

The association between local area immigrant fraction and prevalence of cardiovascular diseases in the United States: an observational study



Deepa Shokeen,^a Natalie Wang,^b Ngan P. Nguyen,^c Ethan Bakal,^d Osika Tripathi,^{e,f} Latha P. Palaniappan,^g and Robert J. Huang^{a,*}

^aDivision of Gastroenterology and Hepatology, Stanford University School of Medicine, Stanford, CA, USA

^bUniversity of California Los Angeles, Los Angeles, CA, USA

^cDepartment of Psychology, Fulbright University Vietnam, Vietnam

^dNew York University College of Arts and Sciences, New York, USA

^eThe Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, La Jolla, CA, USA

^fSan Diego State University School of Public Health, San Diego, CA, USA

^gDivision of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA



Summary

Background Local area immigrant fraction is strongly and positively correlated with local life expectancy in the United States. The aim of the study was to determine the relationship between local area immigrant fraction and local prevalence of coronary heart disease (CHD) and stroke.

Methods Cross-sectional study design, with ZIP code as the unit of observation. Demographic data was obtained from the American Community Survey, and linked to indicators of health access (e.g., insurance, annual check-ups, cholesterol screening), obesity, behavior (smoking, exercise), and cardiovascular outcomes data from the 2020 Population Level Analysis and Community Estimates. Multivariable regression and path analyses were used to assess both direct and indirect relationships among variables.

Findings CHD prevalence was lower in the second (3.9% relative difference, 95% CI: 3.1–4.5%), third (6.5%, 95% CI: 5.8–7.1%), and fourth (14.8%, 95% CI: 14.1–15.8%) quartiles of immigrant fraction compared to the lowest (p-trend <0.001). These effects remained robust in multivariable analysis following adjustment for indicators of access, obesity, and behavioral variables (p-trend <0.0001). For stroke, only the highest quartile demonstrated a significant difference in prevalence (2.1%, 95% CI: 1.2–3.0% with full adjustment). In CHD path analysis, ~45% of the association of immigrant fraction was direct, and ~55% was mediated through lower prevalence of deleterious behaviors (e.g., smoking). In stroke path analysis, the effect was entirely mediated through indirect effects.

Interpretation In the United States, ZIP codes with higher immigrant fractions have lower prevalence of cardiovascular diseases. These associations are partially mediated through differences in health behaviors at the community level.

Funding NIH (K08CA252635, P30AG0059304, K24HL150476), Stanford University, Rutgers University.

Copyright © 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Coronary heart disease; Stroke; ZIP code; Healthy immigrant effect

Introduction

The United States (US) foreign-born (immigrant) population was estimated at 45 million in 2020.¹ Despite representing nearly one in seven Americans, immigrants have been historically underrepresented in

research studies. While immigrants have lower socioeconomic status than non-immigrants, they experience a lower prevalence of age-related diseases (such as atherosclerosis, stroke, diabetes, dementia, and osteoporosis) compared to non-immigrants.^{2–6} This has been

Abbreviations: ACS, American Community Survey; BMI, body mass index; BRFSS, Behavioral Risk Factor Surveillance System; CHD, coronary heart diseases; PLACES, Population Level Analysis and Community Estimates; US, United States; ZCTA, ZIP code tabulated area

*Corresponding author. 300 Pasteur Dr., Alway Building M211, Stanford, CA 94305, USA.

E-mail address: rjhuang@stanford.edu (R.J. Huang).

The Lancet Regional Health - Americas 2023;27: 100613

Published Online xxx
<https://doi.org/10.1016/j.lana.2023.100613>

Research in context

Evidence before this study

Immigrants have been historically underrepresented in research studies. While immigrants have lower socioeconomic status compared to non-immigrants, they experience lower prevalence of age-related diseases (such as atherosclerosis and stroke), a phenomenon termed the “healthy immigrant effect”. There exist sparse data on the association between local area immigrant fraction and community cardiovascular health outcomes. In July 2023, we searched PubMed using the search terms (‘immigrant’) AND (‘neighborhood’ OR ‘community’ OR ‘local’) AND (‘health’ OR ‘heart disease’ OR ‘cardiovascular’ OR ‘stroke’) AND (‘United States’ OR ‘US’ OR ‘USA’) AND (‘behavior’ OR ‘diet’ OR ‘lifestyle’ OR ‘exercise’ OR ‘food’) AND (‘density’ OR ‘fraction’). This data was supplemented by author knowledge. While a few studies have reported on the association between local immigrant density or fraction on the health behaviors of both immigrants and non-immigrants, these studies have mostly been performed within single or a small number of cities. Conversely, prior national database studies focused on the association of immigrant density/fraction and local health outcomes have mostly used larger geographic subdivisions (such as commuting zones or counties). In particular, no study has used small geographic area coding (such as ZIP code) to study the association between immigrant fraction and local health.

Added value of this study

Using ZIP-code linked national data sets of the United States, our study finds that both immigrants and non-immigrants

living in communities of higher immigrant fraction demonstrate lower prevalence of major cardiovascular diseases (coronary heart disease and stroke). Our study leveraged a well-phenotyped data set with access to preventative measures (e.g., annual cholesterol screening, primary care access) as well as behavioral measures (e.g., smoking, physical inactivity) available at the level of ZIP code, allowing us to adjust for confounding and ascertain indirect effects. Using structural equation modeling, we decomposed the direct and indirect (mediated) effects of immigrant fraction on cardiovascular disease prevalence, and found the majority of the association between immigrant fraction and reduced disease prevalence was mediated through lower prevalence of deleterious behaviors (such as inactivity and smoking).

Implications of all the available evidence

Our study suggests that immigrants may create positive externalities within their communities, such as influencing smoking or physical activity patterns, which may impact cardiovascular disease risk. These findings may hold important implications for community health, as there are an estimated 45 million immigrants residing in the United States as of year 2020. Our study was limited both by its cross-sectional nature and by the absence of individual-level data. Future longitudinal studies with individual-level risk factor and outcomes data are needed to confirm these findings.

termed the “healthy immigrant effect”. This effect may be due to differences in health behaviors between immigrants and non-immigrants, including increased physical activity, reduced smoking, reduced alcohol intake, and healthier diet.^{3,5,7–10} Alternatively, immigrant selection (healthier people are more likely to immigrate and less healthy people are more likely to return migrate) may also contribute to this effect.^{11,12}

Recent evidence suggests immigrants not only have better personal health outcomes, but may also impact the health behaviors of the neighborhoods and communities in which they reside. One survey study of Los Angeles County found neighborhood immigrant density to be negatively associated with regular fast food consumption, BMI, and hypertension, and positively associated with fruit/vegetable consumption among native-born Americans.⁵ A separate cross sectional survey study of Los Angeles demonstrated that among immigrants, living in a neighborhood with high immigrant density attenuated the negative effects of acculturation on fast food consumption.¹³ Thus, both immigrant Americans and native-born Americans living in neighborhoods with a high density of immigrants had healthier behaviors compared to those living in

neighborhoods with a low density of immigrants. A national, serial cross-sectional study of the US population used tax records and death records to evaluate factors associated with differences in life expectancy across areas.⁷ Within the bottom income quartile, the greatest single predictor of life expectancy (of both immigrants and non-immigrants) was local area immigrant fraction (on par with median home value and governmental expenditure per capita).⁷ Collectively, these data suggest that immigrant density and fraction within a given community may be associated with healthier behaviors and improved health outcomes among both immigrants and non-immigrants living in that community.

Building on this prior knowledge, we aimed to further explore the association between local area immigrant fraction and local health outcomes in this national cross-sectional study of the US. We chose to focus our analysis on two leading causes of cardiovascular morbidity and mortality which are highly sensitive to dietary and lifestyle patterns, coronary heart disease (CHD) and stroke. This study leverages a unique, national survey which provides detailed data on health behavior and outcomes to the level of ZIP code tabulated area (ZCTA). While prior studies associating neighborhood characteristics

with health outcomes have mostly used county as the unit of analysis, use of ZCTA provides a more granular geographic subdivision (~40,000 ZCTAs vs ~3000 counties in the US) which provides more precise estimates of local disease prevalence. We aimed to decompose the direct and indirect (mediated) association of local area immigrant fraction on cardiovascular disease prevalence through structural equation modeling.

Methods

Data sources and collected variables

This study was conducted on public data sources. A determination was issued by the Stanford University Institutional Review Board that this study did not constitute human subjects research (determination number 61866). This study followed the Strengthening the Reporting of Observational Studies in Epidemiology guidelines for cross-sectional studies. The data sources used in this study were the American Community Survey (ACS) 2014–2018 5-Year Estimates from the United States Census Bureau,¹⁴ and the 2020 Population Level Analysis and Community Estimates (PLACES) data release available through the US Centers for Disease Control and Prevention.¹⁵ These data sources were linked through five-digit ZCTA code. A comprehensive data dictionary of all variables utilized in this study are provided in [Supplementary Table S1](#).

The ACS is the largest source of small area statistics for social, economic, housing and demographic characteristics across the US and is based on sampling of ~3.5 million households annually. ZCTA-level data captured from the ACS included total population estimate, immigrant population estimate (further stratified by Hispanic ethnicity or Asian race), proportion of adults >65 years of age, male proportion, median household income, educational attainment (proportion of population >25 years of age with a high-school equivalent degree), and rural-urban commuting area codes. First released in December 2020, PLACES leverages data from the 2018 Behavioral Risk Factor Surveillance System (BRFSS), while providing model-based small area estimation; this allows for estimation of health behavior and outcomes prevalence to the level of ZCTA. The BRFSS is a large, health-related telephone survey system that collects data from 400,000 adults annually through a standardized core questionnaire. ZCTA-level data captured from PLACES included the following: proportion of adult population with health insurance, proportion of adults receiving annual health checkup, proportion of adults receiving annual cholesterol screening, proportion of adults currently smoking, proportion of adults who demonstrated physical inactivity (no exercise in past 30 days), and proportion of adults with obesity (defined as body mass index ≥ 30). The two outcome variables (CHD and stroke) were also captured from PLACES, and were defined as the

proportion of respondents by ZCTA who reported ever being diagnosed by a physician with CHD, or ever being told by a physician that they had had a stroke, respectively.

Study design

The independent variable of interest was ZCTA immigrant fraction, defined as the estimated immigrant population as the numerator and estimated total population as the denominator, and categorized by quartiles for analysis. We chose quartile categorization of immigrant fraction in order to offer a simpler interpretation of common effect measures from statistical models. We also performed a sensitivity analysis keeping immigrant fraction as a continuous variable (see below). From all total ZCTAs (N = 32,409), we chose to exclude ZCTAs where immigrant fraction was <1% (N = 16,038) in our main analysis. We chose to perform this exclusion for several reasons (see [Table 1](#) for characteristics of both the included and excluded ZCTAs), including small population size and significant difference in characteristics (*e.g.*, excluded ZCTAs were much more rural than included ZCTAs). Based on differences in characteristics, we felt that these low-immigrant ZCTAs did not share much common support with the included ZCTAs, making these observations ungeneralizable.^{16,17} Another reason for excluding these ZCTAs was the high rate of missing data (see [Supplementary Table S2](#)). After excluding low-immigrant ZCTAs, our main analytic cohort contained 16,371 ZCTAs. Notably, we performed sensitivity analyses including all ZCTAs which had complete data (N = 30,603), without restricting to ZCTAs with immigrant fraction $\geq 1\%$. We further performed sensitivity analyses treating the exposure measure (immigrant fraction) as a continuous variable.

The two primary outcomes measures were ZCTA prevalence of CHD and stroke. Estimates were adjusted for ZCTA sociodemographic characteristics (age structure, sex distribution, household income, educational attainment), indicators of health access (health insurance, annual health checkups, routine cholesterol screening), obesity, and health behaviors (smoking and physical inactivity). Subgroup analysis by Asian or Hispanic immigrant origin was performed by assessing the local area fraction of each group (*e.g.*, estimated Asian immigrant population as a fraction of total population). All analyses excluded ZCTAs with <1% of each immigrant group. ZCTAs were then categorized by quartile of both Asian and Hispanic immigrant fraction.

Statistical methods

We compared baseline characteristics between ZCTA quartiles using the Kruskal–Wallis test for continuous variables, and the Chi-squared test for categorical variables. We conducted multivariable linear regression to assess the association between ZCTA immigrant fraction and the primary outcome measures (CHD and

Variable	Excluded (N = 16,038)	Included 1st quartile (N = 4093)	Included 2nd quartile (N = 4092)	Included 3rd quartile (N = 4093)	Included 4th quartile (N = 4093)	p-value ^a	Combined included and excluded (N = 32,409)
Immigrant fraction (range)	<1%	1.0–2.8%	2.8–6.0%	6.0–13.2%	13.2–76.9%		0.0–76.9%
Population size, persons (Q1–Q3)	882 (324–2198)	4899 (2018–11,534)	9572 (3228–21,013)	17,212 (6966–29,866)	27,227 (13,843–41,312)	<0.0001	2918 (759–13,681)
Male, % (Q1–Q3)	50 (47–53)	50 (48–51)	49 (48–51)	49 (48–51)	49 (48–51)	<0.0001	50 (48–52)
>65 years of age, % (Q1–Q3)	19 (15–25)	19 (16–22)	17 (14–21)	16 (13–21)	13 (10–17)	<0.0001	18 (14–23)
>25 years with GED or equivalent, % (Q1–Q3)	89 (83–94)	91 (86–94)	92 (87–95)	92 (87–95)	88 (78–94)	<0.0001	90 (84–94)
Household income, 2018 USD (Q1–Q3)	51,701 (41,458–62,946)	54,750 (45,878–65,761)	60,129 (49,113–76,796)	65,784 (50,827–88,092)	68,848 (51,222–95,930)	<0.0001	55,961 (44,938–71,250)
Fraction classified as rural	44.3%	8.3%	4.9%	2.7%	1.2%	<0.0001	30.6%
Adults with health insurance, % (Q1–Q3)	88 (84–90)	88 (85–90)	88 (84–91)	88 (83–91)	85 (77–90)	<0.0001	88 (84–90)
Receiving annual health checkups, % (Q1–Q3)	76 (74–79)	77 (75–79)	77 (74–79)	76 (73–79)	74 (70–78)	<0.0001	76 (74–79)
Adults receiving annual cholesterol screens, % (Q1–Q3)	80 (78–82)	80 (78–82)	81 (78–83)	81 (78–84)	81 (77–84)	<0.0001	80 (78–83)
Adults currently smoking, % (Q1–Q3)	21 (18–24)	21 (18–24)	19 (16–22)	17 (14–20)	15 (12–19)	<0.0001	20 (17–23)
Low physical activity (no exercise last 30 days), % (Q1–Q3)	27 (24–32)	26 (23–30)	24 (20–28)	22 (19–27)	23 (19–29)	<0.0001	26 (22–30)
Obese (BMI ≥30), % (Q1–Q3)	36 (33–38)	35 (32–38)	33 (30–36)	31 (27–35)	28 (24–34)	<0.0001	34 (30–37)
Ever diagnosed with coronary heart disease, % (Q1–Q3)	8.2 (3.4–4.5)	7.7 (6.7–8.8)	6.9 (5.9–8.1)	6.2 (5.2–7.6)	5.6 (4.7–6.6)	<0.0001	7.5 (6.1–8.8)
Ever diagnosed with stroke, % (Q1–Q3)	3.9 (3.4–4.5)	3.7 (3.2–4.3)	3.3 (2.8–4.0)	3.1 (2.5–3.8)	2.9 (2.4–3.5)	<0.0001	3.6 (3.0–4.2)

ZCTA, ZIP code tabulated area; Q1–Q3, quartile 1 to quartile 3 range; USD, United States Dollars; GED, general educational development (high-school equivalency); BMI, body mass index. ^aSignifies p-value of multiple comparisons between included quartiles.

Table 1: Characteristics of ZCTAs in the United States (N = 32,409).

stroke). Linear trend analysis was performed by immigrant proportion quartile. We assessed for normality in the dependent variables (CHD and stroke) and found both to be parametrically distributed. As the prevalence of CHD and stroke are a continuous variable, we chose a simple linear regression (treating quartile as an ordinal variable) to assess if the distribution of the means of the dependent variable increase or decrease linearly. Models were checked for multi-collinearity between variables, outliers, and non-normality of residuals. For each association, several models were constructed, adjusting for ZCTA sociodemographic features, rurality, indicators of health access, obesity, and behavioral risk factor prevalence. Subgroup analysis was performed using Asian and Hispanic immigrant fraction quartiles as independent variables.

Structural equation modeling was used to visualize pathways and unravel the interrelationships among immigrant fraction, health behaviors, and the outcome measures of CHD and stroke. In this analysis, immigrant fraction, proportion >65 years of age, and male proportion were treated as exogenous variables. Health behaviors (smoking, physical inactivity) and obesity were treated as endogenous potential mediators, and outcomes (CHD and stroke) were treated as endogenous variables in the model. Notably, we did assess all variables as endogenous potential mediators; however, we chose to omit non-significant variables in the final model. The direction of the relationship in the path analysis were based on biological *a priori* and prior literature.^{18–20} Both direct and indirect path coefficients between exogenous and endogenous variables were assessed.

Statistical analysis was performed in R (version 4.1.1). All statistical tests were 2-tailed, and statistical significance was defined at $p < 0.05$.

Role of funding source

The study sponsors had no role in the collection, analysis, or interpretation of the data. The study sponsors had no role in the writing of the report, or in the decision to submit the manuscript for publication.

Results

Analytic cohort

A choropleth plot demonstrating the local area immigrant fraction among all ZCTAs in the US is depicted in [Supplementary Figure S1](#). Following restriction of ZCTAs with fewer than 1% immigrant fraction, the analytic cohort comprised of 16,371 unique ZCTAs. Descriptive characteristics of the included cohort, stratified by quartile, are shown in [Table 1](#). We observed several notable trends with increasing immigrant fraction quartile, including larger ZCTA population size, lower rural fraction, and higher household income. For instance, household income was ~\$55,000 in the lowest quartile and ~\$69,000 in the highest quartile. With

increasing quartile, the prevalence of both CHD and stroke decreased. At the same time, the prevalence of deleterious behaviors (smoking, inactivity) and obesity also decreased.

Histograms of immigrant fraction and socioeconomic characteristics are depicted in [Supplementary Figure S2](#). Scatter plots depicting the crude association between immigrant fraction and prevalence of health behaviors (current smoking, obesity, physical inactivity) and primary outcomes (CHD and stroke) are depicted in [Supplementary Figure S3](#). Increasing immigrant fraction demonstrated inverse correlations with obesity (adjusted r -squared = 0.11, $p < 0.0001$), current smoking (adjusted r -squared = 0.12, $p < 0.0001$), stroke (adjusted r -squared = 0.05, $p < 0.0001$), and CHD (adjusted r -squared = 0.10, $p < 0.0001$). A correlation matrix depicting the strength and direction of crude correlations among all included variables are depicted in [Supplementary Figure S4](#).

Multivariable analysis

The association between ZCTA immigrant fraction (by quartile) and local CHD prevalence (in % prevalence) is presented in [Table 2](#). In the base model (model 1, adjusted for age, sex rurality, income, and education), CHD prevalence was lower in the second (β : -0.26% , 95% CI: -0.30 to -0.21%), third (β : -0.43% , 95% CI: -0.47 to -0.38%), and fourth (β : -0.98% , 95% CI: -1.04 to -0.93%) quartiles of immigrant fraction compared to the lowest quartile, with robust trend (p -trend < 0.0001). This corresponded to a relative difference in CHD prevalence of 3.9% (95% CI: 3.1–4.5%), 6.5% (95% CI: 5.8–7.1%) and 14.8% (95% CI: 14.1–15.8%), respectively. Upon further adjustments for indicators of access factors (model 2), obesity (model 3), and health behaviors (model 4), moderate attenuation of effect estimates were observed. However, even in the fully-adjusted model (model 4) a strong trend persisted (p -trend < 0.0001). In the fully-adjusted model, the relative difference in the highest quartile was 5.0% (95% CI: 4.2–5.8%) compared to the lowest quartile.

In the base model, increasing Asian immigrant fraction was associated with lower CHD prevalence with a strong trend (p -trend < 0.0001). With further adjustments for indicators of health access (model 2), obesity (model 3), and health behaviors (model 4), the estimates attenuated but a robust trend persisted (p -trend < 0.0001 in all models). Similarly, when analyzing the association between Hispanic immigrant fraction and CHD prevalence, a strong inverse relationship was seen in the base model (p -trend < 0.0001). With adjustment for indicators of health access, health behaviors, and obesity, a moderate attenuation was again seen; however, highly significant trends toward decreasing CHD prevalence remained (p -trend < 0.0001 in all models). The fully-adjusted CHD model, with β -coefficients, is depicted in [Supplementary Table S3](#).

	Model 1	Model 2	Model 3	Model 4
All immigrants, % change in prevalence (95% CI)				
First quartile	Reference	Reference	Reference	Reference
Second quartile	-0.26 ^b (-0.30 to -0.21)	-0.27 ^b (-0.32 to -0.23)	-0.17 ^b (-0.22 to -0.13)	-0.02 (-0.06 to 0.02)
Third quartile	-0.43 ^b (-0.47 to -0.38)	-0.44 ^b (-0.49 to -0.40)	-0.24 ^b (-0.29 to -0.20)	-0.02 (-0.06 to 0.03)
Fourth quartile	-0.98 ^b (-1.04 to -0.93)	-0.95 ^b (-1.00 to -0.90)	-0.53 ^b (-0.59 to -0.48)	-0.33 ^b (-0.38 to -0.28)
p-trend	<0.0001	<0.0001	<0.0001	<0.0001
Asian immigrants, % change in prevalence (95% CI)				
First quartile	Reference	Reference	Reference	Reference
Second quartile	-0.21 ^b (-0.27 to -0.15)	-0.19 ^b (-0.25 to -0.14)	-0.15 ^a (-0.20 to -0.09)	-0.09 ^a (-0.14 to -0.04)
Third quartile	-0.41 ^b (-0.47 to -0.35)	-0.36 ^b (-0.42 to -0.30)	-0.25 ^b (-0.31 to -0.20)	-0.19 ^b (-0.24 to -0.14)
Fourth quartile	-0.71 ^b (-0.77 to -0.64)	-0.59 ^b (-0.65 to -0.53)	-0.30 ^b (-0.36 to -0.24)	-0.31 ^b (-0.37 to -0.26)
p-trend	<0.0001	<0.0001	<0.0001	<0.0001
Hispanic immigrants, % change in prevalence (95% CI)				
First quartile	Reference	Reference	Reference	Reference
Second quartile	-0.05 (-0.11 to 0.00)	-0.04 (-0.10 to 0.01)	0.02 (-0.03 to 0.07)	0.01 (-0.04 to 0.05)
Third quartile	-0.17 ^b (-0.23 to -0.11)	-0.15 ^b (-0.20 to -0.09)	-0.03 (-0.08 to 0.03)	-0.04 (-0.09 to 0.01)
Fourth quartile	-0.70 ^b (-0.77 to -0.62)	-0.65 ^b (-0.72 to -0.58)	-0.38 ^b (-0.45 to -0.31)	-0.32 ^b (-0.38 to -0.26)
p-trend	<0.0001	<0.0001	<0.0001	<0.0001

ZCTA, ZIP code tabulated area; CHD, coronary heart disease. Coefficients of association (absolute % change in CHD prevalence per quartile) with 95% CI depicted. Model 1 adjusted for age, sex and ZCTA sociodemographic factors (median household income, education, and rurality). Model 2 adjusted for age and sex, ZCTA sociodemographic factors, and health access factors (health insurance, annual checkup and cholesterol screening). Model 3 adjusted for age and sex, ZCTA sociodemographic factors, health access factors, and obesity. Model 4 adjusted for age and sex, ZCTA sociodemographic factors, health access factors, and health behaviors (e.g., smoking, activity level). ^aDenotes significance at p = 0.01. ^bDenotes significance at p < 0.0001.

Table 2: Association between ZCTA immigrant fraction and CHD prevalence.

The association between local area immigrant fraction and local prevalence of stroke is presented in [Table 3](#). There existed an inverse relationship between immigrant fraction and stroke prevalence in the base model. Stroke prevalence was lower within the second (β : -0.08%, CI: -0.11 to -0.06%), third (β : -0.16%, CI: -0.19 to -0.14%), and fourth (β : -0.41%, CI: -0.44 to -0.38%) quartiles of immigrant fraction compared to the lowest quartile, with significant trend (p-trend <0.0001). This corresponded to relative difference in stroke prevalence of 2.4% (95% CI: 1.8–3.3%), 4.8% (95% CI: 4.2–5.8%), and 12.4% (95% CI: 11.5–13.3%), respectively. This effect attenuated though remained significant following adjustments for indicators of health access (p-trend <0.0001), obesity (p-trend <0.0001), and health behaviors (p-trend = 0.00076). Notably in the fully-adjusted model (model 4), only the highest quartile demonstrated a difference in stroke prevalence (relative difference of 2.1%, 95% CI: 1.2–3.0%). These effects persisted when analyzing both Asian immigrant fraction and Hispanic immigrant fraction, although the effect sizes were quite modest. The fully-adjusted stroke model, with β -coefficients, is depicted in [Supplementary Table S4](#).

Sensitivity analyses

We assessed the sensitivity of the relationship toward the inclusion of all ZCTAs (including those with <1% immigrant fraction). Of the ZCTAs with <1%

immigrant fraction (N = 16,038), 1806 were missing data (see [Supplementary Table S2](#)). The analytic cohort for sensitivity analyses therefore contained 30,603 ZCTAs. The association between ZCTA immigrant fraction (by quartile) and local CHD prevalence is presented in [Supplementary Table S5](#). Similar to the main CHD analysis, increasing immigrant fraction quartile associated with lower CHD prevalence in all models (p-trend <0.0001 for all models). The association between ZCTA immigrant fraction (by quartile) and local stroke prevalence is presented in [Supplementary Table S6](#). Similar to the main stroke analysis, increasing immigrant fraction quartile associated with lower stroke prevalence (p-trend <0.0001 for all models). We also analyzed immigrant fraction as a continuous variable ([Supplementary Table S7](#)). In the fully-adjusted model, every 10% increase in immigrant fraction was associated with a 0.22% (95% CI: 0.20–0.25%) decrease in CHD prevalence. In the fully-adjusted model, every 10% increase in immigrant fraction was associated with a 0.08% (95% CI: 0.07–0.09%) decrease in stroke prevalence.

Path analysis

Path analyses was performed to investigate the relationships between immigrant fraction, ZCTA demographic characteristics, obesity, health behaviors (smoking, inactivity), and outcomes. Notably while these analyses may offer insight as to direct vs indirect

	Model 1	Model 2	Model 3	Model 4
All immigrants, % change in prevalence (95% CI)				
First quartile	Reference	Reference	Reference	Reference
Second quartile	-0.08 ^c (-0.11 to -0.06)	-0.09 ^c (-0.11 to -0.07)	-0.02 (-0.04 to 0.00)	0.04 ^b (0.01-0.06)
Third quartile	-0.16 ^c (-0.19 to -0.14)	-0.17 ^c (-0.20 to -0.15)	-0.03 ^a (-0.05 to -0.01)	0.04 ^b (0.02-0.06)
Fourth quartile	-0.41 ^c (-0.44 to -0.38)	-0.38 ^c (-0.41 to -0.36)	-0.10 ^c (-0.13 to -0.07)	-0.07 ^c (-0.10 to -0.04)
p-trend	<0.0001	<0.0001	<0.0001	0.00076
Asian immigrants, % change in prevalence (95% CI)				
First quartile	Reference	Reference	Reference	Reference
Second quartile	-0.10 ^c (-0.13 to -0.07)	-0.09 ^c (-0.12 to -0.06)	-0.06 ^b (-0.09 to -0.03)	-0.04 ^b (-0.06 to -0.01)
Third quartile	-0.16 ^c (-0.19 to -0.13)	-0.14 ^c (-0.17 to -0.11)	-0.07 ^c (-0.10 to -0.04)	-0.05 ^b (-0.08 to -0.02)
Fourth quartile	-0.29 ^c (-0.32 to -0.26)	-0.23 ^c (-0.26 to -0.20)	-0.06 ^b (-0.09 to -0.02)	-0.09 ^c (-0.12 to -0.06)
p-trend	<0.0001	<0.0001	0.00016	<0.0001
Hispanic immigrants, % change in prevalence (95% CI)				
First quartile	Reference	Reference	Reference	Reference
Second quartile	-0.04 ^a (-0.07 to -0.01)	-0.02 (-0.05 to 0.01)	0.02 (-0.01 to 0.04)	0.00 (-0.02 to 0.03)
Third quartile	-0.09 ^c (-0.12 to -0.06)	-0.06 ^c (-0.09 to -0.04)	0.01 (-0.02 to 0.04)	-0.01 (-0.03 to 0.02)
Fourth quartile	-0.39 ^c (-0.43 to -0.35)	-0.32 ^c (-0.36 to -0.29)	-0.15 ^c (-0.19 to -0.12)	-0.15 ^c (-0.18 to -0.12)
p-trend	<0.0001	<0.0001	<0.0001	<0.0001

ZCTA, ZIP code tabulated area. Coefficients of association (absolute % change in stroke prevalence per quartile) with 95% CI depicted. Model 1 adjusted for age, sex and ZCTA sociodemographic factors (median household income, education, and rurality). Model 2 adjusted for age and sex, ZCTA sociodemographic factors, and health access factors (health insurance, annual checkup and cholesterol screening). Model 3 adjusted for age and sex, ZCTA sociodemographic factors, health access factors, and obesity. Model 4 adjusted for age and sex, ZCTA sociodemographic factors, health access factors, and health behaviors (e.g., smoking, activity level). ^aDenotes significance at p = 0.05. ^bDenotes significance at p = 0.01. ^cDenotes significance at p < 0.0001.

Table 3: Association between ZCTA immigrant fraction and stroke prevalence.

(mediated) effects of exogenous variables on outcomes, no statement of causality can be presumed (as this is a cross-sectional study). Path analysis for CHD prevalence is depicted in Fig. 1. Smoking ($\beta = 7.9$, $p < 0.0001$), inactivity ($\beta = 14.8$, $p < 0.0001$), and obesity ($\beta = 1.3$, $p < 0.0001$) demonstrated positive associations with CHD prevalence. Immigrant fraction demonstrated a subtle though significant direct inverse association with CHD prevalence ($\beta = -1.4$, $p < 0.0001$); that is to say, there is a 0.014% decrease in absolute CHD prevalence for every 1% increase in immigrant fraction. This corresponds to a relative decrease of ~2% in CHD prevalence for every 10% increase in immigrant fraction. In addition to the direct association, immigrant fraction demonstrated significant indirect associations mediated through smoking ($\beta = -0.16$, $p < 0.0001$), inactivity ($\beta = -0.02$, $p < 0.0001$), and obesity ($\beta = -0.17$, $p < 0.0001$). When calculating the cumulative weight of direct vs indirect paths, we found that the direct path constituted ~45%, and the cumulative indirect paths constituted ~55% of the total.

When performing path analysis for stroke prevalence (Fig. 2), smoking ($\beta = 3.4$, $p < 0.0001$), inactivity ($\beta = 8.1$, $p < 0.0001$), and obesity ($\beta = 2.4$, $p < 0.0001$) demonstrated positive associations with stroke prevalence. For stroke, no significant direct path between immigrant fraction was observed; however, significant indirect paths mediated through inverse associations between immigrant fraction and smoking ($\beta = -0.16$, $p < 0.0001$),

inactivity ($\beta = -0.02$, $p < 0.0001$), and obesity ($\beta = -0.17$, $p < 0.0001$) were found. Therefore in the stroke model, the majority of the association of immigrant fraction on disease prevalence was mediated indirectly.

Discussion

Much prior research within immigrant health has focused on the healthy immigrant effect, the seemingly paradoxical observation that immigrants exhibit better health outcomes compared to native-born individuals despite unfavorable socioeconomic conditions.^{21,22} By contrast, relatively sparse literature has been directed at understanding possible externalities immigrants may create for their communities, such as through positively influencing health behaviors of their families and friends. In this study, we utilized a newly-available health behavior and outcomes data set to determine the association between local area immigrant fraction and the prevalence of CHD and stroke on a national level across the US, and with more granular and precise geographic information than was feasible in prior work. The central finding of this study is that increasing immigrant fraction is associated with decreasing local prevalence of CHD, and to a lesser extent stroke. These findings remained consistent when analyzing the two largest immigrant subgroups, Asians and Hispanics. For CHD, there exists a direct contribution of immigrant fraction on disease prevalence, which constitutes

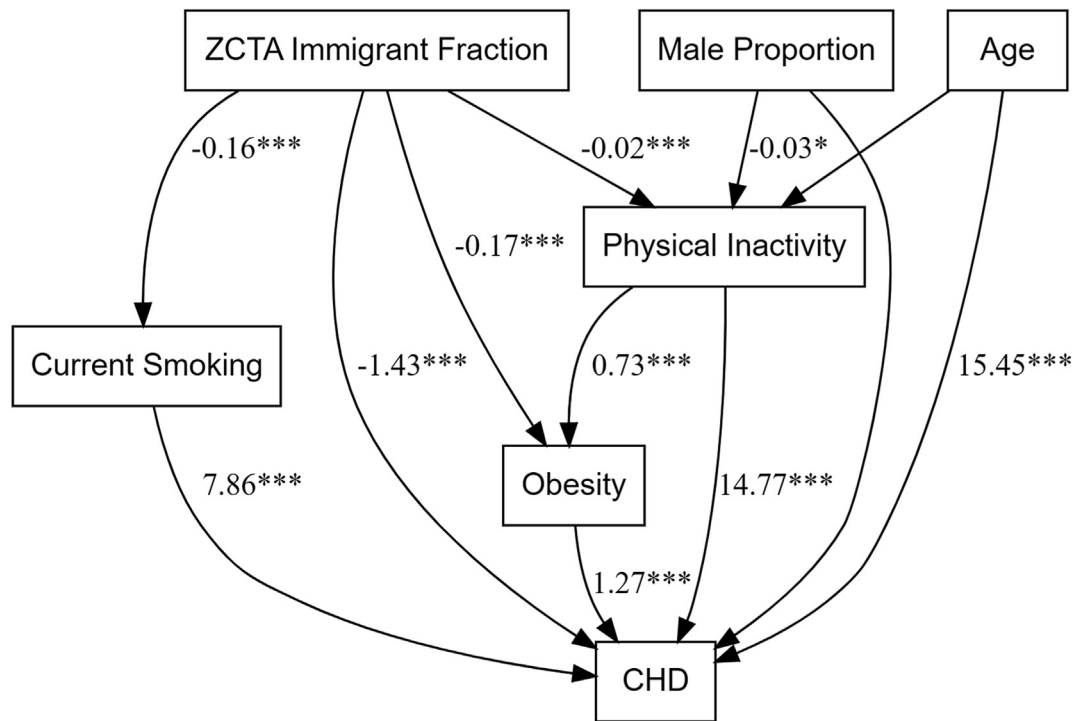


Fig. 1: Multilevel path analysis demonstrating the strength and sign of directional relationships between independent variables (immigrant fraction, male proportion, proportion >65 years old), endogenous behavioral variables (smoking, physical inactivity, obesity), and coronary heart disease (CHD). Lack of path coefficient indicates no significant association. *Denotes significance at $p = 0.05$. ***Denotes significance at $p < 0.0001$.

45% of the total association; by contrast, there also exist significant indirect associations (~55%) mediated through lower prevalence of deleterious health behaviors.

We believe this study to be one of the first to study immigrant fraction at such a detailed geographic level (ZCTA) across the entirety of the US. Prior national database studies focused on immigrants have mostly used larger administrative subdivisions such as counties or commuting zones.⁷ Conversely, smaller geographic or administrative area studies have been mostly performed within single cities or a small number of metropolitan areas.^{5,13,23,24} The findings from our study are both consistent with these prior studies, and expand the literature by providing both neighborhood granularity and generalizability across the diverse landscape of the US. Using national tax records constituting over 1.4 billion person-years of observation, Chetty et al. found local area immigrant fraction (defined at the level of commuting zone, an aggregation of counties) to be the strongest correlate of local life expectancy ($r = 0.72$).⁷ While the absolute reduction in cardiovascular disease prevalence was modest in our study, the relative reduction in CHD prevalence between the highest- and lowest-quartiles was 15% in the base model, and 5% in the fully-adjusted model. To put this in perspective, for

an individual it has been estimated that lowering systolic blood pressure <130 mm Hg reduces 10-year CHD risk by ~35%, and lowering low-density lipoprotein (by at least 30%) reduces CHD risk by ~25%.²⁵

Our study also importantly decomposed the direct and indirect contributions of local immigrant fraction on cardiovascular disease prevalence at the community level. While there appeared a modest direct association of immigrant fraction on CHD, the majority of the contribution appeared mediated through health behaviors. For stroke, the association of immigrant fraction was completely mediated through indirect effects of health behaviors. Others have found similar indirect effects. A cross-sectional study utilizing data from the Los Angeles County Health Survey 2007 linked to 2000 US Census data measured the effects of acculturation (defined by language spoken at home and years since immigration) on unhealthy dietary patterns (defined as fast food eating or low fruit/vegetable consumption).¹³ This study found that living in a neighborhood with a high density of other immigrants attenuates the negative effects of acculturation on healthy eating behaviors. The protective effects of living in high-immigrant communities may also extend to native-born Americans. A separate study of metropolitan Los Angeles based on ZCTA demonstrated that neighborhood

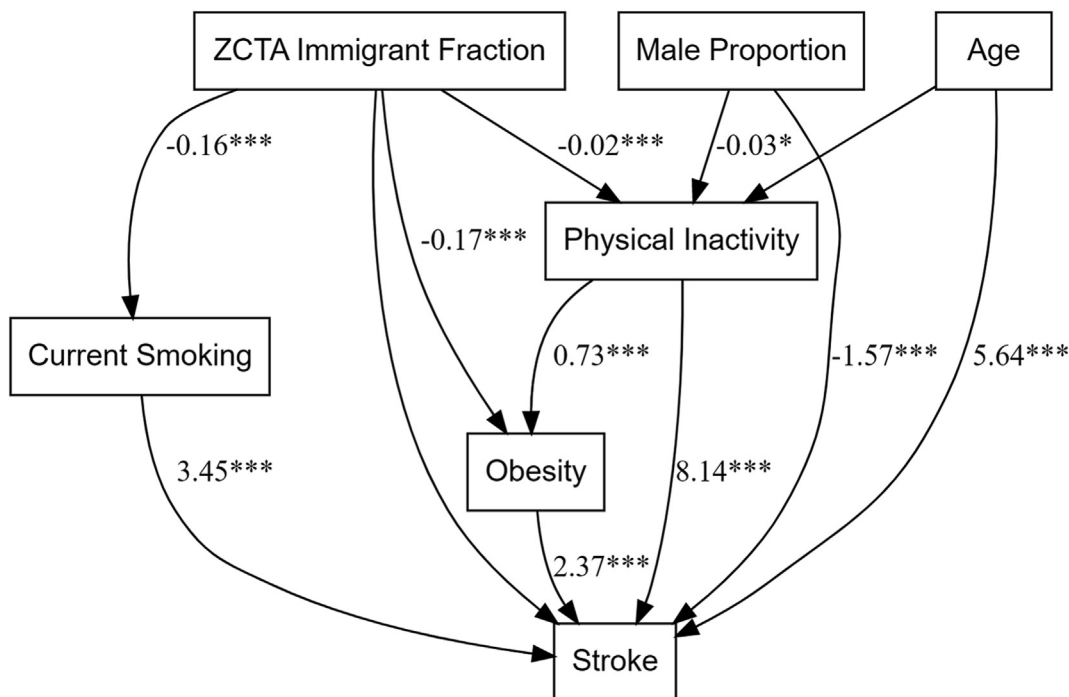


Fig. 2: Multilevel path analysis demonstrating the strength and sign of directional relationships between independent variables (immigrant fraction, male proportion, proportion >65 years old), endogenous behavioral variables (smoking, physical inactivity, obesity), and stroke. Lack of path coefficient indicates no significant association. *Denotes significance at $p = 0.05$. ***Denotes significance at $p < 0.0001$.

immigrant density was negatively associated with regular fast-food consumption, BMI, and hypertension, and positively associated with fruit/vegetable consumption among native-born Americans.⁵ Another study of low-income women demonstrated increases in the consumption of fruits/vegetables with increases in local area immigrant population.²⁶ Immigrant acculturation may also be an important risk factor for smoking by their children.^{27,28} Our study builds on these prior risk-factor studies by not only validating differences in health behaviors between high- and low-immigrant communities, but also demonstrating that these differences in behavior may contribute to differences in major cardiovascular disease prevalence. Notably, the decomposition of direct and indirect effects in this study cannot be construed to imply causality.

Our study separately analyzed immigrant subgroups by origin (Hispanic and Asian). There have been few studies that have reported on immigrants by disaggregated racial/ethnic background. A study based on the Multiethnic Study of Atherosclerosis cohort examined whether neighborhood immigrant composition was associated with health behaviors among Hispanic (recruited from New York, Los Angeles, and St Paul) and Chinese Americans (recruited from Los Angeles and Chicago).²⁹ After adjustment for relevant neighborhood sociodemographic characteristics, living in a

tract with a higher proportion of immigrants was associated with lower consumption of high-fat foods among Hispanics and Chinese, but with being less physically active among Hispanics. Another cross-sectional study using data from the Medical Expenditure Panel Survey linked to the US Census found living in communities of high Hispanic density to be associated with increased obesity among both Hispanics and non-Hispanic Whites; conversely, living in communities of high Asian density was associated with reduced obesity among non-Hispanic Whites.³⁰ Our study similarly found living in communities of high Hispanic-origin immigrant fraction to be correlated with obesity, and high Asian-origin immigrant fraction to be inversely correlated with obesity. However, our study suggests that once accounting for these differences, Americans living in both Asian- and Hispanic-immigrant-predominant neighborhoods may benefit from positive externalities. There may be some common underlying protective mechanism against cardiovascular diseases mediated through anti-acculturation.

Our study has limitations. One important limitation is ecological bias. The independent and dependent variables were not measured at the level of the individual, but rather at the level of the ZCTA. We therefore could not decompose CHD or stroke prevalence between immigrants and non-immigrants (or between racial groups)

within a ZCTA. We also could not decompose health behavioral patterns or demographic profiles within a ZCTA by group. Our study was cross-sectional in nature. This is in part because local area markers of disease prevalence and health behavior at the level of ZCTA were only recently made available through data sets such as PLACES. While direct vs indirect effects were decomposed through structural equation modeling, no inference about causality can be made. We chose to exclude low-immigrant communities in our main analyses. We did this for a number of reasons, including inadequate sampling, missingness of data, and lack of similarity between low-immigrant ZCTAs and our analytic cohort. These excluded ZCTAs (~16,000) were found predominantly in rural communities throughout the US. As such, the generalizability of our findings to less populated and more rural regions of the US may be limited. A related limitation is that while ACS (~3.5 million households sampled annually) and BRFSS (~400,000 adults sampled annually) are very large surveys, they may not be adequate for small population clusters and may be prone to error for even larger population clusters. Certain potential mediators or confounders of the relationship between exposure and outcomes were not included in the model—these included potentially important mediators such as hypertension and diabetes. While we did include certain proxies for these conditions (e.g., prevalence of cholesterol screening), the indirect nature of the data is a limitation.

With these limitations in mind, this study is one of the first to demonstrate in a large, national sample the association between immigrant fraction and both reduced prevalence of deleterious behaviors, and reduced prevalence of major cardiovascular disease at the community level.

Contributors

DS led analysis and manuscript writing. NW, NN, EB performed data analysis and contributed to manuscript writing. OT provided statistical oversight and contributed to manuscript writing. LPP procured funding and provided oversight of the project. RJH procured funding and co-lead data analysis and manuscript writing. All authors approved the final, submitted version of this manuscript. DS and RJH have directly accessed and verified the data reported in the manuscript. RJH had final responsibility to submit the study for publication.

Data sharing statement

ZCTA-level demographic data from ACS and health behavioral and outcomes data from BRFSS/PLACES are publicly available for access.

Editorial disclaimer

The Lancet Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

Declaration of interests

All funding sources have been acknowledged in the above section. The authors declare no actual or potential conflicts of interest.

Acknowledgements

The authors would like to thank Mathew V Kiang, ScD for methodological and statistical advice and guidance.

Sources of funding: RJH received support from the NIH/NCI under Award Number K08CA252635, the Rutgers University Asian Resource Centers for Minority Aging Research Center under NIH/NIA Grant P30-AG0059304, and the Stanford University Department of Medicine Chair Diversity Investigator Award. LPP is supported by NIH/NHLBI grant K24HL150476. NW reports conference and travel support from the Stanford Center for Asian Health Research and Education. The content is solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.lana.2023.100613>.

References

- Budiman A. *Key findings about U.S. Immigrants Pew Research Center*; 2020. <https://www.pewresearch.org/hispanic/2020/08/20/facts-on-u-s-immigrants/>. Accessed September 1, 2022.
- Alese JT, Alese OB. Comparison of physical activity as an obesity-related behavior between immigrants and native-born US adults. *Palgrave Commun*. 2020;6(1):65.
- Kennedy S, Kidd MP, McDonald JT, Biddle N. The healthy immigrant effect: patterns and evidence from four countries. *J Int Migrat Integrat*. 2015;16(2):317–332.
- Kim H, Cao C, Grabovac I, et al. Associations in physical activity and sedentary behaviour among the immigrant and non-immigrant US population. *J Epidemiol Community Health*. 2020;74(8):655–661. <https://jech.bmj.com/content/74/8/655>.
- Shi L, Zhang D, Rajbhandari-Thapa J, Katapodis N, Su D, Li Y. Neighborhood immigrant density and population health among native-born Americans. *Prev Med*. 2019;127:105792.
- Tu JV, Chu A, Rezai MR, et al. The incidence of major cardiovascular events in immigrants to Ontario, Canada: the CANHEART immigrant study. *Circulation*. 2015;132(16):1549–1559.
- Chetty R, Stepner M, Abraham S, et al. The association between income and life expectancy in the United States, 2001–2014. *JAMA*. 2016;315(16):1750–1766.
- Duffey KJ, Gordon-Larsen P, Ayala GX, Popkin BM. Birthplace is associated with more adverse dietary profiles for US-born than for foreign-born Latino adults. *J Nutr*. 2008;138(12):2428–2435.
- Fang J, Yuan K, Gindi RM, Ward BW, Ayala C, Loustalot F. Association of birthplace and coronary heart disease and stroke among US adults: national health interview survey, 2006 to 2014. *J Am Heart Assoc*. 2018;7(7):e008153. <https://www.ahajournals.org/doi/full/10.1161/JAHA.117.008153>.
- McDonald JT, Kennedy S. Insights into the ‘healthy immigrant effect’: health status and health service use of immigrants to Canada. *Soc Sci Med*. 2004;59(8):1613–1627.
- de Castro AB, Hing AK, Lee NR, et al. Cohort profile: the Health of Philippine Emigrants Study (HoPES) to examine the health impacts of international migration from the Philippines to the USA. *BMJ Open*. 2019;9(11):e032966. <https://bmjopen.bmj.com/content/9/11/e032966>.
- Martinez-Cardoso AM, Geronimus AT. The weight of migration: reconsidering health selection and return migration among Mexicans. *Int J Environ Res Public Health*. 2021;18(22):12136.
- Zhang D, van Meijgaard J, Shi L, Cole B, Fielding J. Does neighbourhood composition modify the association between acculturation and unhealthy dietary behaviours? *J Epidemiol Community Health*. 2015;69(8):724–731.
- Census.gov. *American community survey 2014–2018 5-year estimates: Census.gov*; 2019. <https://www.census.gov/newsroom/press-releases/2019/acs-5-year.html>. Accessed September 1, 2021.
- CDC-BRFSS. *Behavioral risk factor surveillance system: CDC*; 2020. <https://www.cdc.gov/brfss/index.html>. Accessed September 1, 2021.
- Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci*. 2010;25(1):1–21.
- Jackson JW, Schmid I, Stuart EA. Propensity scores in pharmacoepidemiology: beyond the horizon. *Curr Epidemiol Rep*. 2017;4(4):271–280.
- de Heer HD, Balcazar HG, Castro F, Schulz L. A path analysis of a randomized promotora de salud cardiovascular disease-prevention trial among at-risk Hispanic adults. *Health Educ Behav*. 2012;39(1):77–86.

- 19 Kerkhof GF, Duivenvoorden HJ, Leunissen RW, Hokken-Koelega AC. Pathways leading to atherosclerosis: a structural equation modeling approach in young adults. *Hypertension*. 2011;57(2):255–260.
- 20 Pichon LC, Arredondo EM, Roesch S, Sallis JF, Ayala GX, Elder JP. The relation of acculturation to Latinas' perceived neighborhood safety and physical activity: a structural equation analysis. *Ann Behav Med*. 2007;34(3):295–303.
- 21 Aldridge RW, Nellums LB, Bartlett S, et al. Global patterns of mortality in international migrants: a systematic review and meta-analysis. *Lancet*. 2018;392(10164):2553–2566.
- 22 Elshahat S, Moffat T, Newbold KB. Understanding the healthy immigrant effect in the context of mental health challenges: a systematic critical review. *J Immigr Minor Health*. 2021;24(6):1564–1579.
- 23 DeCamp LR, Choi H, Fuentes-Afflick E, Sastry N. Immigrant Latino neighborhoods and mortality among infants born to Mexican-origin Latina women. *Matern Child Health J*. 2015;19(6):1354–1363.
- 24 Tam CC, Lui CK, Karriker-Jaffe KJ. Intersections of neighborhood co-ethnic density and nativity status on heavy drinking in a general population sample of US Latinos and Asians. *Alcohol Alcohol*. 2021;56(1):74–81.
- 25 Pencina MJ, Navar AM, Wojdyla D, et al. Quantifying importance of major risk factors for coronary heart disease. *Circulation*. 2019;139(13):1603–1611.
- 26 Dubowitz T, Subramanian SV, Acevedo-Garcia D, Osypuk TL, Peterson KE. Individual and neighborhood differences in diet among low-income foreign and U.S.-born women. *Womens Health Issues*. 2008;18(3):181–190.
- 27 O'Loughlin J, Maximova K, Fraser K, Gray-Donald K. Does the "healthy immigrant effect" extend to smoking in immigrant children? *J Adolesc Health*. 2010;46(3):299–301.
- 28 Bosdriesz JR, Lichthart N, Witvliet MI, Busschers WB, Stronks K, Kunst AE. Smoking prevalence among migrants in the US compared to the US-born and the population in countries of origin. *PLoS One*. 2013;8(3):e58654. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0058654>.
- 29 Osypuk TL, Diez Roux AV, Hadley C, Kandula NR. Are immigrant enclaves healthy places to live? The multi-ethnic study of atherosclerosis. *Soc Sci Med*. 2009;69(1):110–120.
- 30 Kirby JB, Liang L, Chen H-J, Wang Y. Race, place, and obesity: the complex relationships among community racial/ethnic composition, individual race/ethnicity, and obesity in the United States. *Am J Public Health*. 2012;102(8):1572–1578.