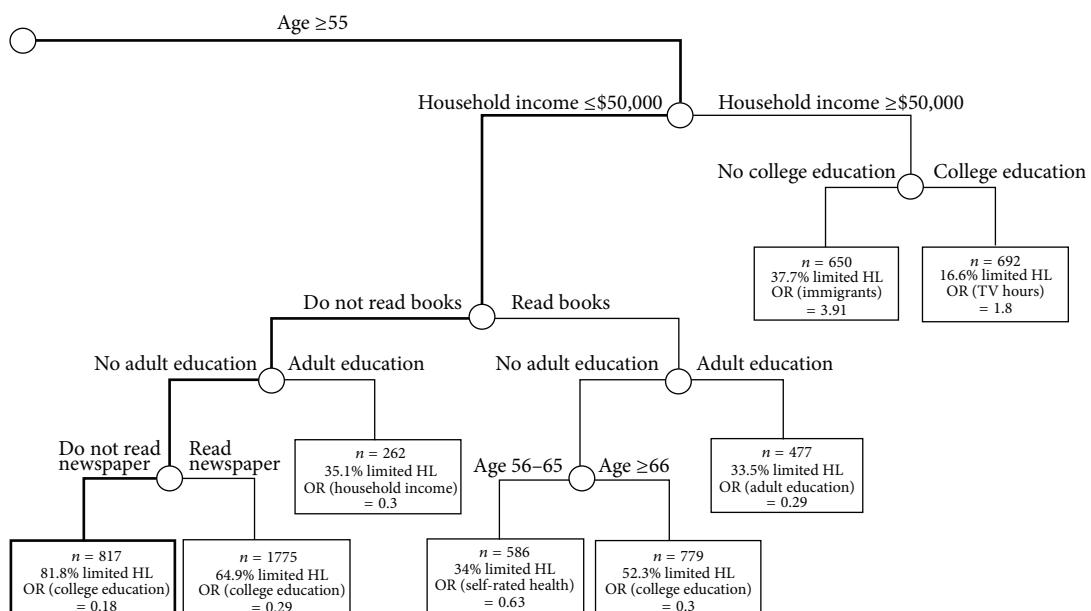


Notes:

Data source: international adult literacy and life skill survey (IALSS)

OR: estimated odds ratio for a predictor variable of simple logistic regression in each terminal node.

(a)



Notes:

Data source: international adult literacy and life skill survey (IALSS)

OR: estimated odds ratio for a predictor variable of simple logistic regression in each terminal node.

Bold lines lead to the sub-population at the highest risk of limited health literacy

(b)

FIGURE 1: (a) Logistic regression tree diagrams of limited health literacy: the entire International Adult Literacy and Life Skills Survey (IALSS) respondents (left half). (b) Logistic regression tree diagrams of limited health literacy: the entire International Adult Literacy and Life Skills Survey (IALSS) respondents (right half).

literacy skills (e.g., health literacy) could impact social status determinants like educational attainment and employment, and vice versa. Therefore, individuals with poor initial literacy skills may experience cumulative disadvantages over the life course [15, 39]. As suggested in Rootman and Ronson's model, formal education contributes to one's general literacy as well as other types of literacy skills (e.g., scientific literacy) and, in turn, HL skills. In this regard, life-long (i.e., mostly occupation-related formal programs) and life-wide (i.e., informal education programs) learning activities such as participation in adult education programs and reading newspapers/books might have acted as mediators of the association between limited HL and possible risk factors, particularly after participation in formal schooling ends [15]. Indeed, those who reported frequent reading/learning activities were less likely to have limited HL.

It should be noted that the national level efforts to address limited HL have been made by the governments and national organizations last few decades [40, 41]. In most cases, the health systems (e.g., healthcare providers and health organizations) provide customized materials and additional services to individuals with limited HL. Such "downstream" approach (i.e., assisting existing health illiterates) is a critical part of national health goals because even individuals with mid-level HL skills often face difficulties in healthcare settings [42]. Especially, unfamiliar health-related literacy tasks may be challenging. At the same time, efforts to prevent limited health literacy (i.e., "upstream" approach) are necessary for sustainable long-term HL interventions as more than half of Canadian adults may have difficulties with health-related literacy tasks [6]. Results of this study could inform not only interventions for existing limited HL prevalence but also for preventing future health illiterates. HL and/or literacy education interventions in earlier lives are key for successful prevention strategies considering the identified risk factors like older age and lack of learning activities. As one's health needs and necessary health knowledge change over the life course, identified risk factors in this study may help target potential health illiterates. Finally, a more health literate population may require less case management and direct HL education which could translate into health care cost savings.

4.1. Implications. The findings from this study have several implications for future research and practice to improve HL. First, this analysis identifies subpopulations with higher risk of limited HL; these findings can guide resource allocations for interventions from public health practitioners and health policy makers. Targeting the highest risk subgroups for interventions is an efficient and, hopefully, effective alternative. Second, the information about multiple risk factors and their relationships with each other in the tree diagrams provide useful insights for designing intervention programs. For instance, for individuals who are older and in lower income households, HL intervention programs need to be sensitive to their needs including familiar communication styles and logistics (e.g., transportation) [43–45]. For example, given the identified reading-related risk factors, multiple interventions such as offering a series of educational opportunities and/or making reading activities more accessible to the high risk

groups may improve one's literacy and ultimately health literacy [46]. Additionally, the risk factor of older age suggests that intervention programs/materials should be accessible (e.g., larger font and familiar examples) to mid-aged to older adults with limited health literacy. On a related note, careful examination of tree diagrams generates possible hypotheses for future inquiries. Arguably, each subpopulation has different processes that lead individuals to limited HL. Therefore, more detailed subpopulation analyses could improve understanding about etiological pathways to limited HL.

Third, logistic regression tree algorithms like LOTUS can be applied for various other populations such as community members, hospital patients, and individuals with diseases (e.g., diabetes and arthritis) and data routinely collected by political boundaries [28]. Finally, the identified subpopulations in order of the risk of limited HL (see Table 2) need to be systematically compared in future research. Such a comparative study may further advance practically relevant HL research. Also, the subpopulation strategy may be further tailored to reflect local demographic characteristics (e.g., race and ethnicity), culture, politics, and education systems. In short, while the findings in this study are not conclusive, they will guide current practice and research related to limited HL prevalence that is a serious public health concern by proposing a novel approach for identifying subpopulations.

4.2. Limitations. Possible omitted variable bias might exist due to the information availability in the IALSS data. Other types of education measurement, cognitive skills, more detailed community-level information (i.e., "Actions" domain in Rootman and Ronson's model) besides *regions*, health care utilization, and other health conditions are important to be included in future data collection. Also, although the IALSS is arguably one of only few nationally representative datasets that provides methodologically sound estimate of health literacy skill levels, the evaluation of specific level has not been established, and therefore, the results may need to be verified in future research. Moreover, the statistical method (LOTUS) employed in this study has some advantages over traditional methods, but LOTUS could still benefit from several features for further improvement. Such options might include more flexible options for split point selection of a continuous variable (instead of specific percentile points), the capability to incorporate survey sampling weights, the provision of additional model fit indices (e.g., c-statistic), and features to analyze longitudinal data. Finally, the logistic regression tree approach does not replace traditional methods but adds new analytic capabilities.

5. Conclusion

This study identified specific subpopulations that share the same characteristics and are at risk of limited HL. Specifically, older Canadian adults (56 years and older) earning less than \$50,000 household income who read infrequently and who do not participate in adult education programs are at the highest risk of having limited HL. As demonstrated in this study, the logistic regression tree algorithm (LOTUS) generates an easily interpreted tree diagram and provides insights

such as splitting points and/or potential interaction effects of key risk factors. In addition, possible hypotheses about distinct associations and etiological relationships between limited HL and risk factors across age groups should be on the future research agenda. The findings in this study, particularly the idea of a subpopulation approach, provide promising leads for targeting future intervention programs to address the issue of highly prevalent limited HL in adult populations.

References

- [1] M. Kutner, E. Greenberg, Y. Jin, and C. Paulsen, *The Health Literacy of America's Adults: Results from the 2003 National Assessment of Adult Literacy*, U.S. Department of Education, Washington, DC, USA, 2006.
- [2] L. Nielsen-Bohlman, A. M. Panzer, and D. A. Kindig, *Health Literacy: A Prescription to End Confusion*, National Academy Press, Washington, DC, USA, 2004.
- [3] D. Schillinger, A. Bindman, F. Wang, A. Stewart, and J. Piette, "Functional health literacy and the quality of physician-patient communication among diabetes patients," *Patient Education and Counseling*, vol. 52, no. 3, pp. 315–323, 2004.
- [4] S. C. Ratzan, "Health literacy: communication for the public good," *Health Promotion International*, vol. 16, no. 2, pp. 207–214, 2001.
- [5] M. Kutner, E. Greenberg, and J. Baer, *A First Look at the Literacy of America's Adult in the 21st Century*, U.S. Department of Education, Washington, DC, USA, 2005.
- [6] T. S. Murray, J. Hagey, J. D. Willms, and R. Shillington, *Health Literacy in Canada: A Healthy Understanding 2008*, Canadian Council on Learning, Ottawa, Canada, 2008.
- [7] C. L. Bennett, M. R. Ferreira, T. C. Davis et al., "Relation between literacy, race, and stage of presentation among low-income patients with prostate cancer," *Journal of Clinical Oncology*, vol. 16, no. 9, pp. 3101–3104, 1998.
- [8] I. Rootman and B. Ronson, "Literacy and health research in Canada: where have we been and where should we go?" *Canadian Journal of Public Health*, vol. 96, no. 2, pp. S62–S77, 2005.
- [9] T. C. Davis, M. S. Wolf, P. F. Bass et al., "Literacy and misunderstanding prescription drug labels," *Annals of Internal Medicine*, vol. 145, no. 12, pp. 887–894, 2006.
- [10] D. W. Baker, "Reading between the lines: deciphering the connections between literacy and health," *Journal of General Internal Medicine*, vol. 14, no. 5, pp. 315–317, 1999.
- [11] M. S. Wolf, J. A. Gazmararian, and D. W. Baker, "Health literacy and health risk behaviors among older adults," *American Journal of Preventive Medicine*, vol. 32, no. 1, pp. 19–24, 2007.
- [12] Statistics Canada and Organisation for Economic Co-operation Development, *Literacy, Economy and Society: Results of the First International Adult Literacy Survey*, OECD Publications and Information Center, 1995.
- [13] D. Schillinger and T. Davis, "A conceptual framework for the relationship between health literacy and health care outcomes: the chronic disease exemplar," in *Understanding Health Literacy: Implications for Medicine and Public Health*, J. G. Schwartzberg, J. VanGeest, and C. Wang, Eds., American Medical Association Press, Chicago, Ill, USA, 2005.
- [14] S. C. Kalichman, B. Ramachandran, and S. Catz, "Adherence to combination antiretroviral therapies in HIV patients of low health literacy," *Journal of General Internal Medicine*, vol. 14, no. 5, pp. 267–273, 1999.
- [15] R. Desjardins, "Determinants of literacy proficiency: a lifelong-lifewide learning perspective," *International Journal of Educational Research*, vol. 39, no. 3, pp. 205–245, 2003.
- [16] S. R. Schecter and J. Lynch, "Health learning and adult education: in search of a theory of practice," *Adult Education Quarterly*, vol. 61, no. 3, pp. 207–224, 2011.
- [17] A. M. Kilbourne, G. Switzer, K. Hyman, M. Crowley-Matoka, and M. J. Fine, "Advancing health disparities research within the health care system: a conceptual framework," *American Journal of Public Health*, vol. 96, no. 12, pp. 2113–2121, 2006.
- [18] G. Rose, K. T. Khaw, and M. Marmot, *Rose's Strategy of Preventive Medicine: The Complete Original Text*, Oxford University Press, Oxford, UK, 2008.
- [19] P. A. Braveman, C. Cubbin, S. Egerter et al., "Socioeconomic status in health research: one size does not fit all," *Journal of the American Medical Association*, vol. 294, no. 22, pp. 2879–2888, 2005.
- [20] B. D. Smedley and S. L. Syme, "Promoting health: intervention strategies from social and behavioral research," in *Committee on Capitalizing on Social Science and Behavioral Research to Improve the Public's Health*, Division of Health Promotion and Disease Prevention, Ed., National Academy Press, Washington, DC, USA, 2000.
- [21] P. Braveman and E. Tarimo, "Social inequalities in health within countries: not only an issue for affluent nations," *Social Science and Medicine*, vol. 54, no. 11, pp. 1621–1635, 2002.
- [22] T. S. Murray, Y. Clermont, and M. Binkley, *Measuring Adult Literacy and Life Skills: New Frameworks for Assessment*, Statistics Canada, Ottawa, Canada, 2005.
- [23] "The international adult literacy and skills survey, 2003," Public Use Microdata File User's Manual, Statistics Canada, Ottawa, Canada, 2002.
- [24] Statistics Canada, *Literacy in the Information Age: Final Report of the International Adult Literacy Survey*, OECD, Ottawa, Canada, 2000.
- [25] I. Kirsch and M. Lennon, "The ICT literacy framework," in *Measuring Adult Literacy and Life Skills: New Framework for Assessment*, T. S. Murray, Y. Clermont, and M. Binkley, Eds., Statistics Canada, Ottawa, Canada, 2005.
- [26] D. Schillinger, K. Grumbach, J. Piette et al., "Association of health literacy with diabetes outcomes," *Journal of the American Medical Association*, vol. 288, no. 4, pp. 475–482, 2002.
- [27] T. L. Scott, J. A. Gazmararian, M. V. Williams, and D. W. Baker, "Health literacy and preventive health care use among medicare enrollees in a managed care organization," *Medical Care*, vol. 40, no. 5, pp. 395–404, 2002.
- [28] S. C. Lemon, J. Roy, M. A. Clark, P. D. Friedmann, and W. Rakowski, "Classification and regression tree analysis in public health: methodological review and comparison with logistic regression," *Annals of Behavioral Medicine*, vol. 26, no. 3, pp. 172–181, 2003.
- [29] K. Y. Chan and W. Y. Loh, "Lotus: an algorithm for building accurate and comprehensible logistic regression trees," *Journal of Computational and Graphical Statistics*, vol. 13, no. 4, pp. 826–852, 2004.
- [30] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Chapman & Hall/CRC, New York, NY, USA, 1984.

- [31] G. V. Kass, "An exploratory technique for investigating large quantities of categorical data," *Applied Statistics*, vol. 29, no. 2, pp. 119–127, 1980.
- [32] W.-Y. Loh, "Logistic regression tree analysis," in *The Handbook of Engineering Statistics*, H. Pham, Ed., Springer, London, UK, 2006.
- [33] K. F. Ferraro and T. P. Shippee, "Aging and cumulative inequality: how does inequality get under the skin?" *Gerontologist*, vol. 49, no. 3, pp. 333–343, 2009.
- [34] I. Rootman and J. Wharf-Higgins, "Literacy and health: implications for active living," *Well Spring*, vol. 18, no. 4, pp. 1–3, 2007.
- [35] V. A. Freedman, L. G. Martin, and R. F. Schoeni, "Recent trends in disability and functioning among older adults in the United States: a systematic review," *Journal of the American Medical Association*, vol. 288, no. 24, pp. 3137–3146, 2002.
- [36] I. S. Kickbusch, "Health literacy: addressing the health and education divide," *Health Promotion International*, vol. 16, no. 3, pp. 289–297, 2001.
- [37] M. Marmot, "Social determinants of health inequalities," *Lancet*, vol. 365, no. 9464, pp. 1099–1104, 2005.
- [38] D. Nutbeam, "The evolving concept of health literacy," *Social Science and Medicine*, vol. 67, no. 12, pp. 2072–2078, 2008.
- [39] D. Dannefer, "Cumulative advantage/disadvantage and the life course: cross-fertilizing age and social science theory," *Journals of Gerontology B*, vol. 58, no. 6, pp. S327–S337, 2003.
- [40] I. Rootman and D. Gordon-El-Bihbety, *A Vision for a Health Literate Canada*, Canadian Public Health Association, Ottawa, Canada, 2008.
- [41] *Healthy People 2010 with Understanding and Improving Health and Objectives for Improving Health*, U.S. Department of Health and Human Services, Washington, DC, USA, 2nd edition, 2000.
- [42] P. Franks and K. Fiscella, "Reducing disparities downstream: prospects and challenges," *Journal of General Internal Medicine*, vol. 23, no. 5, pp. 672–677, 2008.
- [43] D. Haber, *Health Promotion and Aging: Practical Applications for Health Professionals*, Springer, New York, NY, USA, 2007.
- [44] G. A. Mensah and M. J. Glover, "Epidemiology of racial and ethnic disparities in health and healthcare," in *Eliminating Healthcare Disparities in America*, R. A. Williams, Ed., pp. 21–40, Humana Press, Totowa, NJ, USA, 2007.
- [45] S. J. Shaw, C. Huebner, J. Armin, K. Orzech, and J. Vivian, "The role of culture in health literacy and chronic disease screening and management," *Journal of Immigrant and Minority Health*, vol. 11, no. 6, pp. 460–467, 2009.
- [46] N. D. Berkman, S. Sheridan, K. E. Donahue et al. et al., *Health Literacy Interventions and Outcomes: An Updated Systematic Review*, Agency for Healthcare Research and Quality, Rockville, Md, USA, 2007, contract no. 199.

