Commissioned Paper
January 2022

Utilization of Preventative Healthcare by Native and Non-Native Speakers of English

Author:
Tanya Longabach

Suggested Citation: Longabach, T. (2022). *Utilization of preventive healthcare by native and non-native speakers of English*. Retrieved [insert date] from PIAAC Gateway website: [insert link]. Washington, D.C.

This project has been funded by the American Institutes for Research through a contract with the National Center for Education Statistics (NCES) of the U.S. Department of Education. This report is based on PIAAC data released in March 2016. The views expressed in this paper do not necessarily reflect the views or policies of the American Institutes for Research, National Center for Education Statistics, or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement of same by the U.S. Government.

AIR-PIAAC Contact:
Jaleh Soroui (AIR-PIAAC Director)
Saida Mamedova (Senior Research Analyst)
PIAACgateway.com
piaac@air.org

Author Contact:
Kaplan Professional
Tanya Longabach at tlongabach@yahoo.com

Tanya Longabach

UTILIZATION OF PREVENTIVE HEALTHCARE BY NATIVE AND NON-NATIVE SPEAKERS OF ENGLISH

INTRODUCTION

While there is no shortage of research explaining people's health service utilization, our understanding of factors accounting for variations in immigrants' utilization remains underdeveloped (Yang & Hwang, 2016). A number of studies have outlined the underutilization of healthcare by the immigrant population (Derose, Bahney, Lurie, & Escarce, 2009; Lucas, Barr-Anderson, & Kington, 2003; Pandey & Kagotho, 2010; Xu & Borders, 2008). Barriers to accessing preventive care, and healthcare in general, include lack of language proficiency (Chang, Chan, & Han, 2015; Lebrun, 2012; Ponce et al., 2006; DeRose, Escarce, & Lurie, 2007; Jacobs, Chen, Karliner, Agger-Gupta, & Mutha, 2006), inability to afford health care (Aguilera & Massey, 2003), lack of health insurance (Pylypchuk & Hudson, 2009), and lack of information about health conditions and services (Perreira et al., 2012). Underutilization of health care services is directly associated with poorer health conditions, such as longer stays in hospital, more serious health problems, and heightened mortality rates (Pylypchuk & Hudson, 2009). Understanding the factors leading to these disparities is critical to ameliorating them, and consequently being able to provide appropriate care for the country's population, and possibly reducing future health expenditures (Yang & Hwang, 2016).

POPULATION OF INTEREST

This paper examines some of the factors affecting preventive healthcare utilization (PHCU) among non-native speakers of English in the US. A number of studies related to the utilization of healthcare focus on the immigrant status of participants (being born in a country other than the US, in case of immigrants to the US). The term "non-native speakers of English" may sometimes be used to refer to the population identified as immigrants (those who were not born in the United States, but moved here later in life), however non-native speaker status is not necessarily equal to immigrant status. Clearly for some immigrants, English is a native language (consider, for example, immigrants from New Zealand or Australia), while at the same time, there are people who were born in the United States, for whom English is not a native language.

To make this distinction clearer, this paper compares PHCU for the following two groups: native speakers of English, and non-native speakers of English, regardless of their immigration status. The emphasis on language status is made due to the fact that in addition to the demographic factors that impact PHCU, we also examine factors related to language skills within the non-native speaker group.

In previous studies, the variables impacting healthcare utilization are usually explored in relation to either immigrant healthcare utilization, or non-native speaker healthcare utilization. Studies examining the impact of income and education on healthcare utilization often use immigrant status as a characteristic based on which study sample is selected. On the other hand, studies specifically examining language barriers to obtaining health care use non-native speaker status to select the study population. As a result, these two characteristics – immigrant status and non-native speaker status – are necessarily confounded in these studies. Our study examines many variables that are commonly researched both for the immigrant population (e.g., demographic variables, income, education) and non-native speaker population (e.g., language-related variables) as applied to the non-native speaker population. For that reason, when we discuss these previous studies, we sometimes talk about the population of immigrants, and at other times about the population of non-native speakers, depending on which population was emphasized in a specific study in which a given variable was researched. The purpose of this study is not to separate the effects of immigrant status and non-native speaker. We believe that this is an important topic that may deserve a separate study of its own, and therefore is not covered here.

RESEARCH QUESTIONS

Based on results from previous studies, our hypothesis is that non-native speakers of English are less likely to use preventive care services than native speakers of English. Other factors impacting the likelihood of preventive care use are age, gender, ethnicity, income, educational status, employment status, having medical insurance, length of residence in the US, and self-described health status. In addition, we hypothesize that those non-native speakers who learned English later in life (after the age of 16), have a lower likelihood of preventive care utilization than those who learned English before the age of 15. This hypothesis is based on prior research findings concerning the negative impact of low English language proficiency on healthcare utilization. Our assumption is that those non-native speakers who learned English later in life are [may be] less likely to obtain near-native proficiency, which may play a significant role in this population being able to effectively communicate with healthcare providers, or obtain information on what healthcare services are necessary and available, and how to access them. We also speculate that specific language area skills (understanding, speaking, reading, and writing) may have an impact on preventive care utilization, with less skilled individuals receiving less preventive care. Finally, we explore whether length of residence in the US impacts PHCU.

There are 8 variables in the PIAAC questionnaire that pertain to obtaining preventive healthcare:

1. Getting a flu shot in the past year
2. Getting a mammogram in the past year
3. Getting a pap smear in the past year
4. Getting colon cancer screening in the past year
5. Getting vision screening in the past year
6. Getting prostate cancer screening in the past year
7. Getting osteoporosis screening in the past year
8. Seeing a dentist in the past year

Of these eight, only three (getting a flu shot, getting vision screening, and seeing a dentist) are typically performed yearly. Consequently, when a person responds negatively to any except these three questions, it does not necessarily signify that they are not receiving the appropriate preventive care. Because of this, and because these three preventive care interventions are not gender-specific, unlike, for example, getting a mammogram or a prostate cancer screening, the decision was made to use only these three questions as indicators of PHCU.

In addition to looking at native-speaker status and time of learning English (early vs late) the study also looked at a number of demographic, socioeconomic, health-related, cognitive, and language-related factors impacting PHCU. We chose to include cognitive factors as well because demographic, socioeconomic, and linguistic factors are far better researched than cognitive skills.

The demographic factors include age, gender, and ethnicity. Socioeconomic factors include education, employment status, and income. Health-related factors include self-rated health status and having medical insurance. Language-related factors include native speaker status; self-rated language skills these areas (understanding, speaking, reading, and writing), and length of residence in the U.S. Cognitive factors include literacy and numeracy skills. Although demographic and cognitive factors affect the general population, regardless of the native speaker status, they may impact native speakers and non-native speakers differently. Linguistic factors, on the other hand, are specific to the non-native speaker population. Therefore, the first research question will examine the impact of native speaker status on PHCU while controlling for factors common to both native and non-native speakers, while the subsequent research questions will explore the relationship between PHCU and other factors that may impact only the non-native speaker population. These include length of residence, age of language acquisition, and self-rated language skills.

Initially, we compared PHCU of native and non-native speakers of English, while controlling for the demographic, socioeconomic, health-related, and cognitive factors mentioned above, in order to determine which factors impact PHCU for the two groups. Then, we compared PHCU within the non-native speakers of English group based on the age of language acquisition, controlling for demographic, socioeconomic, health-related, cognitive, and language-related factors mentioned above. We also evaluated the impact of self-rated skills in specific linguistic domains (speaking, listening, reading, writing) on PHCU, controlling for the same demographic, socioeconomic, health-related, cognitive, and language-related factors mentioned above. This approach allowed us not only to explore the differences between PHCU of native and non-native speakers, but also to see what specific factors affect PHCU of non-native speakers. Since non-native speakers are not a homogeneous group and often vary significantly based on demographic and socioeconomic factors, this approach is useful in identifying these differences.

Our research questions are as follows:

1. Are there differences in the utilization of preventive care (flu shot, vision checkup, dentist visit) between native speakers of English vs non-native speakers of English, controlling for demographic factors (age, gender, ethnicity), socioeconomic factors (income, educational status, employment status), health-related factors (having medical insurance and self-described health status), and cognitive factors (literacy and numeracy skills)? Is the relationship between literacy and PHCU stronger than the relationship between numeracy and PHCU?
2. For non-native speakers of English, does the age of English language acquisition and self-rated language proficiency have an impact on PHCU, controlling for the demographic, socioeconomic, health-related, and cognitive factors listed above?

DEMOGRAPHIC CHARACTERISTICS

AGE

All else being equal, older people are more likely to seek medical care than younger ones because of their biological needs as a result of aging (Schmidt et al., 2011). At the same time, there may be some interaction between immigrant status and age on PHCU. The impact of age on healthcare utilization has been researched largely based on immigrant status, rather than native speaker status. Some studies found that, on average, new immigrants tend to be younger, and therefore healthier, than the natives, and thus potentially less likely to seek health care services. However, the longer they stay in the host country, the more preventive and other healthcare they start using (Antecol & Bedard, 2006; Hernández-Quevedo & Jiménez-Rubio, 2009; McDonald & Kennedy, 2004, 2005). This appears to be due to several factors, including being diagnosed with new medical conditions or developing an exacerbation of old ones; acculturation and acquisition of poorer health habits; as well as obtaining more education, a better job, and health insurance (Antecol & Bedard, 2006; Hernández-Quevedo & Jiménez-Rubio, 2009; McDonald & Kennedy, 2004, 2005).

GENDER

A number of studies finds that females in general are more likely to seek care than males (Broyles et al., 1999; Ye, Mack, Fry-Johnson, & Parker, 2012). This may be due partly to their need to access care during pregnancy and child birth, or due to socialization. Schmidt et al. (2011) found that women are more likely to be insured, visit a healthcare provider, participate in some preventive health screening, and consequently to report one or more chronic diseases.

ETHNICITY

Study findings vary regarding PHCU differences by race/ethnicity. For example, a number of studies found that while PHCU tends to be lower for racial/ethnic minorities than for Whites, these differences become nonsignificant after controlling for variables such as income, education level, and insurance status (Thorpe et al., 2013; Doty & Weech-Maldonado, 2003; Lees et al., 2005), while others find that the differences persist even after controlling for those variables (Chi & Handcock, 2014), suggesting that the sources of disparities need additional investigation. Doty & Weech-Maldonado (2003) showed that prior to controlling for any variables, significant disparities were noted in preventive dental care by ethnicity, with the white group having the greatest proportion of individuals who reported having regular dental checkups (65.2 percent). Mexican American individuals had the lowest proportion of regular dental checkup users (42.0 percent). Racial/ethnic minorities have lower utilization of preventive dental care even after controlling for other predisposing characteristics (gender, age, education, and health status). However, these disparities were eliminated once they controlled for two of the enabling resources (income level and insurance). This finding suggests that enabling resources are important in estimating preventive dental care utilization differences among racial/ethnic groups. Similarly, Lees et al. (2005) found that while there are some differences between Whites, Blacks, and Spanish- and English-speaking Hispanics in the use of flu vaccination, after adjusting for factors such as personal health characteristics, SES, and measures of access to care, these differences became nonsignificant.

On the other hand, Chi and Handcock (2014) report that among all immigrants, American Indians, Alaskan Natives, Asians, and others reporting a single race, were approximately 70 % less likely to report having had a doctor's visit than were African Americans. White immigrants were less likely to report having had a doctor's visit than African American immigrants. Multirace immigrants were approximately 70 % less likely to report delays in obtaining medical treatment than African American immigrants. Wu et al. (2013) found that despite a significant upward trend of frequency of dental cleaning from 1999 to 2008 in 2008 still only 56–77% of any ethnic or racial group reported having had a dental cleaning in the previous 12 months. Relative to Whites, Blacks were less likely to have a dental cleaning in the previous 12 months. These variations persisted even when SES, health conditions, health behaviors, and number of permanent teeth were controlled. In contrast, Hispanics, Asians, and American Indians/Alaskan Natives did not differ from Whites in dental cleanings. Zhang et al. (2012) found racial/ethnic differences in eye doctor visits. For example, in 2008, non-Hispanic whites were more likely to visit an eye doctor than Hispanics (52.6% vs. 36.9%). They also observed some significant racial/ethnic differences in the inability to afford eyeglasses in several years. For example, in 2008, Hispanics were more likely to report being unable to afford eyeglasses when needed than non-Hispanic whites and non-Hispanic blacks (26.7% vs. 16.0% and 15.3%).

SOCIOECONOMIC FACTORS

INCOME, EDUCATIONAL STATUS, EMPLOYMENT STATUS, AND HEALTH INSURANCE

These four factors are often interrelated. Again, these variables are often researched for the immigrant population, with native speaker status often being a variable confounded with immigrant status. Immigrants, especially new immigrants who have limited English proficiency (LEP), tend to be less educated than native-born English speakers, land in unskilled and low-paying jobs (e.g., Aguilera & Massey, 2003) and are less likely to have medical health insurance (Ku & Matani, 2001; Hall & Cuellar, 2016). People with a higher level of education may be more likely to seek health service (Hernández-Quevedo & Jiménez-Rubio, 2009), especially preventive health services, than those with a lower level of education. One reason for that is that their knowledge of health and their resources allow them to treat health problems in a timely fashion and to prevent health problems from happening (Alegría et al., 1991; Blackwell et al., 2009; Parslow, Christensen, & Jacomb, 2002; Szwarcwald, Souza-Junior, & Damacena, 2010). At the same time, those with higher education tend to have higher paying jobs with employee-sponsored insurance.

Health insurance has been found to be a factor impacting healthcare utilization in general, regardless of immigration and language status (Smith et al., 2018; Thorpe et al., 2013), with those without insurance tending to delay or completely forgo care. It has been found to be a consistent predictor of immigrant health services utilization (Schmidt, 2011, Pylypchuk & Hudson, 2009). Some of the differences between citizen and non-citizen immigrants are attributable to differences in rates of employer sponsored health insurance (Carrasquillo, Carrasquillo, & Shea 2000). Length of time in US, English proficiency, and access to publicly funded health insurance also influence health insurance coverage and access to care for immigrants (Derose, Escarce, & Lurie, 2007). Employment is a key variable in health insurance coverage. Overall immigrants have lower rates of health insurance and receive lower quality of care than US born populations (Derose, Escarce, & Lurie, 2007). Intermittent lack of health insurance also results in less use of preventative services (Sudano & Baker 2003).

At the same time, having health insurance is not a guarantee that immigrants will be using preventive care at the same rates as native-born population. Pylypchuk and Hudson (2009) found that even after detailed control for insurance schemes, the foreign-born population is less likely to use preventive care than natives. The reasons for this are unclear, but perhaps immigrants obtain less comprehensive insurance than natives or are not aware that insurance covers preventive tests. However, given the fact that the results hold even for such inexpensive and simple measures as a visit to a provider and a flu shot, these findings suggest that other factors may be at work as well. Immigrants' time constraints, or exhibition of moral hazard may simply be different from those of natives.

HEALTH STATUS

Clearly, self-rated health status is a valid predictor of seeking care, including preventive care (Morris, Sutton, & Gravelle, 2005; Giltay, Vollaard, & Kromhout, 2012; Mellner & Lundberg, 2003). Health status has been found to be the strongest predictor of health services utilization by some researchers (de Boer, Wijker, & de Haes, 1997; Hernández-Quevedo & Jiménez-Rubio, 2009). The logic is that persons who perceive themselves to be healthy are less likely to seek care. Therefore, those who say that they are not healthy are more likely to seek care. At the same time, if people use preventive care, they tend to stay healthy. The directionality of this relationship therefore is somewhat different for the utilization of preventive care than for the utilization of sick care. Additionally, self-reported health status may be expressed and perceived differently across gender and ethnic lines, with some populations tending to report more prior medical conditions (Shin, Shin, & Rhee, 2012; Campbell & Edwards, 2012).

COGNITIVE VARIABLES

LITERACY

It may be expected that a positive relationship exists between literacy levels and education level. Those with higher education levels and higher income also tend to self-rate their health status higher, although the directionality of this relationship is not clear. In addition, adults with low literacy skills tend to have less secure jobs with lower income (Lunze & Paasche-Orlow, 2014). A study by Yamashita and Kunkel (2015) indicates that literacy skills mediated the effect of education on health, but the degree of mediation varied by different study locations.

NUMERACY

The relationship between numeracy and health outcomes in general, and PHCU in particular, is not well understood (Yamashita & Kunkel, 2015; Berkman et al., 2011).
Numeracy skills are necessary for preventive health care decisions, such as when individuals estimate the risks and benefits of their decisions and behaviors (Cavanaugh et al., 2008). A study by Yamashita et al. (2018a) showed that numeracy is a strong predictor of having had a dental checkup in the past year, but it is not predictive of having had a vision screening, influenza vaccination, or osteoporosis screening after accounting for theoretically meaningful factors.

Not only that, but numeracy is important to be able to evaluate whether a person has sufficient resources to afford care and plan to obtain the care necessary. Smith et al. (2018) report that individuals often avoid preventive care even despite being insured, because they do not know how insurance works and how much they will need to pay for the care. These concerns about the cost of care can cause consumers anxiety and distress and deter them from obtaining care. They found that people who delayed or avoided receiving any medical care were more likely to perceive health-related financial expenditures as burdensome, had lower subjective numeracy skills and less knowledge about how health insurance works, including lack of knowledge that many preventive care services are covered at no out-of-pocket cost.

Overall, adults who struggle with literacy, numeracy, and poor health are disproportionately people of color, the elderly, and those with limited education, income, and English proficiency (Kutner et al., 2006).

LANGUAGE-RELATED VARIABLES

ENGLISH LANGUAGE SKILLS

A higher level of English proficiency enables individuals to have better access to the health care system (Chang, Chan, & Han, 2015; Lebrun, 2012; Ponce et al., 2006; DeRose, Escarce, & Lurie, 2007; Jacobs, Chen, Karliner, Agger-Gupta, & Mutha, 2006). When compared with English speakers, people whose main spoken language is not English are less likely to understand the processes necessary to become insured (Feinberg et al., 2002) and to remain insured (Jang, Lee, & Woo 1998), to receive preventive care (Harlan, Bernstein, & Kessler, 1991; Liao & McIlwaine, 1995; Naish, Brown, & Denton, 1994; Woloshin et al., 1997), to have a regular source of primary care (Jang, Lee, & Woo, 1998; Weinick and Krauss 2001), and to receive timely eye, dental, and physical examinations (Kirkman-Liff & Mondragon, 1991). Factors, such as having English as a second language, which is not frequently spoken or a means of communication in the household, could affect understanding of the consequences of the disease and when to take actions to stop progression of an illness or seek alternative/available health service/treatment (Jacobson, Hund, & Mas, 2016).

DeRose, Escarce, and Lurie (2007) note that adults with limited English proficiency and their children are much less likely to have insurance and a usual source of care, have fewer physician visits, and receive less preventive care than those who only speak English. Similarly, Schmidt et al. (2011) note that individuals with some English speaking ability are several times more likely to participate in regular physical exams than those who do not speak English. Respondents who speak English well are several times more likely to have health insurance.

VARIABLES RELATED TO SECOND LANGUAGE ACQUISITION

Since more advanced English language skills are associated with the ability to obtain a better job, education, and health insurance, it makes sense to examine factors related to English as a second language acquisition in order to control for those variables when examining the relationship between PHCU and native speaker status. While many factors can impact second language acquisition, the important variables that are available in the PIAAC survey data are age of English language acquisition and length of residence in the US. A number of studies have found a significant relationship between language proficiency and either both variables, one of the two, or their interaction (Kasper & Rose, 2002; Krashen et al., 1979).

AGE OF SECOND LANGUAGE ACQUISITION

While it is still not completely clear how the critical period hypothesis applies to second language acquisition, many linguists agree that a person who started learning a second language later in life, specifically after puberty, is less likely to attain native-like proficiency than someone who acquired a second language before that time (Johnson and Newport, 1989; Newport, 1990; Birdsong, 1999). Therefore, the answer to the age of acquisition question serves as somewhat of an objective measure of language skills, as compared to other self-rated measures related to the ability to speak, listen, read, and write. We hypothesize that those who acquired English at the age of 16 or later may have lower language skills.

LENGTH OF RESIDENCE IN THE US

Clearly, some native-born U.S. residents' native language may not be English, just as the native language of some of immigrants' into the U.S. may be English. However, a more common situation is for an immigrant to the U.S. to also be a non-native speaker of English. Studies often find that significant disparities exist in healthcare utilization between recent and non-recent immigrants (Chi & Handcock, 2014; Schmidt et al., 2011). Length of residence is positively related to healthcare utilization in general, and PHCU in particular. Length of residence is related to a number of other factors impacting PHCU. In a study by Chi and Handcock (2014), health care access and utilization differences between recent and non-recent immigrants were driven primarily by limited English proficiency (LEP), insurance status, public assistance usage, and poverty level. Not only are newer immigrants more likely to underutilize health care, but also their underutilization is related to their lack of insurance, lack of adequate financial resources, and inability to navigate the health care system due to LEP.

Chang, Chan, and Han (2015) observed that English proficiency and length of residence are often linked because English proficiency generally improves with more time spent in the United States, yet the findings suggest that these variables represent 2 distinct barriers to health care access. Language barriers suggest underlying difficulties in communication and information seeking, and length of residence likely represents knowledge of the health care system and other cultural norms. Language barriers may persist for some immigrants regardless of time in the United States, whereas other recent immigrants arrive fluent in English.

Schmidt et al. (2011) stipulates that this increase in PHCU occurs for several reasons. On the one hand, length of residence may mean higher acculturation (adoption of the home country cultural ways and habits), which leads to higher awareness of "how things work" in general, such as how health insurance works and where to obtain health services. These factors are likely to cause higher PHCU. On the other hand, usually the longer the person resides in the country, the better his or her language skills become. The improved language skills in turn lead to increased PHCU. Effects of acculturation may be negative as well as positive, however. Some authors note that immigrants may be healthier when they arrive in the US, but their health deteriorates as they acquire poor dietary and exercise habits practiced by US residents (Antecol & Bedard, 2006; Schmidt et al., 2001).

METHODOLOGICAL ISSUES WITH PUTTING SEVERAL LANGUAGE-RELATED VARIABLES IN A REGRESSION MODEL

Methodologically, due to the linear relationship between the individual's age, age of language acquisition, and length of residence in the host country, putting these three variables in the same regression model presents a problem. The general problem of linear dependence among trios of variables such as these is often referred to as the "age-period-cohort" problem (Stevens, 2006). Because the variables are linearly dependent (as person's chronological age increases, the length of residence in a host country typically increases as well, and so does the difference between the person's chronological age and the age of language acquisition), putting them all in the model at once leads to multicollinearity. Researchers deal with this issue in a variety of ways, frequently by including only one of the three variables in the regression model, which may lead to erroneous conclusions. In her study, Stevens (2006) outlines several ways of dealing with this problem. Some of them include collecting additional or better data, which is not possible in our circumstances. A longitudinal study assigning different slopes to the relationship between language proficiency variables and independent variables is also not a possible course of action for this paper. However, Stevens makes an important point that these variables are only proxies for other variables that are really making an impact on language acquisition. For example, it is not the length of residence that really matters, but the environment in which the individual spends their time during that residence period. If they are not integrated into the community that speaks the target language, the length of residence would not have an impact on their language proficiency. Flege and Liu's 2001 study shows that the effects of length of residence on measures of L2 proficiency vary according to the respondents' social roles. Similarly, it is not the chronological age that really matters, but whether the individuals are spending their years in an educational environment that helps them learn the target language. Stevens (2006) suggests that collecting more meaningful information may be the solution to the age-period-cohort problem. For example, it might be appropriate to obtain the number of native speakers living in the same household instead of length of residence. This substitution could be made on the grounds that the longer immigrants live in their country of destination, the more likely they are to live with a spouse or children who are native speakers of the respondent's L2 and, thus, the more intense the exposure to the opportunities to learn and to improve in the L2. Alternatively, the effect of age could be modeled through an aggregate measure of school enrollment (or educational achievement) because numerous studies suggest that educational characteristics are strongly associated with L2 proficiency (e.g., Flege & Liu, 2001; Flege, Yeni-Komshian, & Liu, 1999; Hakuta et al., 2003). For this reason, since we already have a variable for education level, we feel that it is appropriate to exclude length of residence as a variable and keep education level as a proxy for length of residence, thus resolving the issue of the linear relationship between age, age of acquisition, and length of residence.

METHODOLOGY

DATA

We used PIAAC 2012-14 data for this study. PIAAC is a survey organized by the OECD and conducted by each participating country (NCES, 2016a, 2016b). PIAAC examines a range of skills and assesses skills consistently across participating countries. Twenty-four countries, including the U.S., participated in Round 1 of the survey in 2011-2012, and an additional nine countries participated in Round 2 in 2014-2015. Additionally, a national supplemental survey was completed in the U.S. in 2014, focusing on younger workers, older workers (ages 66 to 74), and an increased sample of unemployed adults. The first round collected data from 5,010 adults, and the supplemental survey collected data from 3,660 adults, for a total of 8,670 adults ages 16 to 74 (Hogan et al., 2016).

We find PIAAC data appropriate for answering our research question because it provides information about key variables, such as self-described health status and the use of preventive care, in combination with demographic variables, in one database. PIAAC employs a complex sampling methodology involving a four-stage stratified probability method requiring an extensive system of weights application. The IEA International Database Analyzer (IDB Analyzer) was used for data analysis. It is an application developed by the IEA Data Processing and Research Center (IEA-DPC) in Hamburg, Germany, that can be used to combine and analyze data from IEA's large-scale assessments, including PIAAC. An advantage of using the IDB analyzer is that it automatically applies weights to analyses, computing correct estimation of parameters and variance, which yields correct significance testing results.

MEASURES

Variables related to demographic characteristics

Age was recorded in age groups and was constructed as an ordinal variable with 6 levels, from youngest to oldest: 24 and younger, 25-34, 35-44, 45-54, and 66 and older. This variable is labeled AGEG10LFSEXT. Sex was coded as a dichotomous dummy variable (1 = female, 0 = male) and is labeled GENDER_R. We created 4 dichotomous variables for race based on 4 categories (White, Hispanic, Black, Other) with White as a referent. These four dichotomous variables are labeled RACETHN_4CAT. This is a non-standard classification for the U.S., because it does not include the Asian group. RACETHN_4CAT is a variable created by PIAAC, with 'Other' combining Asian/ Pacific Islander, American Indian/ Alaska Native, and 'Other' race.

Variables related to socioeconomic status

We used the ordinal variable EDCAT6 to indicate the respondent's education status. This variable had 6 levels: Lower secondary or less; Upper secondary; Post-secondary, non-tertiary; Tertiary - professional degree; Tertiary - bachelor degree; and Tertiary - master/research degree. We explored the possibility of using the ordinal variable EMP_6CAT with 6 levels to indicate the respondent's employment status but found that the sample size for non-native speakers at several categories was too small (below 62, which was considered too small in previous studies, cf. Cummins, Yamashita & Arbogast, 2018). For that reason, another variable C_D05 was used as the basis for the 3 dichotomous variables to indicate employment status: employed (1=yes, 0=no), unemployed (1=yes, 0=no), and out of labor force (1=yes, 0=no).

When accounting for income, we used the ordinal variable EARNMTHALLDCL with 10 levels, which represent 10 income deciles. However, for similar reasons of very small samples of non-native speakers at 6 out of 10 deciles, we had to create a dichotomous variable with 0 indicating lowest to 6th decile, and 1 indicating 5th through highest (10th) decile.

Health-related variables

Self-rated health status (PIAAC background questionnaire question ID I_Q08) in the original PIAAC dataset was constructed as an ordinal variable with 5 levels: 1=excellent, 2=very good, 3=good, 4=fair, and 5=poor. The scale was recoded in reverse (5=excellent, 4=very good, 3=good, 2=fair, 1=poor) for a more intuitive interpretation in the regression models. Having medical insurance (PIAAC background questionnaire question ID I_Q10bUSX1) was a dichotomous variable (0=no health insurance; 1=have health insurance). The dependent variables (receiving a flu shot, PIAAC background questionnaire question ID I_Q10bUSX3a, a vision checkup, PIAAC background questionnaire question ID I_Q10bUSX3e, and a dentist visit, PIAAC background questionnaire question ID I_Q10bUSX3h) were dichotomous variables (0=no, 1=yes).

Literacy and numeracy skills

PIAAC data provide 10 plausible values for literacy and numeracy skills. Plausible values represent the statistically estimated means for responses to literacy- and numeracy-related tasks in the PIAAC. These plausible values were used in our analysis in order to estimate literacy and numeracy skill scores. The estimated score ranged from 0 (the lowest proficiency) to 500 (the highest proficiency). Plausible, value-based literacy and numeracy assessment is known to return more accurate estimates of groups than conventional point estimates (Hogan et al., 2016). To determine whether literacy or numeracy has a stronger association with receiving preventive healthcare, both literacy and numeracy variables were entered in the regression model.

Language-related variables

For the determination of native speaker status (research question 1), we used a dichotomous variable NATIVESPEAKER (native speaker=1, non-native speaker=0). For research question 2, we used the variable LANGUAGE to describe age of English language acquisition. In the original PIAAC dataset, this variable has 3 levels (1=learned as first language, 2=learned at the age of 15 or younger, 3=learned at the age of 16 or older). However, when the sample consisted only of non-native speakers, it became a dichotomous variable, which we recoded as 0=at the age of 15 or earlier, 1=after the age of 15.
For research question 2, we also used several ordinal variables that indicated self-rated language skills in specific areas (understanding J_Q05cUSX3a, speaking J_Q05cUSX3b, reading J_Q05cUSX3d, and writing in English J_Q05cUSX3e). All 4 variables were on the following scale: 1=very well, 2=well, 3=not well, 4=not at all. The scale was recoded in reverse (4=very well, 3=well, 2=not well, 1=not at all) for a more intuitive interpretation in the regression models.

ANALYTIC APPROACH

Sampling weights

PIAAC data provide the survey sampling weights and replicate weights in order to estimate the nationally representative figures and correct standard errors. The final weight (SPFWT10) and the 80 replicate weights (SPFWT1 – SPFWT80) were applied in all estimation procedures. We employed the SPSS syntax developed by the IEA International Database Analyzer (IDB Analyzer) for data analysis. The SPSS syntax was designed to take the sampling weights, replicate weights and plausible values into account simultaneously. All analyses were conducted using SPSS version 25.

Weighted descriptive statistics

Weighted descriptive statistics were calculated for all variables by native speaker status. Missing values were excluded from the analysis. However, because the set of variables used for each research question varied, the final sample sizes were not the same for all analyses. Crosstabulations for the variables were calculated. It is important to have each cell populated at all variable levels to avoid situations where levels of one variable are completely conflated with levels of another variable (for example, if all Hispanics are also out of labor force, or if all Whites are employed). This situation may not only result in unexpected results and possibly variable coefficients changing signs, but also in very high correlations among confounded variables and increased variance information factor (VIF) when it is possible to calculate these values. However, it is not critical to have each crosstabulation cell populated at each variable cross section. Unless we're examining interactions between the variables in question, some cells may have to be left blank. For example, if we were to perform crosstabulations between age and employment, it is likely that those over the age of 70 would be only in the employment category 'out of labor force'. That is a clear confounding relationship; as long as the interaction of these variables is not included in the regression, the confounding relationship should not adversely impact the results of the regression.

Weighted correlations between variables

We performed correlations between continuous variables and ordinal variables treated as continuous that were included in the regression models (AGEG10LFSEXT, EDUCATION, health status, literacy and numeracy). Since there is no option to perform chi square test of relationships in IDB Analyzer, correlations between categorical variables were not performed. Instead, crosstabulations were examined to ensure that as many cells as possible in the crosstabulations were populated.

Weighted logistic regression analyses

Weighted logistic regression was used to evaluate statistical significance for all questions.
1. Are there differences in the utilization of preventive care (flu shot, vision checkup, dentist visit) between native speakers of English vs non-native speakers of English, controlling for demographic factors (age, gender, ethnicity), socioeconomic factors (income, educational status, employment status), health-related factors (having medical insurance and self-described

health status), and cognitive factors (literacy and numeracy)? Is the relationship between literacy and PHCU stronger than the relationship between numeracy and PHCU?
For research question 1, three separate binary logistic regressions were used to model three preventive care services (receiving a flu shot, getting a vision checkup, and having a dentist visit). The dependent variables were these dichotomous variables. The independent variable was native speaker status; the covariates were demographic variables (age, gender, ethnicity), socioeconomic variables (employment, income, education), health-related variables (self-rated health status and having medical insurance), and cognitive variables (literacy and numeracy scores). The sample consisted of native speakers and non-native speakers. The rationale for entering variables into the model in a specific way is as follows: in step 1, the primary variable of interest was entered into the model. (e.g., native speaker status or literacy and numeracy variables). In step 2, demographic, socioeconomic, and health-related factors were entered.

A number of models were run for each dependent variable. The first model included only the native speaker status to determine whether this variable by itself was significantly associated with the dependent variable. This analysis would reveal whether non-native speakers had higher odds of receiving preventive care services than native speakers. Additional independent variables were then added into the model in a stepwise fashion to test the statistical significance of their association with the dependent variable. Thus, subsequent regressions modeled preventive care utilization as a function of each of these additional independent variables. The statistical significance indicated that these independent variables were associated with the likelihood of receiving the preventive care service.

2. For non-native speakers of English, does the age of English language acquisition and self-rated language proficiency have an impact on PHCU, controlling for the demographic, socioeconomic, health-related, and cognitive factors listed above?

For research question 2, similarly, three separate, binary logistic regressions were conducted to model the likelihood of engaging in the three preventive care services (receiving a flu shot, getting a vision checkup, and having a dentist visit) for the non-native speaker sample only. For this reason, the sample consisted only of non-native speakers, since we were interested to see how the demographic, socioeconomic, and health factors affect that population. The independent variables were demographic variables (age, gender, ethnicity), socioeconomic variables (employment, income, education), health-related variables (self-rated health status and having medical insurance), and cognitive variables (literacy and numeracy scores). Additional variables specific to non-native speaker population were included in this regression model: age of English language acquisition and self-rated language skills.

As above, a number of models were run for each dependent variable, with several combinations of independent variables being added into the model in a stepwise-like fashion to test the statistical significance of their association with the dependent variable. The rationale for entering variables into the model in a specific way is as follows: in step 1 the primary variable of interest was entered into the model. (e.g., native speaker status or literacy and numeracy variables). In step 2, demographic, socioeconomic, and health-related factors were entered. The statistical significance indicated that these independent variables were associated with receiving the preventive care service for this group of respondents. Statistical significance was evaluated at the alpha 0.05 level.

Interpretation of logistic regression coefficients and odds ratios

In logistic regression, a logistic transformation of the odds (referred to as a logit) serves as a dependent variable. Log odds are expressed by the formula
log(odds) = logit (P) = $\ln(\frac{P}{1-P})$ where P is the probability of an event occurring and (1-P) is the probability of the event not occurring. A simple logistic regression with an independent variable X takes the form of logit (P) = $a+bX$.

Analogous to simple linear regression, b is interpreted as the expected change of logit(P) associated with a unit change in X. When b is positive, increases in X are associated with increases in logits. When b is negative, increases in X are associated with decreases in logits. Since b is the log odds, or the power to which e is taken, to calculate the odds ratio we need to take antilogs of b (Pedhazur, 1997). For example, if the b value is 0.23, then the odds ratio is $e^{0.23}$ = 1.26. The odds ratio is 1.26, which may be interpreted as 1.26/1. Group 1 has the odds of 1.26, as opposed to group 0 has the odds of 1 of incurring the event, resulting in odds ratio of 1.26/1 = 1.26. We can also say that group 1 has odds ratio that is 26% higher than the odds ratio of group 0 of the event occurrence.

When the value of b coefficient is negative, e is taken to the negative power to calculate log odds, resulting in values of log odds that are lower than 1. It is often suggested that for ease of interpretation, when an odds ratio is less than 1, it is preferable to use its reciprocal (Pedhazur, 1997). For example, suppose that the value of b is -0.14. Then the odds ratio is calculated as $e^{-0.14}$ = 0.87. This means that for each 1 point of increase in X, the odds are increasing 0.87 times. But to say that odds are increasing by a factor that is less than 1 means that the odds are decreasing. This interpretation is challenging to understand. Instead, especially when dealing with dummy variables, a reciprocal of the odds ratio is calculated to determine the increase in odds for a referent group, as opposed to the treatment group (or the group coded as 1). For example, if the referent group is 'female' and is coded as 0, and the b coefficient is -0.14, then the group coded as 1 ('male') will have lower odds of the event occurring. The odds of the event occurring for males would be 0.87. Instead, we calculate the reciprocal of odds for males: 1/0.87=1.15. If the odds are lower for males, then they must be higher for females. We interpret it as, 'females have 15% higher odds than males of the event occurring'.

RESULTS

Descriptive statistics

We examined the means and standard deviations (SD) for continuous variables and categorical (ordinal) variables treated as continuous (table 1A). The reason we calculated means and SD for categorical variables treated as continuous is because they were treated as continuous in the regression models. We understand that the means for categorical (ordinal) variables do not have the same meanings as the means for continuous variables. For that reason, both means and SDs, as well as percentages at each level of these variables were also calculated. We examined the same values by native speaker status (tables 2A and 3A). It is suggested that for analyses involving plausible values, an unweighted cell size of 62 cases or more is enough to report statistics including means, standard errors, standard deviations, a set of percentiles, a set of proficiency level percentages, and do regression analysis. For unweighted cell sizes of between 30 and 61, the data should be interpreted with caution. If the unweighted cell size is less than 30, analysis results should be suppressed. For analyses not involving plausible values, an unweighted cell size of 30 cases or more is enough to report descriptive statistics including means, standard errors, standard deviations, a set of percentiles, a set of percentages, and do regression analysis (AIR PIAAC Team, 2019).

For the full sample, all variable levels had a sufficient number of people for robust estimates for all but one variable level, ability to understand English = 'not at all' (45 responders). However, the self-rated language proficiency variables were not used for full sample models. For the crosstabs of native speaker status by different variable levels (table 3A), the minimum number of people at a variable level was not met for 'length of residence in the country' = 'in host country 5 or fewer years' and a number of levels of self-rated language proficiency variables for native speakers only. This is understandable, as very few native speakers would claim not being able to speak, read, write, or understand English, as well as living in the US for less than 5 years. However, the native speaker only sample was not considered in this study. The only variable levels where non-native speaker sample did not meet the minimum sample requirements was 'ability to understand English' = 'not at all (43 responders) and 'health status' = 'poor' (39 responders). Estimates for these groups are reported with caution. In the end, the full sample includes 8670 responders. The non-native speaker sample includes 1280 responders.

The non-native speakers tended to be significantly younger and less educated than native speakers. Perhaps not surprisingly, their self-rated language skills were lower than those of native speakers (table 2A). Approximately the same percentage of respondents were employed, unemployed, and out of labor force in the two groups. Slightly more non-native speakers rated their health as fair or poor. Approximately 32% of non-native speakers claimed to have no health insurance, as opposed to 16% of native speakers. Hispanics were the most prevalent ethnicity among non-native speakers, while whites were the most prevalent race among native speakers. In the weighted sample surveyed by PIAAC, the rate of having health insurance was 81.5%. This is somewhat lower than the health insurance coverage rate according to the US Census (91.2% covered in 2019 at the time of the interview; Keisler-Starkey & Bunch, 2020). The racial breakdown according to the US Census was very similar to the one noted in the PIAAC weighted responder sample (12% Black, 63% White, 16% Hispanic; Keisler-Starkey & Bunch, 2020).

Correlations

Correlations among variables that could be treated as continuous were calculated. For the overall group, there were only 5 such variables: literacy skills (PVLIT) and numeracy skills (PVNUM; both continuous variables), and health status, age, and education, 6 levels (ordinal variables, treated as continuous; the distribution of these variables is shown in tables 2A and 3A). All other variables were categorical, either dichotomous, or including several categories. The correlations are included in table 1.

Table 1. Weighted correlations between variables treated as continuous.

Variable	Education	Age	Health status	Literacy
Education	1.00			
Age	0.13	1.00		
Health status	0.23	-0.21	1.00	
Literacy	0.47	-0.14	0.29	1.00
Numeracy	0.49	-0.10	0.28	0.88

Note: all correlations were significant at $p<0.05$.

As expected, the highest correlation was between literacy and numeracy (.88). The next highest correlations were between literacy and numeracy and education status (.47 and .49, respectively). Correlations between health status and literacy and numeracy were moderate and positive. Correlations between age and literacy and numeracy were negative, but small.
For the non-native speaker group, correlations were calculated between literacy, numeracy, health status, age, education level, and self-rated language variables (ability to speak, read, write, and understand English). The correlations are included in table 2.

Table 2. Weighted correlations between variables treated as continuous for the non-native speakers of English.

Variable	Education	Age	Health status	Ability to speak English	Ability to read English	Ability to write English	Ability to understand English	Literacy	Numeracy
Education	1.00								
Age	0.10	1.00							
Health status	0.19	-0.26	1.00						
Ability to speak English	0.35	-0.20	0.26	1.00					
Ability to read English	0.40	-0.23	0.30	0.87	1.00				
Ability to write English	0.41	-0.23	0.30	0.87	0.91	1.00			

Ability to understand English	0.35	-0.21	0.26	0.88	0.85	0.82	1.00		
Literacy	0.51	-0.28	0.31	0.58	0.63	0.62	0.57	1.00	
Numeracy	0.55	-0.16	0.29	0.48	0.54	0.53	0.47	0.89	1.00

Note: all correlations were significant at $p<0.05$.

The highest correlations were between different self-rated language variables (speaking, reading, writing, and understanding). For this reason, it is best to avoid putting these variables in the same regression model to prevent multicollinearity. Otherwise, similar to the correlations for the whole group, correlations between education, numeracy and literacy, and self-rated language variables were moderate to high. Correlations between age and literacy and numeracy and self-rated language variables were negative but low. All correlations were significant at $p<0.05$ level.

Collinearity diagnostics

To check for the presence of multicollinearity, regression diagnostics of VIF and tolerance were performed in SPSS. Linear regression was performed to check for collinearity, which is considered to be an acceptable approach for measuring the extent of collinearity for logistic regression. The reason for this is as follows: multicollinearity diagnostics are not concerned with the dependent variable type (e.g., continuous vs dichotomous), or the values of the dependent variable. Instead, multicollinearity diagnostics are concerned with the relationships between independent variables. The values of VIF and tolerance will remain the same regardless of the type and values of the dependent variable, as long as all the independent variables' relationships are considered.
Collinearity diagnostics – VIF and tolerance -- were conducted on both weighted and unweighted data.

A range of VIF values has been proposed by researchers to indicate collinearity. While smaller VIF values are considered to be better, a value of 10 has been considered sufficiently low to make the decision that collinearity levels are acceptable, and recommended as the maximum level of VIF (e.g., Hair et al., 1995; Kennedy, 1992; Neter, Wasserman, & Kutner, 1989; Marquardt, 1970). However, a recommended maximum VIF value under 5 (e.g., Rogerson, 2001) or 4 (e.g., Pan & Jackson, 2008) can be found in the literature. While we did encounter some VIF values over 5 (e.g., when literacy and numeracy were entered into the model together), values over 10, or even approaching 10, did not occur. In the dataset including all responders, only the literacy variable had VIF>5.0. It was likely collinear with numeracy, whose VIF was 4.97. Once either literacy or numeracy were included in the model (tables 2B and 3B), all VIF values were below 2.0. Similarly, in the dataset including only nonnative speakers, all variables except for literacy had VIF below 5.0 (table 4B). Literacy VIF was 5.34. If either literacy or numeracy only were included in the model (tables 5B and 6B), both variables' VIF was below 2.0. Since all VIF values were below 10, no specific interventions were conducted to counteract collinearity in the regression models.

Findings: RQ1

Are there differences in the utilization of preventive care (dentist visit, vision check, flu shot) between native speakers of English vs non-native speakers of English, controlling for demographic factors (age, gender, ethnicity), socioeconomic factors (income, educational status, employment status), health-related factors (having medical insurance and self-described health status), and cognitive factors (literacy and numeracy)?

Executive Summary of Findings: RQ1

One common significant predictor for all three DVs (dentist visit, vision check, and flu shot) was having medical insurance; it was also the most consistent predictor of using preventive care. Some demographic/ socioeconomic/ educational predictors were significant for some DVs, but not for others. Their impact was inconsistent from one model to another, and from one preventive care measure to another. Native speaker status and cognitive variables (literacy and numeracy) were not significantly related to the DVs.

Dentist visit

When native speaker status alone was entered into the regression model, it was a significant predictor of dental visits. However, when additional variables were entered, into the model, it was no longer significant. The strongest and most consistent positive predictor, again, was having medical insurance. Females, those who were employed, more educated, and healthier, had higher odds of getting a dentist visit. When literacy and numeracy were entered separately, they had a small positive impact on getting a dentist visit; however, when additional variables were entered into the model, their impact was no longer significant.

Vison check

Native speaker status was not a significant predictor of having a vision check. The strongest and most consistent predictor of getting a vision check was having medical insurance. Older individuals, females, healthier individuals, those with higher education, as well as Blacks and those of Other race (compared to Whites) and those out of labor force (compared to employed) had higher odds of having a vision check. When in the model by itself, literacy had a small positive impact on getting a dentist visit, but when in the model with other variables, its impact was not significant. The impact of numeracy was not significant on obtaining a vision check.

Flu shot

When entered separately, native speaker status was not a significant predictor for getting a flu shot. The strongest and most consistent predictor of obtaining a flu shot was having medical insurance. Older individuals, females, those with higher education, as well as Hispanics and those of Other race (compared to Whites) and those unemployed and out of labor force (compared to employed) had higher odds of getting a flu shot. Literacy and numeracy were not significant predictors of getting a flu shot.

Detailed Findings: RQ1

Dependent variable: Dentist visit

1. Are there differences in the utilization of preventive care (dentist visit) between native speakers of English vs non-native speakers of English, controlling for demographic factors (age, gender, ethnicity), socioeconomic factors (income, educational status, employment status), health-related factors (having medical insurance and self-described health status), and cognitive factors (literacy and numeracy)?

The results of several logistic regression models are provided in table 1C. We used the interpretation of logistic regression coefficients and odds ratios outlined in the Methodology section 'Interpretation of logistic regression coefficients and odds ratios' on p.14. The first logistic regression model (LR1) included only one explanatory variable – native speaker status. The results indicated that native speaker status was significantly related to having a dentist visit; the odds of native speakers going to the dentist were 19% higher than the odds of non-native speakers.

In the next step, demographic variables (age, gender, and race/ ethnicity) were added to the model, and the coefficient of the native speaker variable changed sign, which often signifies collinearity or a common confounder with other variables. In an attempt to find the common confounder, we added demographic variables one by one. The native speaker status remained significant, and the coefficient positive, as the demographic variables of age (LR2) and gender (LR3) were added in the model. Age did not have a significant impact on the DV. Gender was significant; the odds of females having a dentist visit were 26% higher than the odds of males. When the race/ethnicity variable was added in the model (LR4 and LR5), native speaker status changed signs, indicating that these variables may be confounding with each other. In an attempt to understand the cause of this sign change, we performed several exploratory analyses. First, crosstabulations between race and native speaker status was performed (see table 3A). While there was no exact confounding (such as, for example, all Blacks being nonnative speakers), there were clearly more Hispanics among nonnative speakers. An additional analysis was run by creating an interaction variable between race and native speaker status and adding it to the regression model. The impact of this variable was not significant.

Since we were most interested in the impact of native speaker status on the DV (having a dentist visit), and there was no clear explanation for the change of sign when race was included, we decided that the race variable should be excluded to keep the information about the variable of primary interest. In addition, two other factors played a role in this decision. The percent of Hispanics among non-native speakers was quite high (59.4%), which could have led to the confounding of race/ethnicity with the non-native status and the resulting sign switch. Also, the addition of race/ethnicity increased Nagelkerke's pseudo R^2 only very slightly (from 0.01 to 0.02). For these reasons we decided to exclude the race/ethnicity variable from the model.

After the addition of demographic variables (gender and age), variables that were proxies for SES (EDUCATION and employment status), were added (LR6). With the addition of these variables, native speaker status was no longer significant, nor was age. Gender remained significant, with the odds of females having a dentist visit 25% higher than the odds of males. For every point increase in the EDUCATION variable, the person's odds of visiting a dentist increased by 39%. Those who were unemployed had 54% higher odds of having a dentist visit than those who were employed. Nagelkerke's pseudo R^2 increased from 0.02 to 0.09 in this model.

In the next model (LR7), health status was added. For every point increase in health status, the person's odds of visiting a dentist increased by 38%. For every point increase in the EDUCATION variable, the person's odds of visiting a dentist increased by 33%. Those who were unemployed had 44% higher odds of having a dentist visit than those who were employed. Females had 27% higher odds of visiting a dentist than males. Age and native speaker status were not significant. Nagelkerke's pseudo R^2 increased from 0.09 to 0.12.
Lastly, having medical insurance was added to the model (LR8). The impact of this variable was significant, with those having medical insurance having 256% (3.56 times) higher odds of visiting a dentist than those who did not. Age and native speaker status were not significant. Nagelkerke's pseudo R^2 increased from 0.12 to 0.18. which was the model with the highest percent of explained variance. As observed earlier, the variable that had the most impact on the DV was having medical insurance. We concluded that this was the optimal model to estimate the effects of different variables on the DV.

1A. Is the relationship between literacy and PHCU (dentist visit) stronger than the relationship between numeracy and PHCU (dentist visit)?

The results of several logistic regression models pertaining to the impact of literacy and numeracy on the DV are provided in table 2C. Logistic regression with only literacy as an explanatory variable was carried out. The effect of Literacy was significant, but its impact and the resulting variance explained was very small (LR1). Similarly, logistic regression with only numeracy as an explanatory variable was carried out. The effect of Numeracy was significant variable, but its impact and the resulting variance explained was also very small (LR2). With both variables in the regression model (LR3), there was no increase in Nagelkerke's R^2 from the model with literacy only. Only literacy was significant in the combined model. Based on these results, we can infer that when the only predictive variables are literacy and numeracy, literacy does have a significant impact on the DV. We also constructed a model in which both literacy and numeracy were in the model, along with all other explanatory variables (LR4). Both literacy and numeracy were not significant, and the coefficient of the native speaker status variable was negative, possibly due to multicollinearity and/or common confounder issues. Also, there was no increase in Nagelkerke's R^2 from the model without these variables. An additional analysis was conducted (LR5) without the health insurance status, which was the strongest predictor for having a vision check. While the native speaker coefficient became positive, the literacy and numeracy variables were not significant, and the Nagelkerje's R^2 went down from 0.18 to 0.12. For these reasons, we concluded that other variables are much better predictors of the DV than literacy and numeracy, and the optimal model should not include either variable as an explanatory variable.

Dependent variable: Vision check

1. Are there differences in the utilization of preventive care (vision check) between native speakers of English vs non-native speakers of English, controlling for demographic factors (age, gender, ethnicity), socioeconomic factors (income, educational status, employment status), health-related factors (having medical

insurance and self-described health status), and cognitive factors (literacy and numeracy)?

The results of several logistic regression models are provided in table 3C. We used the interpretation of logistic regression coefficients and odds ratios outlined in the Methodology section ‘Interpretation of logistic regression coefficients and odds ratios’ on p.14. The first logistic regression model (LR1) included only one explanatory variable – native speaker status. The results indicated that native speaker status was not significantly related to having a vision check.

In the next step, the following demographic variables were added to the model: gender, age, and race/ethnicity. For age, every point increase in the age group, the odds of having a vision check go up 10%. For gender, the odds of females having a vision check were 14% higher than the odds for males. Finally, for race/ethnicity, the odds for Hispanics were 29% lower than the odds for Whites, while the odds of Blacks were 28% higher, and Other race was 58% higher than the odds of Whites for having a vision check.

Next, education and employment variables were added to the model. For every additional increase in education level, individuals’ odds of having a vision check increased by 38%. The odds for the unemployed group were 50% lower than for those that were employed to have a vision check. In the next step, health status was added in the model. For each point increase in health status, the individuals’ odds of having a vision check increased by 14%. In the last step, having medical insurance was added in the model. Those who had medical insurance had 153% (2.53 times) higher odds of having had a vision check.

With all variables entered into the model the results were: for every point increase in the age group, an individual’s odds of having had a vision check were 8% higher. For every point of increase in education, an individual individual’s odds of having had a vision check were 6% higher. For every point increase in the health status, an individual’s odds of having had a vision check were 12% higher. Blacks and those of Other race had higher odds of having had a vision check than Whites. Those out of labor force also had higher odds of having a vision check than those that were employed. This was considered to be the optimal model; Nagelkerke’s pseudo R^2 was the highest for this model at 0.07.

1A. Is the relationship between literacy and PHCU (vision check) stronger than the relationship between numeracy and PHCU (vision check)?

The results of several logistic regression models pertaining to the impact of literacy and numeracy on the DV are provided in table 4C. Logistic regression with only literacy as an explanatory variable was carried out. The Literacy variable was significant, but its impact and the resulting variance explained was very small (LR1). Similarly, logistic regression with only numeracy as an explanatory variable was carried out. Numeracy was not significant. With both variables in the regression model (LR3), there was no increase in Nagelkerke’s R^ from the model with literacy only, and neither variable was significant. We also constructed a model in which both literacy and numeracy were in the model, along with all other explanatory variables (LR4). Neither literacy and numeracy were significant, and there was no increase in Nagelkerke’s R^2 from the model without these variables. For that reason, we concluded that other variables are much better predictors of the DV than literacy and numeracy, and the optimal model should not include either variable as an explanatory variable.

Dependent variable: Flu shot

1. Are there differences in the utilization of preventive care (flu shot) between native speakers of English vs non-native speakers of English, controlling for demographic factors (age, gender, ethnicity), socioeconomic factors (income, educational status, employment status), health-related factors (having medical insurance and self-described health status), and cognitive factors (literacy and numeracy)?

The results of several logistic regression models are provided in table 5C. The first logistic regression model (LR1) included only one explanatory variable – native speaker status. The results indicated that native speaker status was not significantly related to getting a flu shot. In the next step, demographic variables were added to the model (gender, age, and race/ ethnicity). For age, every point increase in the age group, an individuals' odds of getting a flu shot were 25% higher. For gender, the odds of females getting a flu shot were 29% higher than the odds for males. Representatives of 'Other' race had 52% higher odds of getting a flu shot than Whites. In the next step, education and employment variables were added to the model. For every increase in education level, an individuals' odds of getting a flu shot increased by 11%, while those out of labor force had 48% lower odds of getting a flu shot than those who were employed. In the next step, health status was added in the model. It did not have a significant impact on the DV.

In the last step, having medical insurance was added in the model. The odds of getting a flu shot for those who had medical insurance were close to 200% (2.96 times) higher than for those who did not have health insurance. In this model, for every point of increase in the age group, an individual's odds of getting a flu shot were 20% higher. For every point of increase in education, an individual's odds of getting a flu shot increased by 6%. Hispanics and those of Other race had higher odds of getting a flu shot than Whites. The odds of those who were unemployed and out of labor force were also higher than of those who were employed. This was considered to be the optimal model; Nagelkerke's pseudo R^2 was the highest for this model at 0.10.

1A. Is the relationship between literacy and PHCU (flu shot) stronger than the relationship between numeracy and PHCU (flu shot)?

The results of several logistic regression models pertaining to the impact of literacy and numeracy on the DV are provided in table 6C. Logistic regressions with only literacy (LR1) and only numeracy (LR2) as explanatory variables were carried out. Neither variable was significant. With both variables in the regression model (LR3), there was no increase in Nagelkerke's R^ from the model with literacy or numeracy only, and neither variable was significant. We also constructed a model in which both literacy and numeracy were in the model, along with all other explanatory variables (LR4). Neither literacy and numeracy were significant, and there was no increase in Nagelkerke's R^2 from the model without these variables. For that reason, we concluded that other variables are much better predictors of the DV than literacy and numeracy, and the optimal model should not include either variable as an explanatory variable.

Findings: RQ2

For non-native speakers of English, does the age of English language acquisition and self-rated language proficiency have an impact on PHCU, controlling for the demographic, socioeconomic, health-related, and cognitive factors listed above?

Executive Summary: RQ2

One common significant predictor for all three DVs (dentist visit, vision check, and flu shot) was having medical insurance. Some demographic/ socioeconomic/ educational predictors were significant for some DVs, but not for others. Most self-rated language variables and cognitive variables (literacy and numeracy) were not significantly related to the DVs, with the exception of self-rated ability to write in English which was strongly related to obtaining a vision check.

Dentist visit

Among non-native speakers of English, the strongest predictor of visiting the dentist was having medical insurance. Within this group, females had higher odds of getting a dental visit than males. Non-native speaking Blacks had lower odds of seeing a dentist than non-native speaking Whites. Those with higher education and those who acquired English at an older age (compared to those who acquired English before the age of 16) had higher odds of visiting a dentist. When in the model by themselves, self-rated language skills were positively and significantly related to obtaining a dentist visit. However, once other predictors were added in the model, they were not significant. Literacy and numeracy were not significantly related to the DV.

Vision check

The strongest predictor of having a vision check was having medical insurance. For this group, older individuals had higher odds of getting a vision check. Non-native speakers of Other race had higher odds of getting a vision check than non-native speaking Whites. Those out of the labor force had higher odds of getting a vision check than employed people. Age of language acquisition was not significantly related to getting a vision check. When in the model by themselves, all self-rated language skills were positively and significantly related to obtaining a dentist visit. However, once other predictors were added in the model, only self-rated ability to write in English remained positively and significantly related to getting a vision check. Literacy and numeracy were not significantly related to the DV.

Flu shot

The strongest predictor of obtaining a flu shot was having medical insurance. Those out of labor force had higher odds of getting a flu shot than employed people. Age of language acquisition was not a significant predictor of obtaining a flu shot. When in the model by themselves, none of the self-rated language variables except for the ability to read in English were significant. Once other predictors were added in the model, none of the self-rated language variables were significant. Literacy and numeracy were not significantly related to the DV.

Detailed Findings: RQ2

Dependent variable: dentist visit

The results of several logistic regression models are provided in table 7C. When the age of language acquisition was entered in the model by itself (LR1), it did not have a significant impact on the DV. When the self-rated language proficiency variables were entered in the model by themselves, their impact was significant. For every point increase in the ability to speak English (LR2), individuals' odds of visiting the dentist went up by 42%; for every point of increase in the ability to read in English (LR3), by 51%; for every point increase in the ability to write in English (LR4), by 47%, and for every point increase in the ability to understand English (LR5), by 45%. When age of language acquisition and self-rated language variables were entered in the model together (LR6), those who acquired English before the age of 16 had 75% lower odds of visiting a dentist compared to those who acquired English at the age of 16 or later; in this model, self-rated language variables were not significant.

In the next step, demographic variables (gender, age, race/ ethnicity) were added (LR7), and self-rated language variables were removed since they were not significant in the previous model. With these variables entered into the model, age of acquisition was no longer significant. For gender, males' odds of visiting a dentist were 62% lower than females', while for race/ethnicity, the odds for Hispanics' and Blacks' were lower than for Whites'.

In the next step, SES proxy variables, education level and employment, were added, (LR8). While the employment variable was not significant, for every additional increase in education level, an individuals' odds of visiting a dentist increased by 28%. Health status was added in the next step (LR9). For every additional point in health status, the individuals' odds of visiting a dentist increased by 21%. Having insurance was added in the next step (LR10). The odds of visiting a dentist for those who had health insurance were 4 times higher than for those without health insurance. We also tested models that had either all self-rated language variables (LR11) or both literacy and numeracy (LR12) included. All of these variables were not significant, and while the Nagelkerke's pseudo R^2 increased from LR10, it remained the same for both LR11 and LR12.

Three additional models were tested – one with all demographics, SES proxy, health status and health insurance variables, and age of acquisition variable, and both cognitive variables and all self-rated language variables (LR13), only including literacy (LR14), and only including numeracy (LR15). In these models, the cognitive and self-rated language variables were not significant. The Nagelkerke's pseudo R^2 increased slightly from model LR12, but remained the same for all three of these models. For all three models, males, Blacks (as compared to Whites), and those who acquired English before the age of 16 (as compared to those who acquired English at 16 or older) had lower odds of visiting a dentist. Those with higher education levels had higher odds of visiting a dentist. Those who had health insurance had higher odds of visiting a dentist than those who did not.

Dependent variable: vision check

The results of several logistic regression models are provided in table 8C. When the age of language acquisition was entered in the model by itself (LR1), it did not have a significant impact on the DV. When self-rated language proficiency variables were entered in the model by themselves, their impact was significant. For every point increase in the ability to speak English (LR2), individuals' odds of having a vision check went up by 42%; for every point increase in the ability to read in English (LR3), by 47%; for every point of increase in the ability to write in English (LR4), by 47%; and for every point increase in the ability to understand English (LR5), also by 47%. When age of language acquisition and self-rated language variables were entered in the model together (LR6), for every point increase in self-rated ability to write, an individual's odds of having a vision check increased by 45%. Age of English language acquisition was not significant.

In the next step, demographic variables (gender, age, race/ ethnicity) were added (LR7), and self-rated language variables were removed since they were not significant in the previous model. For every point increase in age, a person's odds of having a vision check increased by 24%. Gender was not significant. The odds for those of 'Other' race were twice as high as those of Whites.

In the next step, SES proxy variables were added, education level and employment (LR8). While the education variable was not significant, the odds of having a vision check for those out of labor force were 78% higher than those who are employed. Health status was added in the next step (LR9). It was not significant. Having insurance was added in the next step (LR10). The odds of having a vision check for those who had health insurance were 2.77 times higher than for the uninsured.

We also tested models that had either all self-rated language variables (LR11) or both literacy and numeracy (LR12) included. The literacy and numeracy variables were not significant; however, self-rated ability to write in English was significant, with every point increase resulting in an increase in the odds of having a vision check by 55%. This model (LR11) had a very slightly higher Nagelkerke's R^2 than model LR12.

Three additional models were tested – one with all demographics, SES proxy, health status and health insurance variables, and age of acquisition variable, and both cognitive variables and all self-rated language variables (LR13), a model only including literacy (LR14), and a model only including numeracy (LR15). In these models, the cognitive variables were not significant. However, self-rated writing ability was significantly and positively related to the odds of having a vision check.

The Nagelkerke's pseudo R^2 increased slightly from model LR12, but remained the same for all three of these models. For all three models, older individuals, those of Other race, those out of labor force, and those with medical insurance had higher odds of having a vision check. The three models resulted in a slightly higher Nagelkerke's R^2 value than the previous models (LR11 and LR12). These three models had the same Nagelkerke's R^2 values of 0.15.

Dependent variable: flu shot

The results of several logistic regression models are provided in table 9C. When the age of language acquisition was entered in the model by itself (LR1), it did not have a significant impact on the DV. When self-rated language proficiency variables were entered in the model by themselves, only the impact of reading was significant. For every point of increase in the ability to read in English (LR3), the odds of getting a flu shot increased by 22%. Other self-rated language variables in models LR2 and LR4 and LR5 were not significant. When age of language acquisition and self-rated language variables were entered in the model together (LR6), their impact was not significant.

In the next step, demographic variables (gender, age, race/ ethnicity) were added (LR7), and self-rated language variables were removed since they were not significant in the previous model. Males' odds of getting a flu shot were 37% lower than females'. The odds of those of 'Other' race were 72% higher than Whites'. In the next step, SES proxy variables were added, education level and employment (LR8). While education variable was not significant, the odds of having a vision check for those out of labor force were 82% higher than for the employed. Health status was added in the next step (LR9). It was not a significant variable. Having insurance was added in the next step (LR10). The odds of getting a vision check for those who had health insurance were 3.45 times higher than for the uninsured. We also tested out models that had either all self-rated language variables (LR11) or both literacy and numeracy (LR12) included. These variables were not significant in either model.

Three additional models were tested – one with all demographics, SES proxy, health status and health insurance variables, and age of acquisition variable, and both cognitive variables and all self-rated language variables (LR13), a model including only literacy (LR14), and a model including only numeracy (LR15). In LR13, the cognitive and language variables were not significant. In models LR14 and LR15 literacy and numeracy, respectively, had a significant, but very small and negative impact on the odds of obtaining a flu shot. The self-rated language variables were not significant in models LR14 and LR15. The Nagelkerke's pseudo R^2 increased slightly from model LR12 to models LR13 and LR15, and remained the same for model LR14. Those who had medical insurance (compared to the uninsured), were out of labor force (compared to employed), and were of Other race (compared to Whites) had higher odds of getting a flu shot in for these models.

DISCUSSION

Dental preventive care

When considered by itself, but also with the inclusion of the demographic variables age and gender, native speaker status significantly and consistently impacted the use of preventive dental care for the general population. Non-native speakers had lower odds of using preventive dental care compared to native speakers. However, once socioeconomic and health-related variables were added, the effect of native speaker status was nonsignificant. Therefore, we conclude that controlling for socioeconomic and health-related factors, the native speaker status was not significantly related to the use of dental preventive care.

Similarly, literacy and numeracy were significant predictors of dental preventive care utilization when in the model by themselves. When both literacy and numeracy were in the model, but those were the only two explanatory variables, literacy was significantly related to dental care, but numeracy was not. However, once other demographic, socioeconomic, and health-related variables were entered in the model, the effect of literacy and numeracy were nonsignificant. Therefore, we conclude that controlling for socioeconomic and health-related factors, the cognitive factors (literacy and numeracy) were not significantly related to the use of dental preventive care.

The finding related to higher odds of having a dentist visit for females and for individuals with higher education is supported by previous research. The finding that individuals who rate themselves as being healthier are more likely to have preventive care is also supported by literature. However, the direction of causality between higher health status and preventive care utilization is not clear, because people may report being healthier because of using preventive care, or they may report using less healthcare in general if they are healthier.

One counterintuitive finding is that those who were unemployed were more likely to see a dentist. One interpretation of this finding may be that the responders interpreted the question about seeing the dentist as a visit for sick care as opposed to preventive care, and thus reported higher utilization of care.

For the non-native speaker group only, when in the model by itself, age of English language acquisition was not a significant predictor of dental preventive care use. However, all four self-rated language skills (the ability to speak, read, write, and understand English) were significantly and positively related to the use of preventive dental care. Once demographic, socioeconomic, and health-related variables were entered in the model, the age of English acquisition became significant, with those acquiring English at an earlier age having lower odds of using dental preventive care compared to those who acquired English at an older age. It is likely that several factors are at play in the explanation of this outcome. Individuals who acquired English earlier in life are likely to be more proficient in it. They are more likely to get better jobs with health/ dental insurance available, and they are more likely to be aware of the services available to maintain good dental health. If these individuals were also interpreting having a dental visit as sick care, not preventive care, they may have reported lower utilization of dental care. In this model, self-rated language skills became nonsignificant. Therefore, we conclude that controlling for demographic, socioeconomic, and health-related variables, only the age of English acquisition, but not the specific language skills, was a significant predictor of dental preventive care use.

Vision preventive care

For the general population, there were no differences in the utilization of preventive vision care by native or non-native speakers either when the native speaker status was in the model by itself or with other variables. Literacy and numeracy were also not found to be significant predictors of the utilization of preventive vision care. Therefore, we conclude that controlling for socioeconomic and health-related factors, the native speaker status and cognitive variables were not significantly related to the use of preventive vision care.

The finding that utilization of preventive vision care increases with age is supported by prior research. Similarly, those with higher education usually report higher preventive care utilization. Individuals of Black and Other race reported higher utilization of vision care, which is not supported by some studies. This finding can possibly be explained by the responders' interpretation of the question as relating to sick care rather than preventive care.

Looking at the non-native speaker group only, age of English language acquisition was not a significant predictor of vision preventive care use. However, all four self-rated language skills (the ability to speak, read, write, and understand English) were significantly and positively related to the use of preventive vision care. Once demographic, socioeconomic, and health-related variables were entered in the model, the effects of age of English acquisition and all self-rated language skills except writing were nonsignificant. Therefore, we conclude that controlling for demographic, socioeconomic, and health-related variables, the age of English acquisition and all specific language skills except writing were not significant predictors of vision preventive care use.

Getting a flu shot

For the general population, there were no differences in getting a flu shot by native or non-native speakers either when the native speaker status was in the model by itself or with other variables. When literacy or numeracy were in the model separately, or when they were in the model together, they were not significant predictors getting a flu shot. Therefore, we conclude that controlling for socioeconomic and health-related factors, the native speaker status and cognitive variables were not significantly related to getting a flu shot.

The finding that older people are more likely to get a flu shot makes sense. People over 65 are encouraged to get flu shots in the media and by health care providers. The finding that responders of Hispanic and Other race were more likely to get a flu shot may not be supported by other studies. It may be possible that flu shots are more heavily marketed to these populations, or they are more likely to say that they received a flu shot.

For the non-native speaker group only, age of English language acquisition and all but one specific language skill (reading) were not significant predictors of getting a flu shot. Once demographic, socioeconomic, and health-related variables were entered in the model, the effects of age of English acquisition and all self-rated language skills were nonsignificant. Therefore, we conclude that controlling for demographic, socioeconomic, and health-related variables, the age of English acquisition and the specific language skills were not significant predictors of getting a flu shot.

General discussion

Preventive health care utilization is impacted by numerous social, economic, geographical, demographic, medical, and other factors. This study attempted to examine the impact of only a handful of variables on healthcare utilization. The factors that were our primary points of interest were native speaker status, self-reported language skills variables, and literacy and numeracy. The impact of native speaker status was significant for obtaining dental care when considered by itself. Once the demographic, socio-economic, and cognitive factors were controlled for, the effect of native speaker status is not significant. The cognitive factors of literacy and numeracy sometimes had a statistically significant but trivial impact when they were in the model by themselves, but once other factors were controlled for, the effect of these variable were no longer significant. Similarly, when the self-rated language skills were only variables in the model, their impact was significant, but once other socio-economic and demographic factors were controlled for, their effect was non-significant.

For the general population, the most consistent factor impacting the use of all three examples of preventive healthcare was having medical health insurance. The responders were not directly asked about having dental or vision insurance. While medical health insurance does not pay for dental care, it is reasonable to assume that those who have medical insurance are more likely to have dental insurance as well, particularly employer-subsidized insurance. The impact of other socio-economic and demographic factors, while present, was less consistent among the types of preventive care utilization. For example, the positive impact of age on PHCU was only present for vision checks and flu shots. The impact of gender was only present for the utilization of dental care and flu shot. Education level, while positive across all three instances of preventive care, had different degrees of impact. Individuals of Black and 'Other' ethnicity had higher odds of obtaining preventive vision care than Whites, but Hispanics and individuals of 'Other' race had higher odds than Whites of obtaining a flu shot. Please see table 10C in the appendix for the B coefficients that served as the basis for the calculation of odds ratios listed in table 3.

Table 3. Weighted logistic regression odds ratios for final selection models for all three dependent variables.

Variable	Dentist visit	Vision check	Flu shot
	Odds ratio	Odds ratio	Odds ratio
Native speaker (referent=nonnative speaker)	NS	NS	NS
Age	NS	1.07	1.21
Gender (referent=female)	0.80	NS	0.82
Race: Hispanic (referent=White)		NS	1.55
Race: Black (referent=White)		1.45	NS
Race: Other (referent=White)		1.58	1.49
Education	1.27	1.05	1.06
Employment: unemployed (referent=employed)	0.82	NS	1.23
Employment: out of labor force (referent=employed)	NS	1.32	1.43
Health status	1.36	1.12	NS
Have medical insurance (referent=do not have insurance)	3.56	2.53	2.97
Nagelkerke's R^2	0.18	0.07	0.1

For the non-native speaker group, similarly, the most consistent factor impacting the likelihood of engaging in all three examples of preventive healthcare was having medical health insurance. The impact of other socio-economic and demographic factors, while present, was less consistent between the types of preventive care utilization. For example, the positive impact of age on PHCU was only present for vision checks, while the impact of gender was only present for the utilization of dental care. Blacks had lower odds of having a dentist visit than Whites, and individuals of ‘Other’ race had higher odds of having a vision checkup than Whites. Higher education and higher age of English language acquisition only increased the odds of visiting a dentist. Of the self-reported language skills, only writing was significant, and only for the vision check model. Please see table 11C in the appendix for the B coefficients that served as the basis for the calculation of odds ratios listed in table 4.

Table 4. Weighted logistic regression odds ratios for final selection models for all three dependent variables with the sample of non-native English speakers.

Variable	Dentist visit	Vision check	Flu shot
	Odds ratio	Odds ratio	Odds ratio
Age	NS	1.20	NS
Gender (referent=female)	0.62	NS	NS
Race: Hispanic (referent=white)	NS	NS	NS
Race: Black (referent=white)	0.42	NS	NS
Race: Other (referent=white)	NS	2.41	NS
Education	1.15	NS	NS
Employment: unemployed (referent=employed)	NS	NS	NS
Employment: out of labor force (referent=employed)	NS	1.77	1.63
Health status	NS	NS	NS
Have medical insurance (referent=do not have insurance)	4.01	2.66	3.60
Numeracy	NS	NS	NS
Literacy	NS	NS	NS
Age of English language acquisition (referent=learned English when age 16 or older)	0.52	NS	NS
Ability to speak English	NS	NS	NS
Ability to read English	NS	NS	NS
Ability to write English	NS	1.57	NS
Ability to understand English	NS	NS	NS
Nagelkerke’s R^2	0.22	0.15	0.14

Note: unweighted sample includes non-native speakers of English only.
Significance level of significant variables is $p < 0.05$. All estimates are significant unless noted with NS (not significant).

These variations may be an artifact of population sampling. At the same time, they may reflect more complex causes of and reasons for why individuals may seek a specific type of health care. For example, it is possible that a vision check was interpreted by some responders as a means of secondary or tertiary, rather than primary, prevention (that is, an action taken to prevent an illness from getting exacerbated, rather than preventing an illness from occurring). It is possible to speculate that since the incidence of diabetes is higher in the Black population, and one of the complications of diabetes is macular degeneration (a condition that affects eyesight), they may be getting vision care for secondary and tertiary prevention, rather than primary prevention, and therefore more frequently report having a vision check.

The amount of variance explained also varies considerably from the model explaining the odds of receiving dental care to the odds of receiving a vision check or a flu shot. The variance explained by the variables in the model for dental care for both the total population and the non-native group only has the highest R^2, indicating that, for these groups, the dental care model explains more variance than the model for vision check and flu shot. This may indicate that there may be fewer factors that impact dental care utilization, each with a stronger impact, whereas vision care and getting a flu shot may be impacted by many more factors, each with a smaller impact. For example, it is possible that a person will only consider getting preventive dental care if they have dental insurance, due to the higher costs of dental care. A flu shot is more easily available, even in remote and poor communities and can in some instances have no cost. It can even be obtained at a walk-in clinic during flexible hours. For that reason, such factors as education or employment may have less of an impact on obtaining a flu shot.

For the nonnative speaker population, similar to the overall population, the one most consistent predictor of obtaining preventive care was having medical insurance (Pylypchuk & Hudson, 2009; Yang & Hwang, 2016). This finding is consistent with previous research studies. It was interesting that the self-rated ability to write was positively related to receiving preventive vision care. Only specific types of vision care (e.g., a visit to an ophthalmologist, but not to an optometrist) are covered by medical insurance. Usually, an individual has to purchase vision insurance separately, and it is often employer-subsidized. It is possible that only those at specific levels of income or in specific professions that use writing more often are more likely to purchase vision insurance. But it is also possible that this relationship between writing proficiency and the use of preventive vision care is spurious.

One major difference between our findings and the findings of previous studies (Chang, Chan, & Han, 2015; Lebrun, 2012; Ponce et al., 2006; DeRose, Escarce, & Lurie, 2007; Jacobs, Chen, Karliner, Agger-Gupta, & Mutha, 2006) is that the native speaker status may not always have a significant impact on preventive care utilization. There may be several explanations of this finding. One, different variables were included in our study compared to previous studies of similar issues. Two, studies that incorporate medical claims data may be able to evaluate preventive care utilization more accurately than studies that only use self-reported preventive care use.

Another difference is the finding that for non-native speakers, age of English language acquisition and specific language skills were not consistent predictors of preventive care use. This effect may be occurring for similar reasons as the unexpected findings for the population as a whole. Studies of non-native speakers may be controlling for some demographic variables, but not for socioeconomic or health-related variables. In addition, self-reported data may not be completely reflective of the true preventive care utilization.

Our findings do not necessarily indicate that there truly are no differences in the use of preventive care between native and non-native speakers. Especially for dental preventive care, they only indicate that if demographic, socioeconomic, and health-related factors were held equal, there would be no differences in preventive care utilization. Similarly, for the non-native speaker group, the findings indicate that if all other factors were held equal, including the age of

English language acquisition and specific language skills, there would be no differences in preventive care utilization. Access to health insurance is a critical factor in the access to preventive health care, but specific language skills, education, and employment are also important modifiable factors.

One reason for the inconsistencies in the significance of different variables from model to model may be the fact that these variables are interrelated, with the direction of causality of the impact of one variable on another being challenging to determine. For example, one causal pathway may be that sufficient knowledge of English allows a non-native speaker to obtain better education, which in turn leads to a better job and having health insurance, which in turn leads to the use of preventive health care. Suppose this person was a male and had sufficient time to devote to studying, instead of raising children and taking care of the elderly parents. In this case, the 'primary' cause of having health insurance is language proficiency, but the most direct cause is having health insurance. But ultimately, all these variables have some impact on each other. The issue of determining causality comes to the foreground when attempts are made to institute policies to increase PHCU by specific population groups. Sometimes addressing the most direct causes of an issue does not eliminate the 'true', more far-reaching disparities. At the same time, the most direct causes may be more manifest and easier to address than the more far-reaching causes of disparities.

Overall, it is important to remember that the immigrant population and non-native speakers are not a homogeneous group. They may vary significantly on all the variables we examined and controlled for. This is another reason that the addition of different variables into the models may be resulting in results that are not always consistent.

POLICY IMPLICATIONS

Health insurance

Since having insurance is a strong predictor of PHCU, policies are needed to expand access to health insurance and to community clinics and other care venues. Access can be addressed to a large extent by reauthorizing and increasing funding for programs such as SCHIP, which has been effective in maintaining coverage among immigrant children and their families post-PRWORA (Personal Responsibility and Work Opportunity Reconciliation Act) (DeRose, Escarce, & Lurie, 2007). Also useful would be increased funding for community health centers and expanded outreach efforts through trusted sources such as community-based organizations whose culturally competent staff can help immigrants navigate the often overly complex application procedures for Medicaid and SCHIP (DeRose, Escarce, & Lurie, 2007). Policies might also aim to increase employer-based health insurance among immigrants.

Flavin et al. (2018) mention that in 1996, the PRWORA barred legal immigrants from obtaining nonemergency Medicaid. The 2010 Affordable Care Act (ACA) denied legal immigrants access to its health insurance plans until they had completed 5 years of lawful residence and denied undocumented immigrants access to plans, although it otherwise increased health insurance coverage to many low- and moderate-income individuals. These restrictions and denials have produced substantial negative health consequences for immigrant populations in the United States.

According to Parmet (2018), even under the ACA, noncitizen immigrants are far less likely than citizens to have health insurance. In part, this is because immigrants are poorer than the native-born population and are less likely to receive insurance through their workplace. In addition, those young immigrants who received work permits under the Deferred Action for Childhood Arrival (DACA) program are not able to purchase their insurance on the exchange through ACA. In addition, the ACA kept in place the 1996 Personal Responsibility and Work Opportunity Reconciliation Act's (PRWORA's) tight restrictions on immigrants' eligibility for Medicaid and CHIP. They are not eligible for either until they have been in the US for 5 years. As a result, by 2015, 7 million of the 33 million Americans without insurance were noncitizen immigrants. Individual states attempt to fill in the gap by providing their own state-funded insurance for adults and children, but that coverage can be inconsistent.

Although the ACA leaves many immigrants uninsured, it does provide relief to some lawful immigrants. First, it permits lawfully present immigrants who are ineligible for Medicaid to purchase insurance on the exchanges, even if their incomes are below the threshold required for citizens to participate on the exchanges. Second, because immigrants have lower wages than native born citizens, those who are not barred from Medicaid or CHIP due to PROWRA are more likely than native-born citizens to benefit from the ACA's Medicaid expansion. More generally, because of their lower wages, immigrants benefit disproportionately from the ACA's progressivity.

Under the proposed new legislation, American HealthCare Act (AHCA), states will be denied the funds they rely on to maintain the Medicaid expansion. AHCA would also replace the ACA's income-based support with tax credits that are based on age, which will make it very difficult for people with low incomes to afford insurance. Some AHCA provisions specifically target immigrants. For example, one provision limits eligibility for tax credits to non-citizens who are "qualified" within the meaning of PRWORA. Because PRWORA's definition of "qualified" is narrower than the ACA's definition of "lawfully present," some immigrants who are able to buy insurance on the exchanges would now be barred.

Addressing limited English language proficiency

Because issues related to English proficiency remain an important source of vulnerability among immigrants, it is beneficial to work on a broader implementation and enforcement of Culturally and Linguistically Appropriate Services (CLAS) standards and the expansion of Medicaid benefits to cover interpreter services. Further, researchers and community activists advocate for the use of Medicaid funds for community-based health promoters that can reach underserved populations, especially immigrants. Finally, since language-concordant providers are generally preferable to interpreters, long-term solutions include investing in medical education of bilingual people or offering financial incentives such as additional pay for bilingual staff members (DeRose, Escarce, & Lurie, 2007).

Ku (2009) stresses the importance of the role of public and private insurers in improving immigrants' access to care. Insurance providers could take steps to reduce language barriers by paying for interpretation or other language services for patients with limited English proficiency. Although federal civil rights policy already requires health care providers to offer free interpretation or language assistance to LEP patients, a primary stumbling block is that insurance usually does not pay for language services, which means that the provider has no incentive to provide these services. In addition, insurers—particularly public payers—could make efforts to increase the supply of providers, particularly primary care clinicians and safety-net providers (those practicing in community health centers), who practice in areas with higher concentrations of immigrants, by providing them with various incentives.

Ye, Mack, Fry-Johnson, and Parker (2012) suggest that in addition to improving overall health insurance coverage, it is important to create more culturally sensitive health education materials about chronic diseases, medications, available health services, and modes of access. There is also an urgent need for more culturally sensitive health services that are acceptable to the immigrant community, probably requiring time and energy being directed to obtaining input and guidance from this particular community and patient population. Linguistically and culturally appropriate services, as recommended by the National Standards for Culturally and Linguistically Appropriate Services in Health and Health Care, can help make health care more accessible and understandable for recent immigrants and adults with limited English proficiency (Chang, Chan, & Han, 2015).

Prins and Monnat (2015) mention that providing ESL instruction for non-native English speakers and efforts to increase college attainment and access to health insurance for low-SES groups, who are most likely to suffer from poor health, are promising areas of policy intervention.

Feinberg, Greenberg, and Frijters (2016) suggest that different stakeholders have unique roles to play in increasing health information seeking behaviors. For example, health professionals could develop more focus and skills in oral participatory care with patients and caregivers; the health care industry (pharmaceutical, hospital systems, insurance companies) could simplify and standardize forms and written materials; policy makers could include health literacy at appropriate funding levels for K-12 and Adult Basic Education curricula; and health educators and researchers must develop interventions to address health information seeking behaviors through differing skill levels in multiple modalities. Dallo and Kindratt (2015) suggest that immigrants may simply not be aware of the recommended services. Therefore, such educational forums may be useful in increasing screening estimates for other preventive behaviors, such as flu and pneumonia vaccines.

A number of previous studies suggested possible interventions to increase healthcare utilization, including preventive care utilization, by the immigrant population. They include matching patients with limited language proficiency with providers who speak their primary language and/or finding an appropriate interpreter (Jacobs et al., 2006); creating culturally sensitive education materials for the immigrant population and providing more culturally sensitive care in general (Ye, Mack, Fry-Johnson, Parker 2012) as well as policy changes, such as expanding health insurance coverage to allow inclusion for the immigrant population; staffing community health centers with providers with interpreters or bilingual providers; having insurance pay for interpreters (DeRose, Escarce, & Lurie, 2007).

Safety nets for healthcare

Many of the new destinations for immigrants are in states with more-restrictive Medicaid policies, fewer interpreters and language-concordant providers, and weaker public health systems. Because of these disparities, it has been suggested that additional support is needed to increase the long-term viability of the safety net in these areas. Many see the federal government as making an impact by improving services for this population, whether by strengthening the safety net (such as building and staffing more community health centers) and public health systems, by developing health-promoting capabilities and resources within immigrant communities, or by expanding insurance coverage to enable immigrants to make better use of health care services.

CONTRIBUTIONS TO RESEARCH

The study contributed to existing research by examining the factors that impact preventive healthcare utilization for the general population, and for non-native speakers of English in the US. It also looked at the relationship between literacy and numeracy and preventive care utilization for the general US population, as well as for non-native speakers of English in the US. While some of these factors have been considered in previous studies, previous research often explored the native speaker status by itself, without considering socioeconomic, health-related, and cognitive factors. Thus, it may be hard to understand what it is specifically about the native speaker status that makes the utilization of preventive healthcare different for native speakers vs nonnative speakers. Putting all these factors in the same model allows for potential insights about the possible sources of these disparities, when they are present.

We found that the factors that impact the utilization of preventive healthcare are not completely different for the population as a whole and for non-native speakers. We confirmed that having health insurance is a very strong factor in determining the utilization of preventive healthcare, both for the general population and for the non-native speaker population. Overall, demographic and socioeconomic factors have a greater impact on preventive healthcare utilization than cognitive and linguistic factors. Overall, the study added to nationally representative findings about the relationship between preventive healthcare utilization, literacy, numeracy, and a number of demographic and socioeconomic factors.

LIMITATIONS

PIAAC data contain multiple variables that could possibly be included in regression models. One factor preventing us from having too many variables in the model is a limited sample size, especially once the sample is divided into groups, for example, based on demographic variables. Omitting from the models the factors that could potentially be significant in explaining healthcare utilization is a limitation of this study. On the other hand, it is also possible that variables that simply were not measured by the PIAAC questionnaire could also be significant in explaining healthcare utilization, and therefore could not be included in the model.
Another limitation was the fact that we had to use self-rated health status and language skills. Those ratings were not confirmed by more objective data. Additionally, all information on education level, employment status, having health insurance, and even obtaining preventive care

were also self-reported. Answers to some of those questions may be subject to a bias to provide more socially desirable information, and thus the accuracy of the data is not known. For example, responders may be reporting that they had a flu shot because that is the expected answer.

Another limitation is that while the questions about visiting a dentist or a vision checkup were intended to be questions about preventive healthcare, it is not clear if the responders interpreted them in that way. Only the question about the flu shot ("In the past year, have you gotten a flu shot?") is clearly a preventive measure. The question about the dentist visit, at least in the US English version of the background questionnaire, is worded as "In the past year, have you visited a dentist?" The question about the vision check in the US English version is worded as 'In the past year, have you had your vision checked?" The responders may have interpreted the questions as asking about a sick visit, rather than a well or preventive care visit. This made the relationship between health status and receiving these interventions less than straightforward: was good health the result of a preventive care intervention, or was poor health the cause of needing a healthcare intervention? We proceeded on the assumption of the former relationship, but the latter relationship is possible as well. Similarly, the relationship between healthcare utilization and employment status was less than straightforward. It is possible that those who were on disability and therefore out of labor force were more likely to use healthcare because of existing medical conditions, rather than to prevent them.

Another limitation is our lack of knowledge about the responders' access to dental and vision insurance. In the US, medical health insurance will most likely pay for a flu shot, but most often not for vision and dental care. The fact that having health insurance was a significant determinant of having had a dental checkup and a vision checkup may allow an assumption that if a person has medical health insurance, they are at least more likely to have dental and vision insurance. However, that is not always the case.

A limitation of using regression models is that regression coefficients can change, sometimes significantly, based on the inclusion or exclusion of variables in the model, and even the order of entry of those variables. In addition, many variables are not completely independent of each other. We tried to account for these dependencies and prevent collinear variables from entering the models, but it is clear that some variable interdependencies still exist. Structural equation modeling may be a better approach to account for these dependencies and different causality pathways.

FURTHER RESEARCH

Further research should include a developing a better understanding of the relationship between literacy, numeracy, and preventive healthcare utilization. Even though our study did not find a significant relationship between these variables, it is possible to look at the relationship between them from a different perspective. For example, an attempt can be made to identify literacy and numeracy thresholds above or below which healthcare utilization differs. This relationship may be explored for different population groups, including non-native speakers. In addition, different and more objective instruments of language skills can be used to assess language proficiency and how it affects healthcare utilization.

The immigrant population is not a homogeneous group, and many factors may be at play in affecting the utilization of preventive healthcare by this population, including demographic and cultural variables that are not examined in this study. Health-seeking behaviors of immigrants from different regions are not very well researched or understood. Further areas of research may be looking at preventive care utilization and factors that impact it by the region from which immigrants arrived. PIAAC data may be useful in comparing what factors impact preventive care utilization by immigrants from different countries.

Appendix A. Descriptive Statistics and correlations.

Table 1A. Continuous variable weighted means for the whole sample.

Variable	Value Range	Mean	Mean SE	S.D.	SD SE
Age	1-6	3.31	0.01	1.58	0.00
Education	1-6	3.05	0.01	1.64	0.01
Health status	1-5	3.58	0.02	1.08	0.01
Literacy	0-500	269.98	0.97	50.83	0.78
Numeracy	0-500	255.33	1.04	56.01	0.79
Ability to understand English	1-4	3.85	0.01	0.47	0.02
Ability to speak English	1-4	3.79	0.01	0.54	0.02
Ability to read English	1-4	3.78	0.01	0.58	0.02
Ability to write English	1-4	3.71	0.01	0.66	0.02

Note: Age has 6 levels (24 or less; 25-34, 35-44, 45-54, 55-65, 66 and older); Education has 6 levels (Lower secondary or less; Upper secondary; Post-secondary, non-tertiary; Tertiary - professional degree; Tertiary - bachelor degree; and Tertiary - master/research degree); Health status has 5 levels (poor, fair, good, very good, excellent); language skills have 4 levels (not at all, not well, well, very well). Literacy and numeracy are continuous variables.

Table 2A. Continuous variable weighted means for native and non-native speaker of English sample.

Variable (range)	Native speaker	Mean	Mean SE	S.D.	SD SE	Mean difference	t-value
Age (1-6)	0	3.04	0.04	1.48	0.02	-0.33	**-7.45**
	1	3.36	0.01	1.59	0.00		
Education (1-6)	0	2.87	0.06	1.78	0.03	-0.21	**-3.15**
	1	3.09	0.01	1.61	0.01		
Health status (1-5)	0	3.54	0.04	1.10	0.03	-0.05	-1.31
	1	3.59	0.02	1.08	0.01		

Literacy (0-500)	0	241.04	2.96	2.96	1.79	-34.70	**-10.81**
	1	275.74	0.98	0.98	0.61		
Numeracy (0-500)	0	231.30	3.10	3.10	1.79	-28.82	**-8.40**
	1	260.12	1.12	1.12	0.74		
Ability to understand English (1-4)	0	3.33	0.04	0.88	0.03	-0.63	**-14.01**
	1	3.96	0.00	0.22	0.01		
Ability to speak English (1-4)	0	3.16	0.05	0.93	0.04	-0.76	**-14.67**
	1	3.92	0.01	0.28	0.01		
Ability to read English (1-4)	0	3.20	0.05	0.97	0.03	-0.70	**-14.98**
	1	3.90	0.01	0.37	0.01		
Ability to write English (1-4)	0	3.02	0.04	1.04	0.02	-0.83	**-18.16**
	1	3.85	0.01	0.43	0.01		

Note: $t>1.96$ indicates significant difference at $p<0.05$. Significant t values are in bold.
Note: Age has 6 levels (24 or less; 25-34, 35-44, 45-54, 55-65, 66 and older); Education has 6 levels (Lower secondary or less; Upper secondary; Post-secondary, non-tertiary; Tertiary - professional degree; Tertiary - bachelor degree; and Tertiary - master/research degree); Health status has 5 levels (poor, fair, good, very good, excellent); language skills have 4 levels (not at all, not well, well, very well). Literacy and numeracy are continuous variables.

Table 3A. Weighted percent of responder sample at each level of variable

Variable		Non-native speaker		Native speaker		All	
	Level	Percent	Percent SE	Percent	Percent SE	Percent	Percent SE
Age	24 or less	18.89	1.04	16.4	0.3	16.82	0.26
	25-34	20.76	0.92	18.04	0.29	18.49	0.23
	35-44	22.26	1.14	16.95	0.28	17.83	0.28
	45-54	19.34	1.03	19.59	0.34	19.55	0.31
	55-65	13.24	0.66	19.5	0.23	18.46	0.21
	66 plus	5.51	0.57	9.52	0.1	8.86	0.07
Employment	Employed	69.6	1.77	69.82	0.65	69.78	0.62
	Unemployed	6.84	0.61	6.33	0.16	6.41	0.13

	Out of the labor force	23.44	1.64	23.83	0.68	23.76	0.64
Education	Lower secondary or less (ISCED 1,2, 3C short or less)	25.54	1.2	11.71	0.23	14.01	0.23
	Upper secondary (ISCED 3A-B, C long)	35.68	1.42	42.11	0.54	41.04	0.42
	Post-secondary, non-tertiary (ISCED 4A-B-C)	5.01	0.75	9.63	0.47	8.86	0.41
	Tertiary - professional degree (ISCED 5B)	6.85	1.0	9.09	0.41	8.72	0.41
	Tertiary - bachelor degree (ISCED 5A)	13.61	0.92	17.31	0.44	16.7	0.44
	Tertiary - master/research degree (ISCED 5A/6)	13.3	1.34	10.14	0.46	10.67	0.44
Gender	Male	48.38	1.25	48.8	0.25	48.73	0.18
	Female	51.62	1.25	51.2	0.25	51.27	0.18
Health status	Excellent	23.93	1.44	22.34	0.78	22.61	0.73
	Very good	26.82	1.34	34.0	0.62	32.81	0.63
	Good	31.5	2.05	28.1	0.67	28.66	0.66
	Fair	14.74	0.97	11.55	0.52	12.08	0.48
	Poor	3.01	0.65	4.0	0.26	3.84	0.25
Have medical insurance	Yes	68.2	1.77	84.15	0.66	81.51	0.67
	No	31.8	1.77	15.85	0.66	18.49	0.67
Had a flu shot	Yes	44.45	1.68	41.39	0.94	41.89	0.87
	No	55.55	1.68	58.61	0.94	58.11	0.87
Had a vision check	Yes	53.5	1.98	57.62	0.71	56.93	0.61
	No	46.5	1.98	42.38	0.71	43.07	0.61
Had a dentist visit	Yes	62.92	1.77	66.97	0.82	66.3	0.75
	No	37.08	1.77	33.03	0.82	33.7	0.75
Native speaker	No	100	0	0	0	16.61	0.74
	Yes	0	0	100	0	83.39	0.74
Race/ ethnicity	Hispanic	59.4	1.76	4.98	0.34		0.36
	White	15.45	1.2	76.55	0.81	66.42	0.74
	Black	3.99	0.69	14.0	0.14	12.34	0.1
	Other race	21.16	1.69	4.47	0.6	7.24	0.67

Age of English acquisition	Learned English as the first language	0	0	100	0	84.9	0.61
	Learned English when age 15 or younger	72.3	1.94	0	0	10.92	0.64
	Learned English when age 16 or older	27.7	1.94	0	0	4.18	0.26
Length of residence in the US	In host country 5 or fewer years	7.8	0.83	0.15	0.04	1.41	0.15
	In host country more than 5 years	62.44	1.69	2.87	0.33	12.71	0.37
	Non-immigrants	29.76	1.62	96.99	0.34	85.88	0.39
Ability to speak English	very well	45.98	2.12	92.37	0.54	84.67	0.72
	well	30.94	1.57	7.26	0.54	11.19	0.54
	not well	16.26	1.19	0.36	0.09	3	0.19
	not at all	6.82	1.64	0.01	0.01	1.14	0.31
Ability to read English	very well	50.35	1.82	91.52	0.45	84.69	0.59
	well	27.06	1.2	6.89	0.39	10.24	0.37
	not well	14.55	1.3	1.25	0.13	3.46	0.29
	not at all	8.04	1.26	0.34	0.08	1.62	0.24
Ability to write English	very well	43.99	1.65	87.34	0.67	80.14	0.72
	well	25.66	1.63	10.44	0.61	12.96	0.57
	not well	18.84	1.37	1.9	0.2	4.71	0.28
	not at all	11.52	1.47	0.32	0.09	2.18	0.29
Ability to understand English	very well	55.56	2.06	95.78	0.35	89.1	0.56
	well	26.42	1.34	4	0.35	7.72	0.35
	not well	13.2	1.21	0.16	0.05	2.33	0.24
	not at all	4.82	0.95	0.06	0.04	0.85	0.19

Note: Columns add to total (100%).

Appendix B. Collinearity Diagnostics
Note: VIF values above 10 indicate multicollinearity.

Table 1B. Weighted collinearity diagnostics, including literacy and numeracy

	Collinearity Statistics	
Variable	Tolerance	VIF
Gender	0.930	1.076
Native speaker	0.900	1.111
Health status	0.862	1.161
Education	0.688	1.454
Literacy	0.194	5.154
Numeracy	0.200	5.012
Age	0.861	1.162
Have medical insurance	0.906	1.104

Note: VIF values above 10 indicate multicollinearity.

Table 2B. Weighted collinearity diagnostics, including only numeracy

	Collinearity Statistics	
Variable	Tolerance	VIF
Gender	0.964	1.038
Native speaker	0.933	1.072
Health status	0.865	1.156
Education	0.701	1.426
Numeracy	0.657	1.521
Age	0.877	1.141
Have medical insurance	0.909	1.100

Note: VIF values above 10 indicate multicollinearity.

Table 3B. Weighted collinearity diagnostics, including only literacy

	Collinearity Statistics	
Variable	Tolerance	VIF
Gender	0.993	1.007
Native speaker	0.902	1.109
Health status	0.863	1.159
Education	0.700	1.428
Literacy	0.639	1.564
Age	0.861	1.162
Have medical insurance	0.906	1.104

Note: VIF values above 10 indicate multicollinearity.

Table 4B. Weighted collinearity diagnostics for the non-native speakers, including self-rated language variables, literacy, and numeracy

	Collinearity Statistics	
Variable	Toleranc	VIF

Gender	0.954	1.049
Health status	0.873	1.146
Education	0.635	1.574
Numeracy	0.207	4.831
Literacy	0.174	5.738
Age	0.726	1.378
Have medical insurance	0.868	1.152
Age of English language acquisition	0.639	1.564
Ability to speak English	0.223	4.484
Ability to read English	0.200	5.002
Ability to write English	0.201	4.972
Ability to understand English	0.290	3.446

Note: VIF values above 10 indicate multicollinearity.

Table 5B. Weighted collinearity diagnostics for the non-native speaker sample, including self-rated language variables, and literacy

	Collinearity Statistics	
Variable	Tolerance	VIF
Gender	0.983	1.017
Health status	0.876	1.142
Education	0.661	1.513
Literacy	0.475	2.106
Age	0.741	1.350
Have medical insurance	0.871	1.148
Age of English language acquisition	0.663	1.508
Ability to speak English	0.223	4.482
Ability to read English	0.200	5.002
Ability to write English	0.202	4.961
Ability to understand English	0.291	3.435

Note: VIF values above 10 indicate multicollinearity.

Table 6B. Weighted collinearity diagnostics for the non-native speaker sample, including self-rated language variables, and numeracy

	Collinearity Statistics	
Variable	Tolerance	VIF
Gender	0.965	1.036
Health status	0.873	1.146
Education	0.641	1.559
Numeracy	0.564	1.774
Age	0.784	1.276
Have medical insurance	0.869	1.150
Age of English language acquisition	0.669	1.495
Ability to speak English	0.223	4.479
Ability to read English	0.202	4.961
Ability to write English	0.201	4.972

Ability to understand English	0.291	3.437

Note: VIF values above 10 indicate multicollinearity.

Appendix C. Regression models

Table 1C. Weighted logistic regression coefficients for models for dependent variable 'Had a dentist visit in the last 12 months'.

Variable	LR1	LR2	LR3	LR4	LR5	LR6	LR7	LR8
Native speaker (referent=nonnative speaker)	0.18	0.17	0.180	-0.21	-0.21	NS	NS	NS
Age		NS	NS		NS	NS	NS	NS
Gender (referent=female)			-0.23		-0.23	-0.22	-0.24	-0.22
Race: Hispanic (referent=White)				-0.68	-0.68			
Race: Black (referent=White)				-0.36	-0.37			
Race: Other (referent=White)				-0.28	-0.29			
Education						0.33	0.29	0.24
Employment: unemployed (referent=employed)						-0.43	-0.36	-0.2
Employment: out of labor force (referent=employed)						NS	NS	NS
Health status							0.32	0.31
Have medical insurance (referent=do not have insurance)								1.27
Nagelkerke's R^2	0.00	0.00	0.01	0.01	0.02	0.09	0.12	0.18

Note: unweighted sample includes all responders.
Significance level of significant variables is $p < 0.05$. All estimates are significant unless noted with NS (not significant).

Table 2C. Weighted regression coefficients for models for dependent variable 'Had a dentist visit in the last 12 months', focusing on the impact of literacy and numeracy.

Variable	LR1	LR2	LR3	LR4	LR5
Native speaker (referent=nonnative speaker)				-0.17	NS
Age				NS	0.04
Gender (referent=female)				-0.23	-0.26
Education				0.2	0.2
Employment: unemployed (referent=employed)				NS	-0.33
Employment: out of labor force (referent=employed)				NS	NS
Health status				0.29	0.29

Have medical insurance (referent=do not have insurance)				1.24	
Literacy	0.01		0.01	NS	NS
Numeracy		0.01	NS	NS	NS
Nagelkerke’s R^2	0.06	0.05	0.06	0.18	0.12

Note: unweighted sample includes all responders.
Significance level of significant variables is $p < 0.05$. All estimates are significant unless noted with NS (not significant).

Table 3C. Weighted regression coefficients for models for dependent variable ‘Had a vision check in the last 12 months’.

Variable	LR1	LR2	LR3	LR4	LR5
Native speaker (referent=nonnative speaker)	NS	NS	NS	NS	NS
Age		0.1	-0.04	0.09	0.07
Gender (referent=female)		-0.13	-0.22	-0.11	NS
Race: Hispanic (referent=White)		-0.26	-0.39	NS	NS
Race: Black (referent=White)		0.25	-0.21	0.32	0.37
Race: Other (referent=White)		0.46	-0.38	0.44	0.46
Education			0.32	0.09	0.05
Employment: unemployed (referent=employed)			-0.4	NS	NS
Employment: out of labor force (referent=employed)			NS	0.34	0.28
Health status				0.13	0.11
Have medical insurance (referent=do not have insurance)					0.93
Nagelkerke’s R^2	0.00	0.02	0.03	0.04	0.07

Note: unweighted sample includes all responders.
Significance level of significant variables is $p < 0.05$. All estimates are significant unless noted with NS (not significant).

Table 4C. Weighted regression coefficients for models for dependent variable ‘Had a vision check in the last 12 months’, focusing on the impact of literacy and numeracy.

Variable	LR1	LR2	LR3	LR4
Native speaker (referent=nonnative speaker)				NS
Age				0.07
Gender (referent=female)				NS
Race: Hispanic (referent=White)				NS
Race: Black (referent=White)				0.35
Race: Other (referent=White)				0.46
Education				0.06
Employment: unemployed (referent=employed)				NS
Employment: out of labor force (referent=employed)				0.28

Health status				0.12
Have medical insurance (referent=do not have insurance)				0.93
Literacy	0.001		NS	NS
Numeracy		NS	NS	NS
Nagelkerke's R^2	0.00	0.00	0.00	0.07

Note: unweighted sample includes all responders.
Significance level of significant variables is p <0.05. All estimates are significant unless noted with NS (not significant).

Table 5C. Weighted regression coefficients for models for dependent variable 'Had a flu shot in the last 12 months'.

Variable	LR1	LR2	LR3	LR4	LR5
Native speaker (referent=nonnative speaker)	NS	NS	NS	NS	NS
Age		0.22	0.19	0.2	0.19
Gender (referent=female)		-0.25	-0.21	-0.21	-0.2
Race: Hispanic (referent=White)		NS	0.29		0.44
Race: Black (referent=White)		NS	NS		NS
Race: Other (referent=White)		0.42	0.38		0.4
Education			0.1	0.1	0.06
Employment: unemployed (referent=employed)			NS	NS	0.21
Employment: out of labor force (referent=employed)			0.39	0.41	0.36
Health status				NS	NS
Have medical insurance (referent=do not have insurance)					1.09
Nagelkerke's R^2	0.00	0.05	0.06	0.06	0.10

Note: unweighted sample includes all responders.
Significance level of significant variables is p <0.05. All estimates are significant unless noted with NS (not significant).

Table 6C. Weighted regression coefficients for models for dependent variable 'Had a flu shot in the last 12 months', focusing on the impact of literacy and numeracy.

Variable	LR1	LR2	LR3	LR4
Native speaker (referent=nonnative speaker)				NS
Age				0.18
Gender (referent=female)				-0.17
Race: Hispanic (referent=White)				0.4
Race: Black (referent=White)				NS
Race: Other (referent=White)				0.38
Education				0.08
Employment: unemployed (referent=employed)				0.2
Employment: out of labor force (referent=employed)				0.34
Health status				NS

Have medical insurance (referent=do not have insurance)				1.11
Literacy	NS		NS	NS
Numeracy		NS	NS	NS
Nagelkerke's R^2	0.00	0.00	0.00	0.10

Note: unweighted sample includes all responders.
Significance level of significant variables is p <0.05. All estimates are significant unless noted with NS (not significant).

Table 7C. Weighted logistic regression coefficients for models for dependent variable 'Had a dentist visit in the last 12 months' with the sample of non-native English speakers.

Variable	LR1	LR2	LR3	LR4	LR5	LR6	LR7	LR8	LR9	LR10	LR11	LR12	LR13	LR14	LR15
Age							NS	NS	NS	NS	NS	NS	NS	NS	NS
Gender (referent=female)							-0.48	-0.44	-0.46	-0.44	-0.44	-0.49	-0.48	-0.44	-0.47
Race: Hispanic (referent=white)							-0.76	NS	NS	NS	NS	NS	NS	NS	NS
Race: Black (referent=white)							-1.15	-0.98	-1.01	-0.94	-0.97	NS	-0.87	-0.91	-0.87
Race: Other (referent=white)							NS	NS	NS	NS	NS	NS	NS	NS	NS
Education								0.24	0.23	0.19	0.17	0.15	0.14	0.15	0.14
Employment: unemployed (referent=employed)								NS	NS	NS	NS	NS	NS	NS	NS
Employment: out of labor force (referent=employed)								NS	NS	NS	NS	NS	NS	NS	NS
Health status									0.19	0.18	NS	NS	NS	NS	NS
Have medical insurance (referent=do not have insurance)										1.41	1.4	1.38	1.39	1.39	1.39
Literacy												NS	NS	NS	
Numeracy												NS	NS		NS
Age of English language acquisition (referent=learned English when age 16 or older)	NS					-0.56	NS	NS	NS	NS	-0.68	NS	-0.66	-0.69	-0.67
Ability to speak English		0.35				NS					NS		NS	NS	NS
Ability to read English			0.41			NS					NS		NS	NS	NS
Ability to write English				0.39		NS					NS		NS	NS	NS
Ability to understand English					0.37	NS					NS		NS	NS	NS
Nagelkerke's R^2	0.00	0.03	0.05	0.05	0.03	0.04	0.05	0.1	0.11	0.2	0.21	0.21	0.22	0.22	0.22

Note: unweighted sample includes non-native speakers of English only.
Significance level of significant variables is $p < 0.05$. All estimates are significant unless noted with NS (not significant).

Table 8C. Weighted logistic regression coefficients for models for dependent variable ‘Had a vision check in the last 12 months’ with the sample of non-native English speakers.

Variable	LR1	LR2	LR3	LR4	LR5	LR6	LR7	LR8	LR9	LR10	LR11	LR12	LR13	LR14	LR15
Age							0.18	0.16	0.18	0.15	0.18	0.16	0.18	0.17	0.17
Gender (referent=female)							NS	NS	NS	NS	NS	NS	NS	NS	NS
Race: Hispanic (referent=white)							NS	NS	NS	NS	NS	NS	NS	NS	NS
Race: Black (referent=white)							NS	NS	NS	NS	NS	NS	NS	NS	NS
Race: Other (referent=white)							0.68	0.71	0.7	0.79	0.9	0.82	0.88	0.88	0.87
Education								NS	NS	NS	NS	NS	NS	NS	NS
Employment: unemployed (referent=employed)								NS	NS	NS	NS	NS	NS	NS	NS
Employment: out of labor force (referent=employed)								0.63	0.63	0.57	0.59	0.57	0.57	0.58	0.58
Health status									NS	NS	NS	NS	NS	NS	NS
Have medical insurance (referent=do not have insurance)										1.03	0.96	1.02	0.98	0.98	0.98
Numeracy												NS	NS	NS	
Literacy												NS	NS		NS
Age of English language acquisition (referent=learned English when age 16 or older)	NS					NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Ability to speak English		0.35				NS					NS		NS	NS	NS
Ability to read English			0.39			NS					NS		NS	NS	NS
Ability to write English				0.39		0.37					0.44		0.45	0.44	0.45
Ability to understand English					0.38	NS					NS		NS	NS	NS
Nagelkerke’s R^2	0.00	0.03	0.04	0.05	0.04	0.03	0.05	0.06	0.07	0.12	0.14	0.13	0.15	0.15	0.15

Note: unweighted sample includes non-native speakers of English only.
Significance level of significant variables is $p < 0.05$. All estimates are significant unless noted with NS (not significant).

Table 9C. Weighted logistic regression coefficients for models for dependent variable ‘Had a flu shot in the last 12 months’ with the sample of non-native English speakers.

Variable	LR1	LR2	LR3	LR4	LR5	LR6	LR7	LR8	LR9	LR10	LR11	LR12	LR13	LR14	LR15
Age							NS	NS	NS	NS	NS	NS	NS	NS	NS
Gender (referent=female)							-0.31	NS	NS	NS	NS	NS	NS	NS	NS
Race: Hispanic (referent=white)							NS	NS	NS	0.46	0.47	NS	NS	NS	NS
Race: Black (referent=white)							NS	NS	NS	NS	NS	NS	NS	NS	NS
Race: Other (referent=white)							0.54	0.58	0.57	0.6	0.62	NS	NS	0.54	0.55
Education								NS	NS	NS	NS	NS	NS	NS	NS
Employment: unemployed (referent=employed)								NS	NS	NS	NS	NS	NS	NS	NS
Employment: out of labor force (referent=employed)								0.6	0.6	0.52	0.51	0.5	0.49	0.5	0.49
Health status									NS	NS	NS	NS	NS	NS	NS
Have medical insurance (referent=do not have insurance)										1.24	1.22	1.3	1.28	1.28	1.27
Numeracy												NS	NS	-0.01	
Literacy												NS	NS		-0.005
Age of English language acquisition (referent=learned English when age 16 or older)	NS					NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Ability to speak English		NS				NS					NS		NS	NS	NS
Ability to read English			0.2			NS					NS		NS	NS	NS
Ability to write English				NS		NS					NS		NS	NS	NS
Ability to understand English					NS	NS					NS		NS	NS	NS
Nagelkerke’s R^2	0.00	0.00	0.01	0.01	0.01	0.01	0.03	0.04	0.04	0.12	0.12	0.13	0.14	0.13	0.14

Note: unweighted sample includes non-native speakers of English only.
Significance level of significant variables is $p < 0.05$. All estimates are significant unless noted with NS (not significant).

Table 10C. Weighted logistic regression coefficients for final selection models for all three dependent variables.

variable	dentist visit	vision check	flu shot
Native speaker (referent=nonnative speaker)	NS	NS	NS
Age	NS	0.07	0.19
Gender (referent=female)	-0.22	NS	-0.2
Race: Hispanic (referent=White)		NS	0.44
Race: Black (referent=White)		0.37	NS
Race: Other (referent=White)		0.46	0.4
Education	0.24	0.05	0.06
Employment: unemployed (referent=employed)	-0.2	NS	0.21
Employment: out of labor force (referent=employed)	NS	0.28	0.36
Health status	0.31	0.11	NS
Have medical insurance (referent=do not have insurance)	1.27	0.93	1.09
Nagelkerke's R^2	0.18	0.07	0.1

Note: unweighted sample includes all responders.
Significance level of significant variables is $p < 0.05$. All estimates are significant unless noted with NS (not significant).

Table 11C. Weighted logistic regression coefficients for final selection models for all three dependent variables with the sample of non-native English speakers.

variable	dentist visit	vision check	flu shot
Age	NS	0.18	NS
Gender (referent=female)	-0.48	NS	NS
Race: Hispanic (referent=white)	NS	NS	NS
Race: Black (referent=white)	-0.87	NS	NS
Race: Other (referent=white)	NS	0.88	NS
Education	0.14	NS	NS
Employment: unemployed (referent=employed)	NS	NS	NS
Employment: out of labor force (referent=employed)	NS	0.57	0.49
Health status	NS	NS	NS
Have medical insurance (referent=do not have insurance)	1.39	0.98	1.28
Age of English language acquisition (referent=learned English when age 16 or older)	-0.66	NS	NS
Ability to speak English	NS	NS	NS
Ability to read English	NS	NS	NS
Ability to write English	NS	0.45	NS
Ability to understand English	NS	NS	NS

Numeracy	NS	NS	NS
Literacy	NS	NS	NS
Nagelkerke's R^2	0.22	0.15	0.14

Note: unweighted sample includes non-native speakers of English only.
Significance level of significant variables is $p < 0.05$. All estimates are significant unless noted with NS (not significant).

References

Aguilera, M. B., & Massey, D. S. (2003). Social capital and the wages of Mexican migrants: New hypotheses and tests. *Social forces*, *82*(2), 671-701.

AIR PIAAC Team. (2019). What you need to consider before working with PIAAC data.
Alegría, M., Robles, R., Freeman, D. H., Vera, M., Jimenez, A. L., Rios, C., & Rios, R. (1991). Patterns of mental health utilization among island Puerto Rican poor. *American Journal of Public Health*, *81*(7), 875-879.

Antecol, H., & Bedard, K. (2006). Unhealthy assimilation: why do immigrants converge to American health status levels? *Demography*, *43*(2), 337-360.

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low health literacy and health outcomes: an updated systematic review. *Annals of Internal Medicine*, *155*(2), 97-107.

Birdsong, D. (Ed.). (1999). *Second language acquisition and the critical period hypothesis*. Routledge.

Blackwell, D. L., Martinez, M. E., Gentleman, J. F., Sanmartin, C., & Berthelot, J. M. (2009). Socioeconomic status and utilization of health care services in Canada and the United States: findings from a binational health survey. *Medical Care*, 1136-1146.

Broyles, R. W., McAuley, W. J., & Baird-Holmes, D. (1999). The medically vulnerable: their health risks, health status, and use of physician care. *Journal of Health Care for the Poor and Underserved*, *10*(2), 186-200.

Campbell, C. M., & Edwards, R. R. (2012). Ethnic differences in pain and pain management. *Pain Management*, *2*(3), 219-230.

Carrasquillo, O., Carrasquillo, A. I., & Shea, S. (2000). Health insurance coverage of immigrants living in the United States: differences by citizenship status and country of origin. *American Journal of Public Health*, *90*(6), 917.

Cavanaugh, K., Huizinga, M. M., Wallston, K. A., Gebretsadik, T., Shintani, A., Davis, D., ... & Rothman, R. L. (2008). Association of numeracy and diabetes control. *Annals of Internal Medicine*, *148*(10), 737-746.

Chang, E., Chan, K. S., & Han, H. R. (2015). Effect of acculturation on variations in having a usual source of care among Asian Americans and non-Hispanic whites in California. *American Journal of Public Health*, *105*(2), 398-407.

Chi, J. T., & Handcock, M. S. (2014). Identifying sources of health care underutilization among California's immigrants. *Journal of Racial and Ethnic Health Disparities*, *1*(3), 207-218.

Cummins, P. A., Yamashita, T., & Arbogast, A. (2018). An Examination of PIAAC Data for Unemployed Adults Aged 45 to 74. Retrieved from PIAAC Gateway website: https://static1.squarespace.com/static/51bb74b8e4b0139570ddf020/t/5ad75ba96d2a73e3d0cd361a/1524063146040/Cummins+Yamashita+Arbogast_2018.pdf

Dallo, F. J., & Kindratt, T. B. (2015). Disparities in preventive health behaviors among non-Hispanic White men: heterogeneity among foreign-born Arab and European Americans. *American Journal of Men's Health*, *9*(2), 124-131.

de Boer, A. G., Wijker, W., & de Haes, H. C. (1997). Predictors of health care utilization in the chronically ill: a review of the literature. *Health Policy*, *42*(2), 101-115.

Derose, K. P., Escarce, J. J., & Lurie, N. (2007). Immigrants and health care: sources of vulnerability. *Health Affairs*, *26*(5), 1258-1268.

Doty, H. E., & Weech-Maldonado, R. (2003). Racial/ethnic disparities in adult preventive dental care use. *Journal of Health Care for the Poor and Underserved*, *14*(4), 516-534.

Feinberg, E., Swartz, K., Zaslavsky, A. M., Gardner, J., & Walker, D. K. (2002). Language proficiency and the enrollment of Medicaid-eligible children in publicly funded health insurance programs. *Maternal and Child Health Journal*, *6*(1), 5-18.

Feinberg, I., Frijters, J., Johnson-Lawrence, V., Greenberg, D., Nightingale, E., & Moodie, C. (2016). Examining associations between health information seeking behavior and adult education status in the US: An analysis of the 2012 PIAAC Data. *PloS one*, *11*(2), e0148751.

Flavin, L., Zallman, L., McCormick, D., & Wesley Boyd, J. (2018). Medical expenditures on and by immigrant populations in the United States: a systematic review. *International Journal of Health Services*, *48*(4), 601-621.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 527-552.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*(1), 78-104.

Giltay, E. J., Vollaard, A. M., & Kromhout, D. (2012). Self-rated health and physician-rated health as independent predictors of mortality in elderly men. *Age and Ageing*, *41*(2), 165-171.

Hair, J. F. Jr., Anderson, R.E., Tathan, R.L., & Black, W.C. (1995). *Multivariate Data Analysis* (3rd ed.). New York: Macmillan.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, *14*(1), 31-38.

Hall, E., & Cuellar, N. G. (2016). Immigrant health in the United States: A trajectory toward change. *Journal of Transcultural Nursing*, *27*(6), 611-626.Harlan, Bernstein, & Kessler, 1991

Hernández-Quevedo, C., & Jiménez-Rubio, D. (2009). A comparison of the health status and health care utilization patterns between foreigners and the national population in Spain: new evidence from the Spanish National Health Survey. *Social Science & Medicine*, *69*(3), 370-378.

Hogan, J., Thornton, N., Diaz-Hoffmann, L., Mohadjer, L., Krenzke, T., Li, J.,...Provasnik, S. (2016). U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014: Main study and national supplement technical report (NCES 2016-036). Retrieved from nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2016036

Holtzman, S., Kandathil, M., Kapur, L., Kline, D., Barone, J., Li, L., Krenzke, T., Hogan, J., Mohadjer, L., Carstens, R., Koehler, H., & Daniel, T. (2016a). PIAAC U.S. household public use file codebook. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2016667

Holtzman, S., Kandathil, M., Kapur, L., Kline, D., Barone, J., Li, L., Krenzke, T., Hogan, J., Mohadjer, L., Carstens, R., Koehler, H., & Daniel, T. (2016b). Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014: U.S. national supplement public use data files-household. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2016667

Jacobs, E., Chen, A. H., Karliner, L. S., Agger-Gupta, N., & Mutha, S. (2006). The need for more research on language barriers in health care: a proposed research agenda. *The Milbank Quarterly*, *84*(1), 111-133.

Jacobson, H. E., Hund, L., & Mas, F. S. (2016). Predictors of English Health Literacy among US Hispanic Immigrants: The importance of language, bilingualism and sociolinguistic environment. *Literacy & Numeracy Studies: An International Journal in the Education and Training of Adults*, *24*(1), 43.

Jang, M., Lee, E., & Woo, K. (1998). Income, language, and citizenship status: factors affecting the health care access and utilization of Chinese Americans. *Health & Social Work*, *23*(2), 136-145.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60-99.

Kasper, G., & Rose, K. R. (2002). Pragmatic Development in a Second Language. *Language Learning: A Journal of Research in Language Studies*, *52*, 1.

Keisler-Starkey, K., & Bunch, L. (2020). Health Insurance Coverage in the United States: 2019. Retrieved from https://www.census.gov/library/publications/2020/demo/p60-271.html#:~:text=The%20percentage%20of%20people%20with,of%202019%20was%2092.0%20percent.&text=Private%20health%20insurance%20coverage%20was,point%20during%20the%20year%2C%20respectively

Kennedy, P. (1992). *A Guide to Econometrics*. Oxford: Blackwell.
Kirkman-Liff, B., & Mondragón, D. (1991). Language of interview: relevance for research of southwest Hispanics. *American Journal of Public Health*, *81*(11), 1399-1404.

Krashen, S. D., Long, M. A., & Scarcella, R. C. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly*, 573-582.

Ku, L. (2009). Health insurance coverage and medical expenditures of immigrants and native-born citizens in the United States. *American Journal of Public Health*, *99*(7), 1322-1328.

Ku, L., & Matani, S. (2001). Left out: immigrants' access to health care and insurance. *Health Affairs*, *20*(1), 247-256.

Kutner, M., Greenburg, E., Jin, Y., & Paulsen, C. (2006). The Health Literacy of America's Adults: Results from the 2003 National Assessment of Adult Literacy. NCES 2006-483. *National Center for Education Statistics*.

Lebrun, L. A. (2012). Effects of length of stay and language proficiency on health care experiences among immigrants in Canada and the United States. *Social Science & Medicine*, *74*(7), 1062-1072.

Lees, K. A., Wortley, P. M., & Coughlin, S. S. (2005). Comparison of racial/ethnic disparities in adult immunization and cancer screening. *American Journal of Preventive Medicine*, *29*(5), 404-411.

Liao, X. H., & McIlwaine, G. (1995). The health status and health needs of Chinese population in Glasgow. *Scottish Medical Journal*, *40*(3), 77-80.

Lucas, J. W., Barr-Anderson, D. J., & Kington, R. S. (2003). Health status, health insurance, and health care utilization patterns of immigrant Black men. *American Journal of Public Health*, *93*(10), 1740-1747.

Lunze, K., & Paasche-Orlow, M. K. (2014). Limited literacy and poor health: The role of social mobility in Germany and the United States. *Journal of Health Communication*, *19*(sup2), 15-18.

Marquardt, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation and non-linear estimation. *Technometrics, 12,* 591-612.

McDonald, J. T., & Kennedy, S. (2004). Insights into the 'healthy immigrant effect': health status and health service use of immigrants to Canada. *Social Science & Medicine*, *59*(8), 1613-1627.

McDonald, J. T., & Kennedy, S. (2005). Is migration to Canada associated with unhealthy weight gain? Overweight and obesity among Canada's immigrants. *Social Science & Medicine*, *61*(12), 2469-2481.

Mellner, C., & Lundberg, U. (2003). Self-and physician-rated general health in relation to symptoms and diseases among women. *Psychology, Health & Medicine*, *8*(2), 123-134.

Morris, S., Sutton, M., & Gravelle, H. (2005). Inequity and inequality in the use of health care in England: an empirical investigation. *Social Science & Medicine*, *60*(6), 1251-1266.

Naish, J., Brown, J., & Denton, B. (1994). Intercultural consultations: investigation of factors that deter non-English speaking women from attending their general practitioners for cervical screening. *BMJ: British Medical Journal*, *309*(6962), 1126.

Neter, J., Wasserman, W., & Kutner, M.H., (1990). *Applied Linear Regression Models*. Homewood, IL: Irwin.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11-28.

Paasche-Orlow, M. K., & Wolf, M. S. (2007). The causal pathways linking health literacy to health outcomes. *American Journal of Health Behavior*, *31*(1), S19-S26.

Pan, Y., & Jackson, R.T. (2008). Ethnic difference in the relationship between acute inflammation and serum ferritin in US adult males. *Epidemiology and Infection*, 136,421-431.

Pandey, S., & Kagotho, N. (2010). Health insurance disparities among immigrants: are some legal immigrants more vulnerable than others? *Health & Social Work*, *35*(4), 267-279.

Parmet, W. (2017). Repeal, replace and leave out immigrants: The American health care act's impact on immigrants. Retrieved from https://blog.petrieflom.law.harvard.edu/2017/03/11/repeal-replace-and-leave-out-immigrants-the-american-health-care-acts-impact-on-immigrants/

Parslow, R., Jorm, A., Christensen, H., & Jacomb, P. (2002). Factors associated with young adults' obtaining general practitioner services. *Australian Health Review*, *25*(6), 109-118.

Pedhazur, E. (1997). Multiple regression in behavioral research. (3rd ed.) Orlando, FL: Harcourt Brace.

Pereira, K., Crosnoe, R., Fortuny, K., Pedroza, J., Ulvestad, K., Weiland, C., & Yoshikawa, H. (2012). Barriers to immigrants' access to health and human services programs. Retrieved from https://www.urban.org/sites/default/files/publication/33666/413260-Barriers-to-Immigrants-Access-to-Health-and-Human-Services.PDF

Pitkin Derose, K., Bahney, B. W., Lurie, N., & Escarce, J. J. (2009). Immigrants and health care access, quality, and cost. *Medical Care Research and Review*, *66*(4), 355-408.

Ponce, N. A., Hays, R. D., & Cunningham, W. E. (2006). Linguistic disparities in health care access and health status among older adults. *Journal of General Internal Medicine*, *21*(7), 786-791.

Prins, E., & Monnat, S. (2015). Examining associations between self-rated health and proficiency in literacy and numeracy among immigrants and US-born adults: Evidence from the Program for the International Assessment of Adult Competencies (PIAAC). *PloS one*, *10*(7), e0130257.

Pylypchuk, Y., & Hudson, J. (2009). Immigrants and the use of preventive care in the United States. *Health Economics*, *18*(7), 783-806.

Rogerson, P.A. (2001). *Statistical Methods for Geography*. London: Sage.

Schmidt, H., Waltner, A. M., & Muller, S. T. (2011). *The Immigrant Health Initiative: A Study of Health Care of Recent Immigrants in Dutchess County, New York*.

Shin, H. Y., Shin, M. H., & Rhee, J. A. (2012). Gender differences in the association between self-rated health and hypertension in a Korean adult population. *BMC Public Health*, *12*(1), 1-6.

Smith, K. T., Monti, D., Mir, N., Peters, E., Tipirneni, R., & Politi, M. C. (2018). Access is necessary but not sufficient: Factors influencing delay and avoidance of health care services. *MDM Policy & Practice*, *3*(1).

Stevens, G. (2006). The age-length-onset problem in research on second language acquisition among immigrants. *Language Learning*, *56*(4), 671-692.

Sudano Jr, J. J., & Baker, D. W. (2003). Intermittent lack of health insurance coverage and use of preventive services. *American Journal of Public Health*, *93*(1), 130-137.

Szwarcwald, C. L., Souza-Júnior, P. R., & Damacena, G. N. (2010). Socioeconomic inequalities in the use of outpatient services in Brazil according to health care need: evidence from the World Health Survey. *BMC Health Services Research*, *10*(1), 1-7.

Thorpe Jr, R. J., Wilson-Frederick, S. M., Bowie, J. V., Coa, K., Clay, O. J., LaVeist, T. A., & Whitfield, K. E. (2013). Health behaviors and all-cause mortality in African American men. *American Journal of Men's Health*, *7*(4_suppl), 8S-18S.

Weinick, R. & Krauss, N. (2001). Racial/Ethnic differences in children's access to care. *American Journal of Public Health*, 90 (11), 1771-1774.

Woloshin, S., Schwartz, L. M., Katz, S. J., & Welch, H. G. (1997). Is language a barrier to the use of preventive services? *Journal of General Internal Medicine*, *12*(8), 472-477.

Wu, B., Liang, J., Luo, H., & Furter, R. (2013). Racial and ethnic variations in preventive dental care utilization among middle-aged and older Americans, 1999–2008. *Frontiers in Public Health*, *1*, 65.

Xu, K. T., & Borders, T. F. (2008). Does being an immigrant make a difference in seeking physician services? *Journal of Health Care for the Poor and Underserved*, *19*(2), 380-390.

Yamashita, T., & Kunkel, S. R. (2015). An international comparison of the association among literacy, education, and health across the United States, Canada, Switzerland, Italy, Norway, and Bermuda: implications for health disparities. *Journal of Health Communication*, *20*(4), 406-415.

Yamashita, T., Bardo, A. R., Millar, R. J., & Liu, D. (2018a). Numeracy and preventive health care service utilization among middle-aged and older adults in the US. *Clinical Gerontologist*, *43*(2), 221-232.

Yamashita, T., Keene, J. R., Moonie, S. A., Shen, J. J., Pharr, J. R., & Yoo, J. W. (2018b). Literacy activity and health among older adults in the USA. *Educational Gerontology*, *44*(10), 627-638.

Yang, P. Q., & Hwang, S. H. (2016). Explaining immigrant health service utilization: a theoretical framework. *Sage Open*, *6*(2).

Ye, J., Mack, D., Fry-Johnson, Y., & Parker, K. (2012). Health care access and utilization among US-born and foreign-born Asian Americans. *Journal of Immigrant and Minority Health*, *14*(5), 731-737.

Zhang, X., Cotch, M. F., Ryskulova, A., Primo, S. A., Nair, P., Chou, C. F., ... & Saaddine, J. B. (2012). Vision health disparities in the United States by race/ethnicity, education, and economic status: findings from two nationally representative surveys. *American Journal of Ophthalmology*, *154*(6), S53-S62.