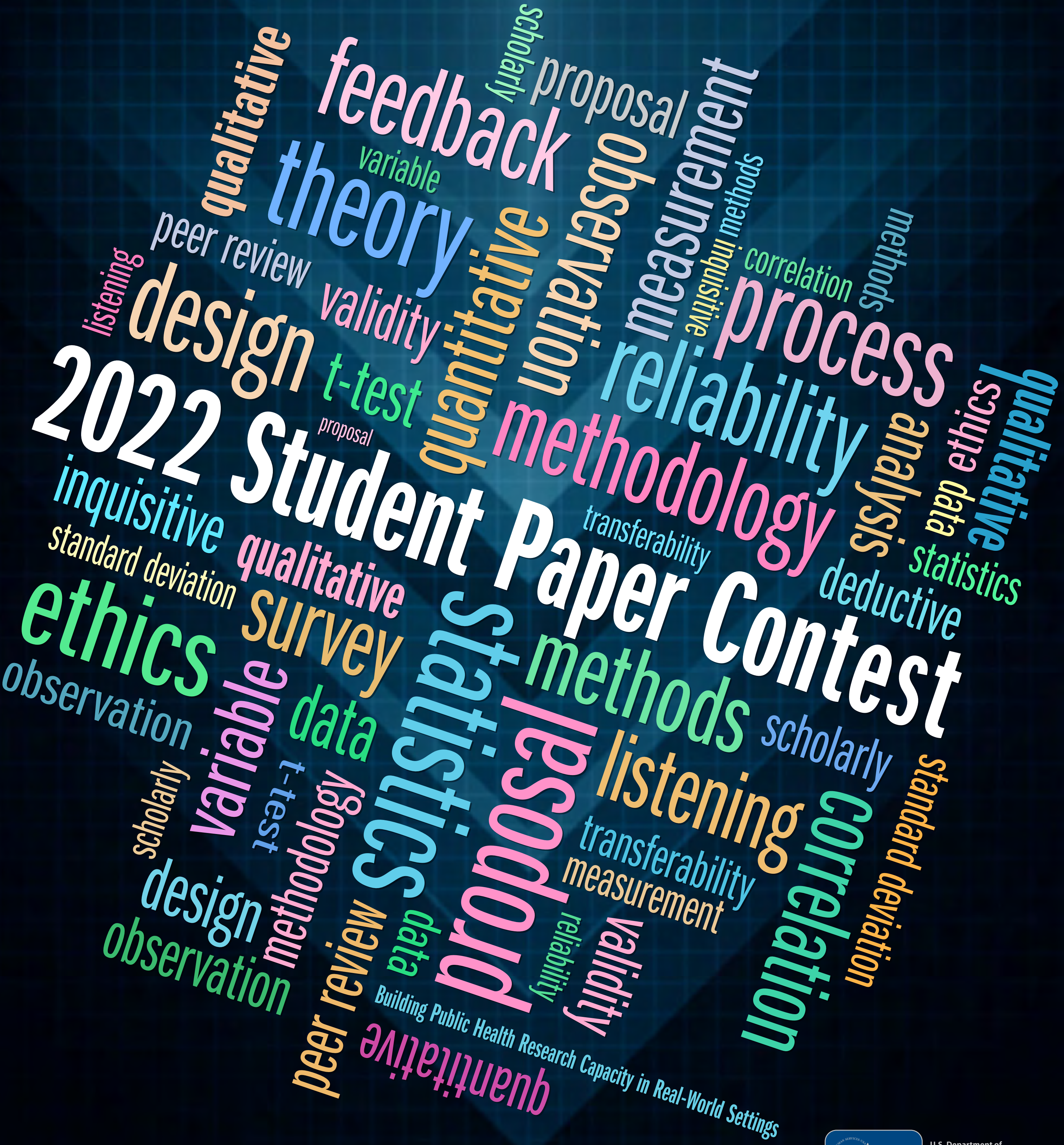
PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY

2022 Student Paper Contest



Building Public Health Research Capacity in Real-World Settings

U.S. Department of
Health and Human Services
Centers for Disease
Control and Prevention

About the Journal

Preventing Chronic Disease (PCD) is a peer-reviewed public health journal sponsored by the Centers for Disease Control and Prevention and authored by experts worldwide. PCD was established in 2004 by the National Center for Chronic Disease Prevention and Health Promotion with a mission to promote dialogue among researchers, practitioners, and policy makers worldwide on the integration and application of research findings and practical experience to improve population health.

PCD's vision is to serve as an influential journal in the dissemination of proven and promising peer-reviewed public health findings, innovations, and practices with editorial content respected for its integrity and relevance to chronic disease prevention.

PCD Staff

Associate Editors

Table of Contents

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
Volume 19, E73 NOVEMBER 2022

EDITORIAL

Maintaining Our Commitment to Building Scientific Publishing Capacity of Students: PCD's 2022 Student Paper Contest Winners and Release of 2023 Call for Papers

Leonard Jack Jr, PhD, MSc[1]

Accessible Version: www.cdc.gov/pcd/issues/2022/22_0303.htm

Suggested citation for this article: Jack L Jr. Maintaining Our Commitment to Building Scientific Publishing Capacity of Students: PCD's 2022 Student Paper Contest Winners and Release of 2023 Call for Papers. Prev Chronic Dis 2022;19:220303. DOI: https://doi.org/10.5888/pcd19.220303.

NON–PEER REVIEWED

Preventing Chronic Disease (PCD) initiated its first Student Paper Contest in 2011 (1). Since that time, the journal has received hundreds of student research papers from around the world. Last year, PCD celebrated its 10-year anniversary of building scientific publishing skills and abilities among students (2). Research mentors have used this journal as an avenue to introduce their students to the rigors of generating scholarly writings that focus on conceptualizing research; collecting, analyzing, and reporting data; and discussing the public health implications of research findings. Since 2011, regardless of whether papers are accepted, PCD has provided students with extensive feedback on their submissions.

Goals and Submission Requirements

Eligibility for PCD's student paper contest has evolved over the years (2). Participants must be currently enrolled students or have completed one of the following programs within the last 12 months: high school, undergraduate or graduate degrees, medical residency, or a postdoctoral fellowship conducted under the supervision of a principal investigator or research advisor. PCD also requires that the student author serve as the first and corresponding author. PCD only accepts original research that has not been published previously or submitted elsewhere for publication. Papers submitted for consideration must use one of two PCD article types: Original Research or GIS Snapshots.

Over the years, PCD has refined the purpose of the student research publication opportunity to include these 5 primary goals (3):

- Provide students with an opportunity to become familiar with a journal's manuscript submission requirements and peer-review process
- Assist students in connecting their knowledge and training on conducting quality research with a journal's publication expectations
- Develop students' research and scientific writing skills to become producers of knowledge rather than just consumers of knowledge
- Provide students with an opportunity to become a first author on a peer-reviewed article
- Promote supportive, respectful, and mutually beneficial student–mentor relationships that strengthen students' ability to generate and submit scholarly manuscripts throughout their professional careers

2022 Winners and Submissions

Sixteen student research papers were submitted for the contest in 2022, and all submissions underwent the same peer-review process as any other manuscript submitted to the journal for consideration. Eight of the 16 papers were successful in making it through a rigorous review process before being accepted for publication (4–11). Student research papers addressed a range of topics: COVID-19 and various aspects of health, food insecurities among caregivers in southern states, the relationship between physical activity and depression among high school students, perceptions of neighborhood development on active living among community residents, use of a cancer index as a predictor of common cancers, and spatial analysis of breast cancer mortality rates in rural states. PCD is pleased to announce 2022 student paper winners in 4 categories: high school, undergraduate, master's degree, and doctoral degree. After careful review, PCD did not select a winner in the postdoctoral category.

In the high school category, Wang and coauthors of the article "Association Between Physical Activity and Sedentary Behavior With Depressive Symptoms Among US High School Students, 2019" point out that depression among high school students has increased over the past decade (4). Their research analyzed 2019 data from the Youth Risk Behavior Survey that consisted of 13,526 high school students. They found that increased exposure to computer or device use was associated with higher odds of inadequate physical activity and excessive sedentary behavior and depressive symptoms. Their findings suggest the importance of public health interventions that focus on implementing physical activity to reduce risk of depression among high school students.

This year's winning paper in the undergraduate category was generated by Creech et al: "Physical Activity Among Adults in Rural Western North Carolina During the COVID-19 Pandemic" (5). This article examines the locations, reasons, benefits, and barriers to engaging in physical activity among adults living in rural communities during the COVID-19 pandemic. Using nonparametric measures, authors identified that, among a sample of 297 adults, most engaged in physical activity at home, in parks, and in neighborhoods. Primary reasons for participating in physical activity included getting out of the house, maintaining mental health, and engaging in some form of physical activity.

Schulz and colleagues' article, "Spatial Analysis of Breast Cancer Mortality Rates in a Rural State," is the winner in the master's degree category (6). This article reports findings from a study in South Dakota that assessed what sociodemographic factors contribute to mortality rates and, using spatial analysis, explored how counties' observed age-adjusted mortality rates compared with expected rates. A linear regression model was used to identify sociodemographic factors associated with breast cancer mortality rates and to compute new standardized incidence ratios (SIRs), after controlling for significant factors affecting mortality. Findings indicated that educational level and breast cancer incidence rates were significant factors associated with breast cancer mortality rates at the county level. Authors also used the SIR model to show the spatial distribution of mortality rates by county.

Seto and coauthors of the article "Differences in COVID-19 Hospitalizations by Self-Reported Race and Ethnicity in a Hospital in Honolulu, Hawaii" were selected as winners in the doctoral category (7). Pointing out that COVID-19 has exacted a tremendous toll on racial and ethnic groups in the US, they sought to identify the extent to which race and ethnicity were misclassified in COVID-19 hospitalizations. They assessed the responses of 847 patients at randomly selected hospital and ambulatory units who completed a survey that asked them to self-identify their race and ethnicity and compared their responses with data in electronic medical records (EMRs). Authors found that using self-identified data on race and ethnicity rather than hospital EMR data may help uncover further disparities in COVID-19 hospitalizations.

PCD congratulates this year's impressive winners. In addition, we invite you to join us in celebrating all student authors who submitted manuscripts to the contest, regardless of whether their manuscripts were accepted. We hope all student authors who submitted papers gained a tremendous amount of experience serving as first and corresponding authors and gained more knowledge of journal submission guidelines and in the correspondence process with the journal's editor in chief, peer reviewers, associate editors, and technical editors.

In closing, PCD is proud to release its 2023 Student Paper Contest call for papers. PCD is interested in student research papers relevant to the prevention, screening, surveillance, or population-based intervention of chronic diseases, including but not limited to arthritis, asthma, cancer, depression, diabetes, obesity, and cardiovascular disease. PCD is also interested in papers in these areas that explore the role of social determinants of health to include the impact of racism in shaping health outcomes. Readers are asked to encourage students at the high school, undergraduate, master's, and doctoral levels, as well as medical residents and postdoctoral fellows, to submit a student research paper to the journal for consideration. Students are encouraged to submit an inquiry (www.cdc.gov/pcd/for_authors/submit_inquiry.htm) to the journal before the submission deadline to determine whether a given topic area and research focus align with the intent of the contest. For more information about the journal and previous collections of student papers, please visit the PCD website at www.cdc.gov/pcd.

Author Information

Corresponding Author: Leonard Jack, Jr, PhD, MSc, Editor in Chief, Preventing Chronic Disease, Centers for Disease Control and Prevention, 4770 Buford Hwy, MS S107–8, Atlanta, GA 30341. Email: ljj0@cdc.gov.

Author Affiliations: [1]Preventing Chronic Disease, Office of Medicine and Science, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, Georgia.

References

1. Posner SF. PCD's first annual student research contest: Lui and Wallace examine hospitalization rates for at-risk populations. Prev Chronic Dis 2011;8(5):A103.

2. Jack L Jr. PCD 2021 student research collection: building public health research capacity in real-world settings and the 2022 call for papers. Prev Chronic Dis 2021;18:E68.
3. Jack L Jr. Shaping future generations of public health researchers: Preventing Chronic Disease's student research paper contest. Prev Chronic Dis 2017;14:E96.
4. Wang C, Peiper N. Association between physical activity and sedentary behavior with depressive symptoms among US high school students, 2019. Prev Chronic Dis 2022;19:E76.
5. Creech W, Towner B, Battista R. Physical activity among adults in rural western North Carolina during the COVID-19 pandemic. Prev Chronic Dis 2022;19:E74.
6. Schulz M, Spors E, Bates K, Michael S. Spatial analysis of breast cancer mortality rates in a rural state. Prev Chronic Dis 2022;19:220113.
7. Seto B, Nishizaki L, Akaka G, Kimura J, Seto T. Differences in COVID-19 hospitalizations by self-reported race and ethnicity in a hospital in Honolulu, Hawaii. Prev Chronic Dis 2022;19:E72.
8. Goswami S, Korgaonkar S, Bhattacharya K, Rosenthal M. Food insecurity in a sample of informal caregivers in 4 southern US states. Prev Chronic Dis 2022;19:E51.
9. Tsuzaki K, Taira D. Forgone health care for non–COVID-19–related needs among Medicare beneficiaries during the COVID-19 pandemic, summer 2020–winter 2021. Prev Chronic Dis 2022;19:220110.
10. Dsouza N, Serrano N, Watson KB, McMahon J, Devlin HM, Lemon SC, et al. Exploring residents' perceptions of neighborhood development and revitalization for active living opportunities. Prev Chronic Dis 2022;19:E56.
11. Guo L, Wright M, Osias M, Vaezi M, Hughes M. Creation and evaluation of the Illinois cancer risk index as a predictor of four common cancers. Prev Chronic Dis 2022;19:E75.

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
Volume 19, E76 NOVEMBER 2022

ORIGINAL RESEARCH

Association Between Physical Activity and Sedentary Behavior With Depressive Symptoms Among US High School Students, 2019

Charles H. Wang[1]; **Nicholas Peiper, PhD, MPH**[2]

Accessible Version: www.cdc.gov/pcd/issues/2022/22_0003.htm

Suggested citation for this article: Wang CH, Peiper N. Association Between Physical Activity and Sedentary Behavior With Depressive Symptoms Among US High School Students, 2019. Prev Chronic Dis 2022;19:220003. DOI: https://doi.org/10.5888/pcd19.220003.

PEER REVIEWED

Summary

What is already known on this topic?

The prevalence of depression among US adolescents has increased. Depression has severe and potentially lasting effects. Meanwhile, sedentary behaviors, such as screen time, limit physical activity and are common among US adolescents.

What is added by this report?

We used a large, nationally representative sample of US high school students to demonstrate that being physically active and spending less time on a computer or digital device are significantly associated with reduced odds of reporting depressive symptoms.

What are the implications for public health practice?

Because adolescence is a crucial time for establishing positive health habits, intervention strategies designed to promote physical activity and reduce sedentary behaviors among high school students may be warranted to improve the mental health and the general well-being of this population.

Abstract

Introduction

The prevalence of depression among US adolescents has increased during the past decade. Previous studies found relationships among physical activity, sedentary behavior, and depression, but more recent information is needed to inform research and practice. We used national surveillance data to assess the association of physical activity and sedentary behavior with depressive symptoms among US high school students.

Methods

This study included 13,526 high school students from the 2019 Youth Risk Behavior Survey. The dependent variable was the presence of depressive symptoms in the past year that lasted almost every day for at least 2 weeks in a row and interfered with usual activities. The independent variables were physical activity (overall activity, muscle-strengthening exercises, participation on sports teams) and sedentary behavior (watching television, using a computer or digital device). We used weighted multivariable logistic regression to evaluate the association of physical activity and sedentary behavior variables with depressive symptoms, while controlling for demographic characteristics and other health behaviors.

Results

The prevalence of depressive symptoms was 36.7%. Participating in physical activity 5 or more days in the past 7 days (adjusted odds ratio [aOR], 0.81; 95% CI, 0.68–0.97) and participating on 1 or more sports teams in the past year (aOR, 0.66; 95% CI, 0.55–0.78) were associated with reduced odds of depressive symptoms. Using a computer or digital device for 3 or more hours per school night was associated with higher odds of reporting depressive symptoms (aOR, 1.61; 95% CI, 1.41–1.85).

Conclusion

Inadequate physical activity and excessive sedentary behavior are associated with depressive symptoms among US high school students. Interventions targeting physical activity and sedentary behavior may be a public health strategy to reduce depressive symptoms in this population.

Introduction

Depression is a common mental disorder among adolescents that is characterized by various psychological symptoms such as persistent sadness, lack of enjoyment, and deterioration of happiness (1). These symptoms start to become pronounced during the transition to high school in middle adolescence and can lead to lack of sleep, disturbance in appetite, a sedentary lifestyle, and even suicide. Depressive symptoms among adolescents in high school have been increasing since approximately 2012, with the most dramatic increases starting in 2015 (2,3). While increases in depressive symptoms have been noted across many age groups since 2012, the prevalence of past-year major depressive episodes increased most rapidly among adolescents in high school (aged 14–17 y) (4,5). These recent trends may be indicative of emergent risk factors and mental health problems among adolescents in high school.

Although considerable research has focused on psychological risk factors such as stress and traumatic childhood experiences, the relationships among physical activity, sedentary behaviors, and depression is a growing area of interest (6,7). National surveillance data indicate that the percentage of adolescents in high school who were physically active for at least 60 minutes per day on 5 or more days in the past week decreased from 2011 to 2019 (8). Similarly, sedentary behaviors such as using digital devices for something other than schoolwork, which limits physical activity, also became more common among adolescents in high school during this time (9,10). According to the American Academy of Child and Adolescent Psychiatry, children aged 8 to 12 years spend roughly 4 to 6 hours per day observing a screen, and adolescents in high school often spend up to 9 hours (11). Excessive screen time can lead to inactivity, which then can lead to weight problems, self-image issues, depressive symptoms, and poor health-related quality of life (12,13). More recently, the COVID-19 pandemic led to decreases in opportunities for regular physical activity and increases in depressive symptoms (14).

Although many factors contribute to adolescents' mental well-being, dietary behavior has been gaining attention as a key factor in the prevention and management of depression (15). In addition, depressive symptoms among adolescents may also be related to factors such as substance use and behaviors that contribute to violence (eg, feeling unsafe at school, fighting, bullying, cyberbullying), although the relationships may be bidirectional (16–21). For example, depression is associated with substance use among males during adolescence and young adulthood (21), and exposure to acute and chronic stressful events, such as bullying by peers and maltreatment, can be associated with depression among adolescents (20).

Many public health studies have demonstrated associations among physical activity, sedentary behaviors, and depressive symptoms among high school students, although more recent investigations are needed, given recent increases in depressive symptoms among adolescents in high school (2–5). A meta-analysis in 2017 showed that evaluations of physical activity and depression among US adolescents were gender-specific, focused on only one aspect of physical activity or sedentary behavior, or employed small sample sizes (22). By analyzing data from a large sample of the most recent (2019) Youth Risk Behavior Survey (YRBS), this study provides updated evidence on the associations among multiple aspects of physical activity, sedentary behavior, and depressive symptoms among adolescents. By using YRBS's nationally representative sample, which accurately represents the characteristics of high school students in the US, the results of this study can be generalized to all US students in grades 9 through 12 (23). Our primary objective was to describe the association between different measures of physical activity and sedentary behavior with depressive symptoms among US high school students. We hypothesized that higher levels of physical activity and lower levels of sedentary behavior would be associated with lower levels of depressive symptoms after controlling for demographic characteristics. Our secondary hypothesis was that the associations would hold after additionally controlling for dietary behaviors, substance use, and behaviors that contribute to violence.

Methods

Data source and study sample

The YRBS is a cross-sectional national survey that has been conducted by the Centers for Disease Control and Prevention (CDC) since 1991 to monitor health behaviors among US high school students. The YRBS follows local procedures for obtaining parental permission, and students voluntarily complete the surveys. All responses to the survey are recorded anonymously on a computer-scannable questionnaire booklet. CDC's institutional review board approved this protocol for the national YRBS (23). For this study, we focused on the 2019 national data set that consisted of 13,526 high school students in grades 9 through 12. The 2019 sampling frame included all regular public, Catholic, and private school students, in grades 9 through 12, in the 50 States and the District of Columbia. The school response rate was 75.1%, and the student response rate was 80.3%. Multiplying these 2 percentages (75.1% × 80.3%) together yields an overall response rate of 60.3% (23).

Study measures

The dependent variable of interest was self-reported depressive symptoms (24). The survey item asked if in the past year students ever felt so sad or hopeless almost every day for 2 weeks or more in a row that they stopped doing some usual activities (yes/no). The main independent variables were physical activity and sedentary behavior (25). Physical activity was measured by using 3 dichotomized variables: being physically active for a total of at least 60 minutes per day on 5 or more days in the past 7 days, doing muscle-strengthening exercises on 3 or more days in the past 7 days, and playing on at least 1 sports team during the past 12 months (8). Sedentary behavior was measured by using 2 dichotomized variables: watching 3 or more hours of television on an average school day and using a computer or digital device for 3 or more hours for something that was not schoolwork on an average school day (9). In addition to demographic variables, we included dietary behaviors, substance use, and behaviors contributing to violence as covariates (15,21). Dietary behaviors included dichotomized variables for skipping breakfast on all days in the past 7 days, not eating any fruits in the past 7 days, and not eating any vegetables (eg, green salad, potatoes, carrots, other vegetables) in the past 7 days (8). Dichotomous variables for substance use were cigarette use in the past 30 days, e-cigarette use in the past 30 days, cannabis use in the past 30 days, alcohol use in the past 30 days, lifetime prescription pain medicine misuse, and lifetime illicit drug use (cocaine, inhalants, methamphetamines, heroin, ecstasy, or hallucinogens) (26). Behaviors contributing to violence included dichotomized variables for school safety concerns in the past 30 days, getting into a physical fight in the past 12 months, being bullied at school in the past 12 months, and being cyberbullied in the past 12 months (18,20).

The operationalizations of depressive symptoms, physical activity variables, sedentary behaviors, and other behavioral risk factors were informed by the extant literature and are consistent with standard definitions and reporting practices developed by CDC (8,9,23). Demographic variables included sex (male/female), grades (9–12), and race and ethnicity (American Indian or Alaska Native, Asian American and Native Hawaiian/Other Pacific Islander, Hispanic or Latino, non-Hispanic Black, non-Hispanic White, and non-Hispanic multiracial). Details about the established reliability and validity of the YRBS are available elsewhere (24,25,27).

Statistical analysis

Our initial analyses examined weighted frequencies for depressive symptoms, physical activity, sedentary behavior, dietary behavior, substance use, behaviors contributing to violence, and demographic characteristics among US high school students reporting their grade (N = 13,526). We then conducted cross-tabulations to examine bivariate differences in the prevalence of depressive symptoms among demographic groups, physical activity variables, sedentary behaviors, dietary behaviors, substance use, and behaviors that contribute to violence. We conducted pairwise comparisons with Bonferroni correction to determine differences in the prevalence of depressive symptoms within demographic groups. To test the primary hypothesis, we used a multivariable logistic regression model to evaluate the relationship between depressive symptoms with physical activity variables and sedentary behaviors while controlling for demographic variables. For the secondary hypothesis, we used the first model and added dietary behaviors, substance use, and behaviors that contribute to violence as covariates. We calculated adjusted odds ratios (aORs) and 95% CIs for demographic characteristics, physical activity variables, sedentary behaviors, dietary behaviors, substance use, and behaviors that contribute to violence. All analyses incorporated the sample weights and poststratification variables to account for the complex sampling methods of the YRBS. A 2-sided P value of $\leq .05$ denoted significance. We used Stata version 15.1 (StataCorp LLC) for all analyses.

Results

Overall, students were distributed approximately equally across grades and sex (Table 1). Slightly more than half (51.2%) were non-Hispanic White. The physical activity measures ranged from 44.1% for being active on 5 or more days in the past 7 days to 57.4% for being on 1 or more sports teams. Spending 3 or more hours on a computer or digital device for something other than schoolwork on an average school day was the most commonly reported sedentary behavior (46.1%) followed by watching 3 or more hours of television on an average school day (19.7%). The prevalence of depressive symptoms in the past year was 36.7%.

Depressive symptoms in the past year increased linearly across grades (33.2% in grade 9 to 39.0% in grade 12) (Table 2). In pairwise comparisons with Bonferroni correction, the prevalence of depressive symptoms was significantly higher in grade 10, grade 11, and grade 12 than in grade 9 ($P < .05$ for all). Nearly half (46.7%) of female students reported depressive symptoms in the past year compared with 26.8% of male students. By race and ethnicity, we found the highest prevalence of depressive symptoms among American Indian/Alaska Native (45.5%) and non-Hispanic multiracial students (45.3%). In pairwise comparisons with Bonferroni correction, we found significant differences ($P < .01$ for all) in the prevalence of depressive symptoms among non-Hispanic Black versus non-Hispanic White students, Hispanic versus non-Hispanic White students, non-Hispanic multiracial versus non-Hispanic White students, Hispanic versus non-Hispanic Black students, non-Hispanic multiracial versus non-Hispanic Black students, Asian American and Other Pacific Islander versus Hispanic students, and Asian American and Native Hawaiian and Other Pacific Islander versus non-Hispanic multiracial students. Among students engaging in physical activity on 5 or more days in the past 7 days, 30.1% reported depressive symptoms in the past year, while approximately one-third of students participating in 1 or more sports teams (32.7%) and engaging in muscle strengthening on 3 or more days in the past 7 days (33.4%) reported depressive symptoms. For sedentary behaviors, the prevalence of depressive symptoms was 40.5% among students who watched 3 or more hours of television in the past 7 days and 42.7% among those who spent 3 or more hours using a computer or digital device for something besides schoolwork on an average school day.

In the first multivariable logistic regression model (Table 3, Model 1), students in grade 11 were 21% (aOR, 1.21; 95% CI, 1.03–1.43) more likely than students in grade 9 to report depressive symptoms. Male students were 56% (aOR, 0.44; 95% CI, 0.39–0.49) less likely than female students to report depressive symptoms in the past year. Compared with non-Hispanic White students, Asian American and Other Pacific Islander students were less likely to report depressive symptoms in the past year (aOR, 0.76; 95% CI, 0.58–0.99), while non-Hispanic multiracial students were more likely (aOR, 1.51; 95% CI, 1.12–2.02). For physical activity variables, students engaging in physical activity on 5 or more days in the past 7 days (aOR, 0.81; 95% CI, 0.69–0.93) or participating on 1 or more sports teams in the past year (aOR, 0.71; 95% CI, 0.62–0.82) were less likely to report depressive symptoms in the past year. Students reporting 3 or more hours using a computer or digital device for something other than schoolwork on an average school day (aOR, 1.66; 95% CI, 1.44–1.91) were more likely to report depressive symptoms in the past year.

In the model controlling for dietary behaviors, substance use, and behaviors contributing to violence in addition to demographics (Table 3, Model 2), we found inverse associations between engaging in physical activity on 5 or more days in the past 7 days (aOR, 0.81; 95% CI, 0.68–0.97) and participating on 1 or more sports teams in the past year (aOR, 0.66; 95% CI, 0.55–0.78) with depressive symptoms; these associations were similar to those found in the first model. The increased odds of depressive symptoms among students reporting 3 or more hours of computer or device use for something that was not schoolwork on an average school day (aOR, 1.61; 95% CI, 1.41–1.85) were also comparable to the increased odds found in the first model. E-cigarette use (aOR, 1.29; 95% CI, 1.07–1.57), alcohol use (aOR, 1.23; 95% CI, 1.02–1.49), prescription pain medicine use (aOR, 1.99; 95% CI, 1.65–2.41), and illicit drug use (aOR, 1.67; 95% CI, 1.36–2.04) were associated with increased odds of depressive symptoms. School safety concerns in the past 30 days (aOR, 1.93; 95% CI, 1.50–2.49), getting into a physical fight in the past 12 months (aOR, 1.40; 95% CI, 1.18–1.66), being bullied at school in the past 12 months (aOR, 2.13; 95% CI, 1.77–2.57), and being cyberbullied in the past 12 months (aOR, 2.14; 95% CI, 1.72–2.66) were associated with increased odds of depressive symptoms.

Discussion

This study provides updated data on the associations between physical activity and sedentary behavior with depressive symptoms among a nationally representative sample of high school students. In support of our main hypothesis, we found that after controlling for respondent demographic characteristics, being active on 5 or more days in the past 7 days and being a member of 1 or more sports teams were associated with reduced odds of reporting depressive symptoms in the past year. Similarly, spending 3 or more hours using a computer or digital device for something other than schoolwork was associated with increased odds of reporting depressive symptoms. The secondary hypothesis was also supported: the associations observed under the first hypothesis remained significant after controlling for demographic characteristics as well as dietary behaviors, substance use, and behaviors contributing to violence.

Consistent with current literature, our study demonstrated that physical activity may be a protective factor for depressive symptoms among adolescents. Physical activity may exert protective effects through psychosocial and behavioral mechanisms. For example, several hypothetical frameworks propose that well-being is achieved by satisfying basic psychological needs for social connectivity, autonomy, self-acceptance, environmental mastery, and purpose in life (15). Physical activity, specifically participating on sports teams, may facilitate opportunities for social connectivity among adolescents through cooperation and shared goals. Being able to follow a team schedule, listen to coaches, and participate in team workouts and competitions facilitates physical changes which can enhance self-perception (self-acceptance), discipline (purpose in life), and independence (autonomy) (15). Consistent physical activity may also reduce screen time and create more interactions with nature (12–15), which can potentially increase overall well-being.

Similar to psychosocial mechanisms, behavioral mechanisms are improved by increasing physical activity. Specifically, physical activity may improve sleep volume and quality and coping and self-regulation skills. Participation in physical activity is highly recommended for adolescents experiencing malaise and fatigue (28). It is reasonable to assume that energy expenditure during the day can improve sleep patterns, which may be associated with improved mental health. Therefore, adolescents who are physically active at least 5 days per week may have a great probability of developing better sleep habits, as their fatigued bodies will need ample rest. Participation in physical activity can lead to development of self-regulation and coping skills. For example, many physical activity programs and sports necessitate a healthy diet. Combining regular physical activity and a self-regulated healthy diet may increase self-efficacy and resiliency, which may in turn reduce the likelihood of developing depressive symptoms (29,30). Moreover, physical activity and related health behaviors may be associated with well-being through beneficial physiological mechanisms such as increases in hormones and growth factors (eg, endorphins, brain-derived neurotrophic factor), immune function, and anti-inflammatory effects (31).

Our findings on sedentary behavior and self-reported depressive symptoms among high school students are consistent with previous literature. For example, a systematic review based on a mix of cross-sectional and longitudinal studies of the association between sedentary behavior and mental health among adolescents provided strong evidence for the positive association between screen time and depressive symptoms among adolescents (32). Excessive sedentary behavior, such as watching 3 or more hours of television on an average school day, can result in bad sleep habits, lower grades, weight and mood problems, and less time with family and friends (6). Along with sedentary behavior, skipping breakfast and other unhealthy dietary behaviors, which sedentary behavior indirectly encourages, may result in malnutrition, consumption of low-nutrient foods such as processed snacks or fast foods, and obesity or weight gain (33). Regular physical activity is correlated with less sedentary behavior and healthier eating choices, potentially avoiding these problems.

Our study has several strengths. It adds to the existing body of literature by analyzing data from a nationally representative sample of high school students in the 2019 YRBS to study the associations between physical activity, sedentary behaviors, and depression among adolescents. Our results, therefore, are generalizable to all students in grades 9 through 12 enrolled in public and private schools in the 50 states and the District of Columbia (23) Our study also provides updated evidence on the associations between physical activity, sedentary behaviors, and depressive symptoms during a period characterized by dramatic increases in depressive symptoms among adolescents in high school (2–5).

Our study also has several limitations. The cross-sectional design of this study precludes causal relationships between the observed associations between physical activity, sedentary behavior, and depressive symptoms (34). Bidirectional associations are possible such that adolescents with higher levels of depressive symptoms may be less likely to engage in physical activity and more likely to engage in sedentary behaviors. Another limitation is the use of existing data from the YRBS, which depends on the use of single-item measures with limited clinical and diagnostic utility. For example, although the question on depressive symptoms has demonstrated reliability and validity, it is not clinically calibrated for detecting major depressive disorder. Other questions about physical activity and sedentary behavior may also be too general to yield more specific inferences about the types of joint activities students may commonly engage in on school days (eg, exercising while watching television or using a computer or digital device). In addition, the YRBS is administered in schools, and thus, does not include adolescents who have dropped out or been expelled, 2 populations more likely than populations that have stayed in school to engage in unhealthy health behaviors. Lastly, the YRBS is based on self-reported data and may be subject to recall and social desirability biases, especially for items asking about behaviors and cognitions beyond 1 month.

Our study provides updated evidence of the association between physical activity and sedentary behavior with depressive symptoms among high school students in the US. On the basis of these findings, we note several implications for policy and program development. Since high school is a crucial time for establishing healthy behaviors, school-based interventions should continue to focus on physical activity and sedentary behavior as ways to improve both physical and mental well-being (35,36). Universal prevention interventions could focus on educating students on the mental health benefits of physical activity and providing resources for students who are interested in team-based physical activity (37). Other interventions could be improved by integrating research about the social contexts that foster problematic sedentary and dietary behaviors, such as engagement with online content that promotes body dysmorphia and harmful gender norms (38,39). In addition, the associations between depressive symptoms and substance use and behaviors contributing to violence warrant continued attention in public health interventions for adolescents (40,41). Promoting physical activities and reducing sedentary behaviors in the daily lives of adolescents may reduce the number of adolescents who experience depressive symptoms and improve the general well-being of adolescents in high school.

Acknowledgments

The authors received no external financial support for the research, authorship, and/or publication of this article. The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. No copyrighted materials were used in this article.

Author Information

Corresponding Author: Charles H. Wang, Noble and Greenough School, 10 Campus Dr, Dedham, MA 02026. Telephone: 857-707-0378. Email: cwang2023@gmail.com.

Author Affiliations: [1]Noble and Greenough School, Dedham, Massachusetts. [2]Peiper Education and Consulting, LLC, Georgetown, Kentucky.

References

1. World Health Organization. Depression. Accessed November 30, 2021. https://www.who.int/news-room/fact-sheets/detail/depression
2. Keyes KM, Gary D, O'Malley PM, Hamilton A, Schulenberg J. Recent increases in depressive symptoms among US adolescents: trends from 1991 to 2018. Soc Psychiatry Psychiatr Epidemiol 2019;54(8):987–96.
3. Keyes KM, Hamilton A, Patrick ME, Schulenberg J. Diverging trends in the relationship between binge drinking and depressive symptoms among adolescents in the US from 1991 through 2018. J Adolesc Health 2020;66(5):529–35.
4. Lu W. Adolescent depression: national trends, risk factors, and healthcare disparities. Am J Health Behav 2019;43(1):181–94.
5. Weinberger AH, Gbedemah M, Martinez AM, Nash D, Galea S, Goodwin RD. Trends in depression prevalence in the USA from 2005 to 2015: widening disparities in vulnerable groups. Psychol Med 2018;48(8):1308–15.
6. Udwin O, Boyle S, Yule W, Bolton D, O'Ryan D. Risk factors for long-term psychological effects of a disaster experienced in adolescence: predictors of post traumatic stress disorder. J Child Psychol Psychiatry 2000;41(8):969–79.
7. Mills MS, Embury CM, Klanecky AK, Khanna MM, Calhoun VD, Stephen JM, et al. Traumatic events are associated with diverse psychological symptoms in typically-developing children. J Child Adolesc Trauma 2019;13(4):381–8.
8. Merlo CL, Jones SE, Michael SL, Chen TJ, Sliwa SA, Lee SH, et al. Dietary and physical activity behaviors among high school students — Youth Risk Behavior Survey, United States, 2019. MMWR Suppl 2020;69(1):64–76.
9. Burns RD. Energy balance-related factors associating with adolescent weight loss intent: evidence from the 2017 National Youth Risk Behavior Survey. BMC Public Health 2019;19(1):1206.
10. Beach LB, Turner B, Felt D, Marro R, Phillips GL 2d. Risk factors for diabetes are higher among non-heterosexual US high-school students. Pediatr Diabetes 2018;19(7):1137–46.
11. American Academy of Child and Adolescent Psychiatry. Facts for families: screen time and children. Accessed October 26, 2021. https://www.aacap.org/AACAP/Families_and_Youth/Facts_for_Families/FFF-Guide/Children-And-Watching-TV-054.aspx
12. Wu XY, Han LH, Zhang JH, Luo S, Hu JW, Sun K. The influence of physical activity, sedentary behavior on health-related quality of life among the general population of children and adolescents: a systematic review. PLoS One 2017;12(11):e0187668.
13. Rodriguez-Ayllon M, Cadenas-Sánchez C, Estévez-López F, Muñoz NE, Mora-Gonzalez J, Migueles JH, et al. Role of physical activity and sedentary behavior in the mental health of preschoolers, children and adolescents: a systematic review and meta-analysis. Sports Med 2019;49(9):1383–410.
14. Lu C, Chi X, Liang K, Chen ST, Huang L, Guo T, et al. Moving more and sitting less as healthy lifestyle behaviors are protective factors for insomnia, depression, and anxiety among adolescents during the COVID-19 pandemic. Psychol Res Behav Manag 2020;13:1223–33.
15. Chopra C, Mandalika S, Kinger N. Does diet play a role in the prevention and management of depression among adolescents? A narrative review. Nutr Health 2021;27(2):243–63.
16. Chaiton MO, Cohen JE, O'Loughlin J, Rehm J. A systematic review of longitudinal studies on the association between depression and smoking in adolescents. BMC Public Health 2009;9(1):356.
17. Christian C, Keshishian AC, Levinson CA, Peiper NC. A network examination of risky behaviours in a state-level and national epidemiological sample of high school students. Early Interv Psychiatry 2021;15(6):1650–8.
18. David-Ferdon C, Clayton HB, Dahlberg LL, Simon TR, Holland KM, Brener N, et al. Vital signs: prevalence of multiple forms of violence and increased health risk behaviors and conditions among youths — United States, 2019. MMWR Morb Mortal Wkly Rep 2021;70(5):167–73.
19. Mitchell KJ, Ybarra M, Finkelhor D. The relative importance of online victimization in understanding depression, delinquency, and substance use. Child Maltreat 2007;12(4):314–24.
20. Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. Lancet 2012;379(9820):1056–67.
21. Crane NA, Langenecker SA, Mermelstein RJ. Risk factors for alcohol, marijuana, and cigarette polysubstance use during adolescence and young adulthood: a 7-year longitudinal study of youth at high risk for smoking escalation. Addict Behav 2021;119:106944.
22. Korczak DJ, Madigan S, Colasanto M. Children's physical activity and depression: a meta-analysis. Pediatrics 2017;139(4):e20162266.
23. Underwood JM, Brener N, Thornton J, Harris WA, Bryan LN, Shanklin SL, et al. Overview and Methods for the Youth Risk Behavior Surveillance System — United States, 2019. MMWR Suppl 2020;69(1):1–10.

24. May A, Klonsky ED. Validity of suicidality items from the Youth Risk Behavior Survey in a high school sample. Assessment 2011;18(3):379–81.
25. Brener ND, Collins JL, Kann L, Warren CW, Williams BI. Reliability of the Youth Risk Behavior Survey Questionnaire. Am J Epidemiol 1995;141(6):575–80.
26. Jones CM, Clayton HB, Deputy NP, Roehler DR, Ko JY, Esser MB, et al. Prescription opioid misuse and use of alcohol and other substances among high school students — Youth Risk Behavior Survey, United States, 2019. MMWR Suppl 2020; 69(1):38–46.
27. Brener ND, Kann L, McManus T, Kinchen SA, Sundberg EC, Ross JG. Reliability of the 1999 youth risk behavior survey questionnaire. J Adolesc Health 2002;31(4):336–42.
28. Lubans D, Richards J, Hillman C, Faulkner G, Beauchamp M, Nilsson M, et al. Physical activity for cognitive and mental health in youth: a systematic review of mechanisms. Pediatrics 2016;138(3):e20161642.
29. Belcher BR, Zink J, Azad A, Campbell CE, Chakravartti SP, Herting MM. The roles of physical activity, exercise, and fitness in promoting resilience during adolescence: effects on mental well-being and brain development. Biol Psychiatry Cogn Neurosci Neuroimaging 2021;6(2):225–37.
30. Efthymiou V, Charmandari E, Vlachakis D, Tsitsika A, Pałasz A, Chrousos G, et al. Adolescent self-efficacy for diet and exercise following a school-based multicomponent lifestyle intervention. Nutrients 2021;14(1):97.
31. Mandolesi L, Polverino A, Montuori S, Foti F, Ferraioli G, Sorrentino P, et al. Effects of physical exercise on cognitive functioning and wellbeing: biological and psychological benefits. Front Psychol 2018;9:509.
32. Hoare E, Milton K, Foster C, Allender S. The associations between sedentary behaviour and mental health among adolescents: a systematic review. Int J Behav Nutr Phys Act 2016;13(1):108.
33. Jacka FN, Kremer PJ, Berk M, de Silva-Sanigorski AM, Moodie M, Leslie ER, et al. A prospective study of diet quality and mental health in adolescents. PLoS One 2011;6(9):e24805.
34. Rodriguez-Ayllon M, Cadenas-Sánchez C, Estévez-López F, Muñoz NE, Mora-Gonzalez J, Migueles JH, et al. Role of physical activity and sedentary behavior in the mental health of preschoolers, children and adolescents: a systematic review and meta-analysis. Sports Med 2019;49(9):1383–410.
35. Adrian M, Charlesworth-Attie S, Vander Stoep A, McCauley E, Becker L. Health promotion behaviors in adolescents: prevalence and association with mental health status in a statewide sample. J Behav Health Serv Res 2014;41(2): 140–52.
36. Suryavanshi MS, Yang Y. Clinical and economic burden of mental disorders among children with chronic physical conditions, United States, 2008–2013. Prev Chronic Dis 2016; 13:E71.
37. Martin A, Booth JN, Laird Y, Sproule J, Reilly JJ, Saunders DH. Physical activity, diet and other behavioural interventions for improving cognition and school achievement in children and adolescents with obesity or overweight. Cochrane Database Syst Rev 2018;1:CD009728.
38. Kandola A, Del Pozo Cruz B, Hayes JF, Owen N, Dunstan DW, Hallgren M. Impact on adolescent mental health of replacing screen-use with exercise: a prospective cohort study. J Affect Disord 2022;301:240–7.
39. Vidal C, Lhaksampa T, Miller L, Platt R. Social media use and depression in adolescents: a scoping review. Int Rev Psychiatry 2020;32(3):235–53.
40. Morgan-Lopez AA, Elek E, Graham PW, Saavedra LM, Bradshaw M, Clarke T. A quasi-experimental evaluation of partnerships for success's impact on community-level ethanol and prescription drug poisoning rates. Addict Behav 2019;95: 220–5.
41. Espelage DL, Low SK, Jimerson SR. Understanding school climate, aggression, peer victimization, and bully perpetration: contemporary science, practice, and policy. Sch Psychol Q 2014;29(3):233–7.

Tables

Table 1. Characteristics of High School Students, 2019 Youth Risk Behavior Survey

Characteristic	Overall, % (SE) (N = 13,526)
Grade	
9	26.7 (0.6)
10	25.5 (0.4)
11	24.3 (0.6)
12	23.6 (0.6)
Sex	
Female	49.4 (0.7)
Male	50.6 (0.7)
Race and ethnicity	
American Indian or Alaska Native	0.7 (0.08)
Asian American and Native Hawaiian/Other Pacific Islander	5.4 (1.4)
Hispanic	26.0 (2.3)
Non-Hispanic Black	12.2 (1.1)
Non-Hispanic White	51.2 (2.4)
Non-Hispanic multiracial	4.5 (0.4)
Physical activity	
Active ≥5 days in past 7 days	44.1 (1.1)
≥1 Sports team during past 12 months	57.4 (1.5)
Muscle strengthening ≥3 days in past 7 days	49.5 (0.9)
Sedentary behavior on average school day	
≥3 Hours of television	19.7 (0.7)
≥3 Hours on a computer or digital device	46.1 (0.8)
Dietary behaviors in past 7 days	
Skipped breakfast all days	16.6 (0.7)
No fruits	11.8 (0.7)
No vegetables[a]	7.7 (0.4)
Substance use	
Cigarettes in past 30 days	5.9 (0.5)
E-cigarettes in past 30 days	32.7 (1.0)
Cannabis in past 30 days	21.7 (0.1)
Alcohol in past 30 days	29.2 (0.1)
Prescription pain medicine in lifetime	14.2 (0.8)
Illicit drugs in lifetime[b]	14.6 (0.9)
Behaviors contributing to violence	

[a] Did not eat green salad, potatoes, carrots, and other vegetables 1 or more times during the past 7 days.
[b] Lifetime use of cocaine, inhalants, heroin, methamphetamines, ecstasy, or hallucinogens.

(continued on next page)

(continued)

Table 1. Characteristics of High School Students, 2019 Youth Risk Behavior Survey

Characteristic	Overall, % (SE) (N = 13,526)
School safety concerns in past 30 days	8.7 (0.6)
Physical fight in past 12 months	21.8 (0.8)
Bullied at school in past 12 months	19.5 (0.7)
Cyberbullied in past 12 months	15.7 (0.5)

[a] Did not eat green salad, potatoes, carrots, and other vegetables 1 or more times during the past 7 days.
[b] Lifetime use of cocaine, inhalants, heroin, methamphetamines, ecstasy, or hallucinogens.

Table 2. Weighted Prevalence of Depressive Symptoms Among High School Students, 2019 Youth Risk Behavior Survey

Characteristic	% (SE)
Overall	36.7 (0.8)
Grade	
9	33.2 (1.2)
10	37.0 (1.5)
11	37.9 (1.2)
12	39.0 (1.4)
Sex	
Female	46.7 (1.1)
Male	26.8 (0.8)
Race and ethnicity	
American Indian or Alaska Native	45.5 (6.6)
Asian American and Native Hawaiian/Other Pacific Islander	31.9 (2.1)
Hispanic	40.0 (1.0)
Non-Hispanic Black	31.5 (1.4)
Non-Hispanic White	36.0 (1.0)
Non-Hispanic multiracial	45.3 (3.3)
Physical activity	
Active ≥5 days in past 7 days	30.1 (1.0)
≥1 Sports team during past 12 months	32.7 (1.0)
Muscle strengthening ≥3 days in past 7 days	33.4 (1.1)
Sedentary behaviors on average school day	
≥3 Hours of television	40.5 (1.3)
≥3 Hours on a computer or digital device	42.7 (1.0)
Dietary behaviors in past 7 days	
Skipped breakfast all days	48.8 (1.6)
No fruits	37.5 (1.8)
No vegetables[a]	32.6 (2.0)
Substance use	
Cigarettes in past 30 days	58.0 (2.7)
E-cigarettes in past 30 days	48.7 (1.0)
Cannabis in past 30 days	50.4 (1.4)
Alcohol in past 30 days	48.7 (1.2)
Prescription pain medicine in lifetime	59.8 (1.4)
Illicit drugs in lifetime[b]	59.8 (1.9)
Behaviors contributing to violence	
School safety concerns in past 30 days	60.0 (2.4)

[a] Did not eat green salad, potatoes, carrots, and other vegetables one or more times during the past 7 days.
[b] Lifetime use of cocaine, inhalants, heroin, methamphetamines, ecstasy, or hallucinogens.

(continued on next page)

(continued)

Table 2. Weighted Prevalence of Depressive Symptoms Among High School Students, 2019 Youth Risk Behavior Survey

Characteristic	% (SE)
Physical fight in past 12 months	47.2 (1.3)
Bullied at school in past 12 months	61.9 (1.5)
Cyberbullied in past 12 months	65.1 (1.5)

[a] Did not eat green salad, potatoes, carrots, and other vegetables one or more times during the past 7 days.
[b] Lifetime use of cocaine, inhalants, heroin, methamphetamines, ecstasy, or hallucinogens.

Table 3. Multivariate Associations Between Physical Activity and Sedentary Behaviors With Depressive Symptoms Among High School Students, 2019 Youth Risk Behavior Survey

Characteristic	Adjusted odds ratio (95% CI)	
	Model 1[a]	Model 2[a]
Grade		
9	1 [Reference]	1 [Reference]
10	1.18 (0.97–1.43)	1.20 (0.99–1.45)
11	1.21 (1.03–1.43)	1.25 (1.08–1.44)
12	1.16 (0.94–1.44)	1.11 (0.90–1.37)
Sex		
Female	1 [Reference]	1 [Reference]
Male	0.44 (0.39–0.49)	0.48 (0.41–0.56)
Race and ethnicity		
American Indian or Alaska Native	1.82 (0.88–3.78)	1.53 (0.68–3.45)
Asian American and Native Hawaiian/Other Pacific Islander	0.76 (0.58–0.99)	1.15 (0.86–1.56)
Hispanic	1.12 (0.97–1.30)	1.33 (1.14–1.54)
Non-Hispanic Black	0.79 (0.61–1.01)	0.90 (0.66–1.21)
Non-Hispanic White	1 [Reference]	1 [Reference]
Non-Hispanic multiracial	1.51 (1.12–2.02)	1.49 (1.08–2.04)
Physical activity		
Active ≥5 days in past 7 days	0.81 (0.69–0.93)	0.81 (0.68–0.97)
≥1 Sports team during past 12 months	0.71 (0.62–0.82)	0.66 (0.55–0.78)
Muscle strengthening ≥3 days in past 7 days	1.04 (0.92–1.19)	0.99 (0.86–1.13)
Sedentary behaviors on average school day		
≥3 Hours of television	1.08 (0.92–1.26)	1.06 (0.87–1.30)
≥3 Hours on a computer or digital device	1.66 (1.44–1.91)	1.61 (1.41–1.85)
Dietary behaviors in past 7 days		
Skipped breakfast all days	—	1.45 (1.23–1.71)
No fruits	—	1.00 (0.77–1.31)
No vegetables[b]	—	0.80 (0.65–1.00)
Substance use		
Cigarettes in past 30 days	—	1.25 (0.90–1.73)
E-cigarettes in past 30 days	—	1.29 (1.07–1.57)
Cannabis in past 30 days	—	1.15 (0.87–1.51)
Alcohol in past 30 days	—	1.23 (1.02–1.49)
Prescription pain medicine in lifetime	—	1.99 (1.65–2.41)

Abbreviation: — , does not apply.

[a] Respondent's demographics were controlled for in Model 1; Model 2 included demographics, dietary behaviors, substance use, and behaviors contributing to violence.

[b] Did not eat green salad, potatoes, carrots, and other vegetables one or more times during the past 7 days.

[c] Lifetime use of cocaine, inhalants, heroin, methamphetamines, ecstasy, or hallucinogens.

(continued on next page)

(continued)

Table 3. Multivariate Associations Between Physical Activity and Sedentary Behaviors With Depressive Symptoms Among High School Students, 2019 Youth Risk Behavior Survey

Characteristic	Adjusted odds ratio (95% CI)	
	Model 1[a]	Model 2[a]
Illicit drugs in lifetime[c]	—	1.67 (1.36–2.04)
Behaviors contributing to violence		
School safety concerns in past 30 days	—	1.93 (1.50–2.49)
Physical fight in past 12 months	—	1.40 (1.18–1.66)
Bullied at school in past 12 months	—	2.13 (1.77–2.57)
Cyberbullied in past 12 months	—	2.14 (1.72–2.66)

Abbreviation: — , does not apply.

[a] Respondent's demographics were controlled for in Model 1; Model 2 included demographics, dietary behaviors, substance use, and behaviors contributing to violence.

[b] Did not eat green salad, potatoes, carrots, and other vegetables one or more times during the past 7 days.

[c] Lifetime use of cocaine, inhalants, heroin, methamphetamines, ecstasy, or hallucinogens.

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
Volume 19, E74 NOVEMBER 2022

ORIGINAL RESEARCH

Physical Activity Among Adults in Rural Western North Carolina During the COVID-19 Pandemic

Wade L. Creech[1]; Brooke C. Towner, PhD[1]; Rebecca A. Battista, PhD[1]

Accessible Version: www.cdc.gov/pcd/issues/2022/22_0112.htm

Suggested citation for this article: Creech WL, Towner BC, Battista RA. Physical Activity Among Adults in Rural Western North Carolina During the COVID-19 Pandemic. Prev Chronic Dis 2022;19:220112. DOI: https://doi.org/10.5888/pcd19.220112.

PEER REVIEWED

Summary

What is already known on this topic?

Physical activity could be important to physical and mental health during the COVID-19 pandemic. Rural communities often lack access to locations and facilities for physical activity.

What is added by this report?

We sought to learn about the locations, reasons, barriers, and benefits related to physical activity in rural areas to better understand the distinct needs of rural communities compared to nonrural communities.

What are the implications for public health practice?

Our research may inform physical activity education and the development of physical activity infrastructure in rural communities, thereby improving residents' levels of physical activity and overall health and wellness.

Abstract

Introduction

During the COVID-19 pandemic, measures implemented to protect community health may have influenced how and where people engaged in physical activity. In rural communities, access to resources, the environment, and socioeconomic status could play a role in how adults are physically active. Our study examined locations where rural residents of a county in western North Carolina engaged in physical activity early in the COVID-19 pandemic, their reasons for being physically active, and their perceptions of benefits and barriers related to engaging in physical activity.

Methods

Rural adults (N =297) completed an online survey from August 3 through September 15, 2020, describing their physical activity during the summer of 2020. Data were analyzed using nonparametric measures.

Results

Survey respondents frequently engaged in physical activity in the home (57.8%), at parks or on trails (45.3%), and around their neighborhood (39.4%). The most common types of physical activities at parks or on trails were walking and hiking (99.5%). Across all locations, the most frequently reported reasons for engaging in physical activity were getting out of the house, maintaining fitness and mental health, and exercising.

Conclusion

Our study showed many locations where rural residents were physically active and their reasons for participating in physical activity during the pandemic. Data about perceived benefits of and barriers to physical activity during the pandemic can assist in meeting the current need to increase physical activity levels in rural communities.

Introduction

Physical activity yields both physical and mental health benefits (1). These benefits became even more critical during the COVID-19 pandemic (2,3). However, in the early part of the pandemic, government stay-at-home orders (4) resulted in closure of many common locations for physical activity (eg, fitness centers, parks, trails), which led to a decline in physical activity among many adults (5–8). As a result, people had to adapt their physical activity behaviors.

Walking paths, trails, sidewalks, parks, and recreation facilities are common locations for physical activity (9). Access to such places is an important factor (10) and influences how and whether residents are physically active (7,11). For example, rural residents often lack developed neighborhoods (eg, sidewalks, safe streets) (7), and the pandemic may have restricted access to or closures of trails and parks.

Respondents to recent surveys during the COVID-19 pandemic reported various reasons for engaging in physical activity and various perceptions of benefits. These included having more time available, schedule flexibility, an increased desire to be outside, wanting to decrease stress or anxiety, and other health-related concerns (6,12). Additionally, outdoor activity became especially desirable (5,13) because it facilitated social distancing (12). At the same time, concerns about exposure to COVID-19, lack of available exercise equipment, state or local restrictions, and living conditions were barriers to being physically active (7,8), as were reported “worry/stress” and “lack of motivation” (6).

In 2017, an estimated 1 in 5 people in the US lived in a rural area (14). Although the sparse population of these areas facilitated social distancing, lack of resources such as sidewalks, parks, and fitness clubs may have served as barriers to regular physical activity (10). Although urban and suburban neighborhoods may have such resources (5), rural areas may not (10). To fully interpret the needs of subgroups of rural respondents to national surveys, a full understanding of their available physical activity resources is needed.

The primary objective of our study was to examine the types of available physical activity resources in rural areas and the reasons adults in these areas engaged in physical activity early in the COVID-19 pandemic. A secondary objective was to describe the benefits and barriers to physical activity residents encountered. This information could be useful to public health professionals and clinicians and to land use and public recreation planners in developing an infrastructure to facilitate outdoor physical activity in rural areas (15).

Methods

An online survey completed from August 3 through September 15, 2020, appraised locations, reasons for physical activity during the COVID-19 pandemic, and related benefits and barriers.

Participants and recruitment

By using convenience sampling, we recruited from the general adult population of a western North Carolina county with a rural–urban continuum code (RUCC) of 5 and a nonmetropolitan–nonurban population of 20,000 or more that was not adjacent to a metropolitan area (16). Our inclusion criteria were adults aged 18 years or older, residence in rural western North Carolina, the ability to read English, and access to an electronic device to complete the survey. We contacted prospective participants electronically by using listservs from the county health department, 7 public kindergarten through 8th grade schools, 1 public high school, and 3 family-based homeschool organizations. To obtain the lists, we contacted public school administrators, the health department, and local homeschool organizations via email to explain the purpose of our study and to ask them to participate by distributing our survey to their constituents. Using these listservs allowed us to select names from our county of interest. The county health department list included names from programs for people of low socioeconomic status, which allowed us to reach that audience, and homeschool programs let us reach people with children who did not attend public schools. Those who agreed to distribute the survey received an email describing the study and an email message for them to distribute to their listserv along with a link to the online survey (Qualtrics© XM survey software system). Of the approximately 4,500 people we reached through these listservs, 337 met our criteria and agreed to participate after reviewing the survey’s online consent page describing the study's procedures, risks, and benefits. The institutional review board of Appalachian State University determined this study was exempt.

Participants completed the survey from August 3 through September 15, 2020, during phase 2 of pandemic restrictions, which in North Carolina consisted of face mask requirements and limited gatherings (maximum 10 people indoors and 25 people outdoors). Stay-at-home orders were lifted during this period. The survey consisted of 50 questions and took approximately 30 minutes to complete. After completing the survey, participants had the option of entering a drawing for one of 150 $25 gift cards.

Measures

Locations for physical activity. To determine where participants engaged in physical activity, we asked “Please tell us how often you used the location for physical activity during the COVID-19 pandemic,” defined as the current COVID-19 phase 2 in North Carolina. Options were parks and trails, neighborhoods, and in or around the home. Answers were given on a 3–point Likert scale of 1, frequently; 2, sometimes; and 3, never used for physical activity.

Types of physical activity. To assess types of activities for these locations, participants were asked, “During the COVID-19 pandemic, what physical activities did you do?” Participants selected all types of physical activity they participated in at each location. Answers for parks and trails and for neighborhood were walking or hiking, jogging or running, biking, playing with kids, sports activities, kayaking or canoeing, and swimming; for home, answers were indoor housework, workouts or exercise, playing in the yard, outdoor sports or games, yard work, and gardening. The 2018 Physical Activity Guidelines for Americans included many of these activities as examples of moderate to vigorous aerobic activities (1).

Reasons for physical activity. Participants were asked "Please choose your reason(s) for participating in physical activity" at each location. Response options were as follows: for exercise, to experience nature, to get out of the house, to do something normal, for mental health, to take kids outside, to walk the dog, to connect with others, to maintain health and fitness, to prepare for future events, and to reduce sitting time.

Benefits and barriers. Participants were asked to rate their beliefs about the benefits of physical activity and barriers to engaging in it on a Likert scale of 1 to 5, with 1 indicating strongly disagree and 5 indicating strongly agree. The following were benefit options: something I can do without being exposed to COVID-19, offers a routine or something normal, helps manage stress, helps manage mental health, helps manage chronic disease, supports my immune system, reduces the risk of COVID-19, and will reduce the severity of COVID-19. Barrier options were the following: appropriate social distancing, closures of park and recreation resources, cancellation of sport/intramurals, neighborhood environment, closure of fitness facility, lack of fitness equipment, and stay-at-home orders.

Demographics. Participants reported their gender (male, female, other), age (categorized as 18–25 y, 26–35 y, 36–45 y, 46–55 y, 56–65 y, ≥66 y), race, education level, annual household income (≤$20,000, $20,000– $59,999, $60,000– $99,999, $100,000–$149,999, ≥$150,000), and number of children aged 18 years or younger in the home (0, 1, 2, 3, 4, ≥5).

Analyses were conducted in 2021 by using SPSS Statistics 27 (IBM Corp) with nonparametric measures. We ran frequencies for demographic information and subjective variables with single-response items. Responses were grouped by the location of physical activity (parks or trails, neighborhood, home) during the pandemic. The percentage of respondents who participated in various activities and the reason for participation were determined by the location of the activity.

Results

Of the 337 people who agreed to take the survey, 297 completed at least 50% and were included in our sample. Most participants were middle-aged (36–55 y), White, and female (Table 1). Approximately 74.4 % of participants held a post–secondary-education degree (eg, Bachelor, Master, Doctorate, or other professional degree). Participants' household income varied; 53.9% of earned from $60,000 to $150,000 annually.

More than half (57.8%) of participants frequently used their homes for physical activity, 39.7% used their homes sometimes, and 2.5% never used their homes. We saw similar numbers in the use of parks and neighborhoods: 45.3% frequently used parks, and 39.4% frequently used their neighborhood; 36% sometimes used parks and 43.2% sometimes used their neighborhood; 18.6% never used parks and 17.4% never used their neighborhood.

The most prevalent types of physical activity in the home were indoor housework (67.8%), yard work/landscaping (60.6%), and workouts/exercise (57.2%) (Table 2). The most prevalent reason for physical activity at home was to maintain health and fitness (76.0%). Other common reasons were for mental health (74.5%) and exercise (66.3%) (Table 3).

Among the 183 respondents who said they used parks and trails, 99.5% used them for walking and hiking. Other popular types of physical activity were playing with kids (44.8%), kayaking or canoeing (43.2%), and jogging or running (40.4%) (Table 2). The most frequent reasons given for participating in physical activity at parks or trails were to get out of the house (95.1%) and for exercise (88.5%). Other reasons were for mental health (79.8%), to experience nature (77.0%), and to maintain health and fitness (76.0%) (Table 3).

For physical activity in neighborhoods, 176 participants reported engaging in some form. Walking and hiking were the most prevalent types (98.3%) followed by bicycling (36.4%) and jogging or running (28.4%) (Table 2). The most prevalent reasons given for physical activity in neighborhoods were to get out of the house (89.1%), for exercise (83.6%), and for mental health (70.5%) (Table 3).

Perceived barriers to and benefits of physical activity varied (Table 4). Respondents either agreed (38.9%) or strongly agreed (33.0%) that closure of parks and recreation resources was a barrier to physical activity. The second leading barrier to participation was stay-at-home orders (26.6% agreed and 29.1% strongly agreed). Respondents strongly agreed about other barriers to physical activity, including the cancellation of recreational sports or intramural games (34.0%) and the closure of fitness facilities (28.6%). More than half of respondents strongly agreed that physical activity helped to manage mental health (56.2%) and stress (53.7%); almost half (46.8%) strongly agreed that physical activity was beneficial to the immune system during the COVID-19 pandemic.

Discussion

Our evaluation showed that adults living in rural western North Carolina primarily used the home for physical activity in the early COVID-19 pandemic to maintain fitness and mental health. Activities in and around the home were indoor housework, yardwork, and workout or exercise. This finding was similar to other surveys administered during the pandemic in which the home or garage were the most commonly reported locations (5,7). However, our results differed from others in that the second most common location was parks and trails rather than neighborhoods (5,7). Participants in our study gave reasons for using these locations as "to get out of the house," "for exercise," and "for mental health." Our participants also reported going outside to be active and to "do something normal." Because rural communities often do not have the infrastructure of sidewalks, safe streets, greenways, and playing fields, which are more common in nonrural communities, parks and trails may have been more accessible to them during the pandemic, especially during the stay-at-home phase. Although some of these locations experienced closures during the initial pandemic wave, our data suggest that these spaces were some of the few locations available for physical activity in rural communities.

Our results also showed that various perceived barriers and benefits to physical activity existed and contributed to use of the locations chosen and why. Because of enforced public health regulations, many people could not engage in their preferred recreational activities (eg, travel, eating out at restaurants, going to movies, using fitness centers) (15). Our collected data support this conclusion, because prominent barriers to physical activity were the closure of fitness facilities and because stay-at-home orders made outdoor locations such as parks and trails and neighborhoods increasingly popular. The increase in outdoor recreation and park use was potentially the result of a recreation substitution, as new outdoor recreationists either tried outdoor recreation for the first time or returned to outdoor recreation after a prolonged hiatus (15).

The benefits of engaging in physical activity in our sample included managing overall health. Our participants reported "helps manage my mental health," "helps manage stress," and "is beneficial to my immune system" as reasons. Similarly, others reported stress as a reason for changing health behaviors such as becoming physically active. One study found adults who positively changed health behavior by increasing physical activity during the COVID-19 pandemic did so because they had more time available and used activity to relieve stress and for other health-related concerns (7).

The disparate effects of the COVID-19 pandemic on physical activity levels of potentially vulnerable subgroups, including people with low incomes or who live in rural communities, underscore the need for population-specific physical activity programs and policies over the next several months to years as the pandemic continues (5). Recent evidence from similar studies supports the need for further understanding of subgroups within a population, assessments of their resource concerns, and more information about their physical activity choices to better determine ways to increase physical activity (5,7,8). Although rural communities may lack infrastructure, our results point to the potential use of parks and trails as immediate resources for physical activity. Young adults (18–29 y) engaged in physical activity at parks and open spaces during the early part of the pandemic and continued to be active at these locations 13 months later (13). Similar results suggest the outdoors as an option for future behavioral interventions (17).

Our study had limitations. First, study participants were mainly female, middle-aged, White, with an annual income above $60,000, and a post–secondary-education degree. The lack of diversity reflects the demographics of rural western North Carolina. Other participants in other studies had similar demographic characteristics: a high percentage of females (5,6,8), non-Hispanic people, and residents with post-secondary education (5). According to the Western North Carolina Health Network, the racial and ethnic population of western North Carolina is 89.9% White, 4.3% Black or African American, 1.5% American Indian/Alaska Native, and 6.0% Hispanic or Latino (of any race) (18); 50.2% of the population in the county we studied were female (19). This lack of diversity in our sample is a possible limitation in generalizing results on a broader scale. In addition, our sample included a disproportionately large number of participants with a post–secondary-education degree. In 2019, approximately 35% of Americans held some type of post-secondary academic degree (20), whereas in our sample 74.4% of participants (n = 221) held a post-secondary degree. Thus, education level may influence regular exercise, with the findings showing that people with advanced degrees were more likely to exercise regularly (21). Another limitation was that we developed this survey urgently to address the immediate need for information related to the COVID-19 pandemic. We did not test the survey's validity or reliability before administering it.

Our study results suggest many locations and reasons for participating in physical activity in rural western North Carolina that could be applied to other rural areas of the US. Perceived barriers and benefits related to physical activity can aid further understanding of why participants engaged in physical activity during the COVID-19 pandemic. Because use of parks and trails contributed to outdoor physical activity among rural residents, access to such spaces is key to increasing and maintaining levels of physical activity after the threat of the COVID-19 pandemic subsides. Understanding where adults want to be physically active may also be helpful when developing interventions and resources (22). The early phases of the COVID-19 pandemic (eg, shelter-in-place, stay-at-home) affected many health-related behaviors and forced adults to change where and why they engaged in physical activity. However, much can be learned from behaviors during this time frame. In the future, public health agencies could assist in creating awareness of available physical activity locations.

Acknowledgments

This research was funded by an internal COVID-19 Research Clusters Grant from Appalachian State University. The funding office was not involved in the study design, data analysis, interpretation of results, or decision to submit this article for publication. The findings of this article are the authors' and do not necessarily represent the views of the funding office. No copyrighted materials were used in this article.

Author Information

Corresponding Author: Wade L. Creech, 5464 NC Highway 231, Zebulon, NC 27597-7661. Telephone: (919) 426-7130. Email: Wadecreech1@gmail.com.

Author Affiliations: [1]Appalachian State University, Boone, North Carolina.

References

1. Centers for Disease Control and Prevention. 2018 Physical activity guidelines for Americans. Hyattsville (MD): US Department of Health and Human Services; 2018. Accessed June 20, 2022. http://www.health.gov/paguidelines/pdf/paguide.pdf
2. Hacker KA, Briss PA. An ounce of prevention is still worth a pound of cure, especially in the time of COVID-19. Prev Chronic Dis 2021;18:E03.
3. Nieman DC. Coronavirus disease–2019: a tocsin to our aging, unfit, corpulent, and immunodeficient society. J Sport Health Sci 2020;9(4):293–301.
4. nc.gov. COVID-19 orders & directives; 2022. Accessed March 21, 2022. https://www.nc.gov/covid-19/covid-19-orders-directives
5. Dunton GF, Wang SD, Do B, Courtney J. Early effects of the COVID-19 pandemic on physical activity locations and behaviors in adults living in the United States. Prev Med Rep 2020;20:101241.
6. Knell G, Robertson MC, Dooley EE, Burford K, Mendez KS. Health behavior changes during COVID-19 pandemic and subsequent "stay-at-home" orders. Int J Environ Res Public Health 2020;17(17):6268.
7. Watson KB, Whitfield GP, Huntzicker G, Omura JD, Ussery E, Chen TJ, et al. Cross-sectional study of changes in physical activity behavior during the COVID-19 pandemic among US adults. Int J Behav Nutr Phys Act 2021;18(1):91.
8. Wijngaards I, Del Pozo Cruz B, Gebel K, Ding D. Exercise frequency during the COVID-19 pandemic: A longitudinal probability survey of the US population. Prev Med Rep 2022;25:101680.
9. Abildso CG, Perry CK, Jacobs L, Umstattd Meyer MR, McClendon M, Edwards MB, et al. What sets physically active rural communities apart from less active ones? A comparative case study of three US counties. Int J Environ Res Public Health 2021;18(20):10574.
10. Christiana RW, Bouldin ED, Battista RA. Active living environments mediate rural and non-rural differences in physical activity, active transportation, and screen time among adolescents. Prev Med Rep 2021;23:101422.
11. Sallis JF, Cervero RB, Ascher W, Henderson KA, Kraft MK, Kerr J. An ecological approach to creating active living communities. Annu Rev Public Health 2006;27(1):297–322.
12. Mel AE, Stenson MC. Physical activity changes during the Spring 2020 COVID-19 shutdown in the United States. Transl J Am Coll Sports Med 2021;6(4):1–8.
13. Yi L, Wang SD, Chu D, Ponnada A, Intille SS, Dunton GF. Examining whether physical activity location choices were associated with weekly physical activity maintenance across 13 months of the COVID-19 pandemic in emerging adults. J Phys Act Health 2022;19(6):446–55.
14. Census.gov. One in five Americans live in rural areas. 2017. Accessed March 21, 2022. https://www.census.gov/library/stories/2017/08/rural-america.html
15. Rice WL, Mateer TJ, Reigner N, Newman P, Lawhon B, Taff BD. Changes in recreational behaviors of outdoor enthusiasts during the COVID-19 pandemic: analysis across urban and rural communities. J Urban Econ 2020;6(1):1–7.
16. Economic Research Service. Rural-urban continuum codes. Washington (DC): U.S. Department of Agriculture; 2013. Accessed June 20, 2022. https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx
17. Folk AL, Wagner BE, Hahn SL, Larson N, Barr-Anderson DJ, Neumark-Sztainer D. Changes to physical activity during a global pandemic: a mixed methods analysis among a diverse population-based sample of emerging adults in the U.S. Int J Environ Res Public Health 2021;18(7):3674.
18. WNC Health Network. WNC data/regional key health issues. 2020Accessed March 21, 2022. https://www.wnchn.org/wnc-data/about-wnc-2/#:~:text=Western%20North%20Carolina%20is%2089.9,Latino%20(of%20any%20rac)
19. US Census Bureau. Quick Facts: Watagua County, North Carolina, United States. 2021. Accessed July 18, 2022. https://www.census.gov/quickfacts/fact/table/wataugacountynorthcarolina,US/EDU685220#EDU685220

20. Jiménez-Pavón D, Carbonell-Baeza A, Lavie CJ. Physical exercise as therapy to fight against the mental and physical consequences of COVID-19 quarantine: Special focus in older people. Prog Cardiovasc Dis 2020;63(3):386–8.
21. Zajacova A, Lawrence E. Postsecondary educational attainment and health among younger U.S. adults in the "college-for-all" era. Socius 2021;7.
22. Whitfield GP, Carlson SA, Ussery EN, Watson KB, Adams MA, James P, et al. Environmental supports for physical activity, National Health Interview Survey-2015. Am J Prev Med 2018;54(2):294–8.

Tables

Table 1. Demographic Characteristics, Adults (N = 297) in Rural North Carolina During the COVID-19 Pandemic, August 3–September 15, 2020[a]

Characteristic	n (%)
Gender	
Male	43 (14.5)
Female	254 (85.5)
Other	1 (0.3)
Age, y	
18–25	6 (2.0)
26–35	33 (11.1)
36–45	110 (37.0)
46–55	104 (35.0)
56–65	34 (11.4)
66–75	10 (3.4)
Race	
American Indian or Alaska Native	3 (1.0)
Asian	8 (2.7)
Black or African American	2 (0.7)
White	285 (96.0)
Annual household income, $	
<20,000	8 (2.7)
20,000–59,999	67 (22.6)
60,000–99,999	92 (31.0)
100,000–149,999	68 (22.9)
≥150,000 or more	35 (11.8)
Prefer not to respond	28 (9.4)
Education	
High school graduate	12 (4.0)
Some college	62 (20.9)
Bachelor’s degree	103 (34.7)
Master’s degree	93 (31.3)
Doctorate/professional degree	25 (8.4)
Prefer not to respond	2 (0.7)

[a] Values are number (percentage). Percentages may not total 100 because of rounding.

Table 2. Type of Physical Activity, by Location, Among Adults (N = 297) in Rural North Carolina During the COVID-19 Pandemic, August 3–September 15, 2020[a]

Variable	Parks/trails, n = 183	Neighborhood, n = 176	Home, n = 208
Walking/hiking	182 (99.5)	173 (98.3)	—
Jogging/running	74 (40.4)	50 (28.4)	—
Biking	70 (38.3)	64 (36.4)	—
Playing with kids	82 (44.8)	—	—
Sports	29 (15.8)	29 (16.5)	—
Kayaking/canoeing	79 (43.2)	—	—
Swimming	47 (25.7)	—	—
Indoor housework	—	—	141 (67.8)
Workouts/exercise	—	—	119 (57.2)
Playing in the yard	—	—	116 (55.8)
Outdoor sports/games	—	—	84 (40.4)
Yard work/landscaping	—	—	126 (60.6)
Gardening	—	—	108 (51.9)

Abbreviation: —, Variable was not offered as a response for the listed location.
[a] Values are number (percentage).

Table 3. Reason for Physical Activity, by Location, Among Adults (N = 297) in Rural North Carolina During the COVID-19 Pandemic, August 3–September 15, 2020[a]

Variable	Parks/trails, n = 183	Neighborhood, n = 183	Home, n = 208
For exercise	162 (88.5)	153 (83.6)	138 (66.3)
To experience nature	141 (77.0)	105 (57.4)	—
To get out of the house	174 (95.1)	163 (89.1)	—
To do something normal	123 (67.2)	106 (57.9)	125 (60.1)
For mental health	146 (79.8)	129 (70.5)	155 (74.5)
To take the kids outside	113 (61.7)	107 (58.5)	99 (47.6)
To walk the dog	92 (50.3)	106 (57.9)	—
To connect with others	46 (25.1)	44 (24.0)	28 (13.5)
To maintain health/fitness	139 (76.0)	113 (61.7)	158 (76.0)
To prepare for future events	13 (7.1)	9 (4.9)	17 (8.2)
To reduce sitting time	124 (67.8)	113 (61.7)	134 (64.4)

Abbreviation: —, Variable was not offered as a response for the listed location.
[a] Values are number (percentage).

Table 4. Perceived Barriers to Physical Activity Among Adults (N = 203) in Rural North Carolina During the COVID-19 Pandemic, August 3–September 15, 2020[a]

Response	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
Barriers					
Ability to maintain appropriate social distancing	44 (21.7)	48 (23.6)	27 (13.3)	51 (25.1)	30 (14.8)
Closure of parks and recreation resources	13 (6.4)	21 (10.3)	23 (11.3)	79 (38.9)	67 (33.0)
Cancellation of sports or intramurals	45 (22.2)	13 (6.4)	46 (22.7)	30 (14.8)	69 (34.0)
Neighborhood environment	77 (37.9)	50 (24.6)	44 (21.7)	21 (10.3)	11 (5.4)
Closure of fitness facility	41 (20.2)	12 (5.9)	57 (28.1)	33 (16.3)	58 (28.6)
Lack of fitness equipment	45 (22.2)	36 (17.7)	47 (23.2)	43 (21.2)	32 (15.8)
Stay-at-home orders	30 (14.8)	39 (19.2)	33 (16.3)	54 (26.6)	49 (29.1)
Benefits					
Is something I can do to leave the house without being exposed to COVID-19	8 (3.9)	11 (5.4)	26 (12.8)	93 (45.8)	65 (32.0)
Is something routine or normal that I can still do while social distancing	6 (3.0)	9 (4.4)	13 (6.4)	104 (51.2)	71 (35.0)
Helps manage stress during the COVID-19 pandemic	3 (1.5)	2 (1.0)	6 (3.0)	83 (40.9)	109 (53.7)
Helps manage my mental health (eg, stress, anxiety) during the COVID-19 pandemic	1 (0.5)	1 (0.5)	6 (3.0)	79 (38.9)	114 (56.2)
Helps me manage my chronic condition(s) (eg, diabetes, hypertension, heart disease)	8 (3.9)	3 (1.5)	78 (138.4)	60 (29.6)	48 (23.6)
Is beneficial to my immune system during the COVID-19 pandemic	3 (1.5)	0 (0)	13 (6.4)	92 (45.3)	95 (46.8)
Will help reduce my risk of getting COVID-19	12 (5.9)	17 (8.4)	64 (31.5)	53 (26.1)	56 (27.6)
Will reduce the severity of COVID-19 if I get it	8 (3.9)	15 (7.4)	63 (31.0)	65 (32.0)	51 (25.1)

[a] Number reflects survey attrition. Values are number (percentage).

PREVENTING CHRONIC DISEASE

PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY

Volume 19, E65 OCTOBER 2022

ORIGINAL RESEARCH

Spatial Analysis of Breast Cancer Mortality Rates in a Rural State

Marisa Schulz, MSc[1]; Emma Spors, MSc[1]; Kari Bates[2]; Semhar Michael, PhD[1]

Accessible Version: www.cdc.gov/pcd/issues/2022/22_0113.htm

Suggested citation for this article: Schulz M, Spors E, Bates K, Michael S. Spatial Analysis of Breast Cancer Mortality Rates in a Rural State. Prev Chronic Dis 2022;19:220113. DOI: https://doi.org/10.5888/pcd19.220113.

PEER REVIEWED

Summary

What is known on this topic?

Breast cancer affects 1 in 8 women in the US and is the most frequently diagnosed cancer in women.

What is added by this report?

Because South Dakota is a rural state, sociodemographic factors affect the population differently than in the general US population. We assessed the spatial distribution of breast cancer mortality rates by county, and our findings add insight on educational attainment as a risk factor for breast cancer.

What are the implications for public health practice?

Our results can be used to help allocate resources to the South Dakota counties that need them most.

Abstract

Introduction

Breast cancer affects 1 in 8 women in the US and is the most frequently diagnosed cancer in women. In South Dakota, 102 women die from breast cancer each year. We assessed which sociodemographic factors contributed to mortality rates in South Dakota and used spatial analysis to investigate how counties' observed age-adjusted mortality rates compared with expected rates.

Methods

We computed standardized incidence ratios (SIRs) of all counties in South Dakota by using the age-adjusted mortality rates, the 2000 US standard population, and the South Dakota estimated population. We used a linear regression model to identify sociodemographic factors associated with breast cancer mortality rates and to compute a new SIR value, after controlling for relevant factors.

Results

Educational level and breast cancer incidence rates were significantly associated with breast cancer mortality rates at the county level. The SIR values based on age-adjusted counts showed which counties had more deaths due to breast cancer than what might be expected using South Dakota as the reference population. After controlling for sociodemographic factors, the range of SIR values decreased and had lower variability.

Conclusion

The regression model helped identify factors associated with mortality and provided insights into which risk factors are at play in South Dakota. This information, in combination with the spatial distribution of mortality by county, can be used to help allocate resources to the counties in South Dakota that need them most.

Introduction

In South Dakota, breast cancer is the most commonly diagnosed cancer and the second leading cause of cancer death among women (1,2). In 2022, an estimated 750 new cases and 110 deaths attributed to female breast cancer will occur in South Dakota. In general, a woman in the US has a 1-in-8 lifetime risk of developing breast cancer (3,4). Since 1989, the US breast cancer mortality rate has decreased 40%, but from 2010 to 2019 the rate slowed to a low of decreasing by 1.3% per year (5).

Characteristics such as age and race and ethnicity affect a woman's chances of being diagnosed with or dying of breast cancer, but new evidence has established that sociodemographic factors, including education level, also play a role (6). Albano et al noted a negative relationship between number of years of education and cancer mortality and found that the level of education and race vary considerably with mortality rates (7). Of the South Dakota population aged 25 years or older, 92.2% are high school graduates (higher than the national average) and 29.3% have a bachelor's degree or higher (lower than national average) (8).

Olson et al acknowledged that communities exist in which geographic disparities are more prominent because of rural isolation and small population size (9). Furthermore, 64 of 66 counties in South Dakota are categorized as rural or frontier, and South Dakota contains 9 American Indian reservations (10). Finally, 61.6% of women receiving breast services are White, and 16.7% are American Indian (11); most of the population in South Dakota is White, and the leading minority is American Indian at 8.8% (8).

The study aimed to describe the spatial distribution of female breast cancer mortality at the county level in South Dakota and assess the association between mortality rates and risk factors reported in the literature.

Methods

Data source

The 66 counties of South Dakota have boundaries that are defined by the South Dakota Legislature and accepted by the US Census. The counties range in population from 183,439 in Minnehaha County to 917 in Jones County, and the median population per county is 5,413. Most residents of South Dakota were White; the median percentage of non-White residents by county was 6.6% and the maximum was 95.2%

Cancer data

Breast cancer incidence and mortality rates from 2001 through 2015 were extracted from the South Dakota State Cancer Registry provided by the South Dakota Department of Health (12). Both rates were per 100,000 people and age-adjusted to the 2000 US standard population and the South Dakota estimated population. The proportion of mammography screening rates in South Dakota was based on the numbers reported by Holzhauser et al for the All Women Count! mammography program (13). The average number of participants for 1997 through 2016 was reported for the program; then, the average number of participants was adjusted for the total number of women older than 40 years in the county to get an estimated screening rate for each county (13,14).

Demographics

We used 2015 data from the US Census Bureau to obtain information on the 66 South Dakota counties, including the number of providers and the education level, poverty level, percentage of uninsured, median age, and race of residents (8).

Data on educational attainment were obtained from the US Census Bureau's American Community Survey (ACS). These data were count estimates for the population of each county aged 25 years or older. Levels were categorized as less than 9th grade, 9th through 12th grade with no diploma, high school graduate or equivalent, some college but no degree, associate degree, bachelor's degree, and graduate or professional degree. These values were modified into an educational attainment statistic of the percentage of the population with less than a bachelor's degree of education. The statistic used in this study was the percentage of the population with less than a bachelor's degree, by county.

We collected data on poverty estimates, by county, from the ACS; these data adhered to the standards specified by the Office of Management and Budget in Statistical Policy Directive 14 (8). Poverty was determined by a set of income thresholds that consider the living situation (alone or with nonrelatives), age, and number of people per household. For example, the poverty threshold for 2-person families varies by the age of the primary householder and differs from the poverty threshold for people living alone or with nonrelatives, which also varies by age.

Insurance coverage percentages were collected from Small Area Health Insurance Estimates (SAHIE) (15). The uninsured percentage included residents who were not covered by insurance, which excluded those on government assistance such as Medicaid or Medicare. Finally, the data set summarizing racial distributions in a county included estimated population counts for American Indian and Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, other race, and 2 or more races. Because of South Dakota's predominantly White population, the data were configured into White and non-White, which determined the non-White percentage per county (8) (Table 1).

Statistical analysis

Data manipulation and missing value imputation

Mortality rates and the various independent variables were combined into 1 data set; 15 of the 66 counties were missing mortality rate data. Mortality rates are often suppressed from public availability when 3 or fewer deaths are reported in a county, to protect patient identity. To remedy the missing data, k-nearest-neighbor (KNN) imputation was used to estimate the missing mortality values. KNN imputation compares a data point x_i with its k nearest neighbors and then approximates x_i using the majority vote of these k neighbors in multidimensional space. For the data, $k = 9$ nearest neighbors were used, and a weighted mean of the k nearest values was placed for each missing x_i (16,17). This was done with the function "knn()" from the R package VIM version 6.1.1 (R Foundation for Statistical Computing) (18).

Multiple linear regression

We used multiple linear regression to model the relationship between the factors in this study and breast cancer mortality rates

(19). We considered several potential predictor variables with observed correlation, hence a stepwise variable selection technique was used, in both the forward and backward direction, to perform feature selection. As a result, a subset of the factors that were associated with the mortality rates was obtained based on Akaike information criterion (AIC) (19,20).

The resulting model is of the form $\widehat{Y_i} = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \cdots + \hat{\beta}_p X_{ji}$ for $i = 1, 2, \ldots, 66$ where $\widehat{Y_i}$ represents the estimated mortality rate for the i^{th} county and $\hat{\beta}_0$ is the intercept. The variables X_{1i} through X_{ji} represent the values of the factors for the i^{th} county and $\hat{\beta}_1$ through $\hat{\beta}_p$ are the coefficients that were estimated using the least squares regression method (19,21). To explore the data, 4 linear regression models were created, which differ by factors included in the model.

The first regression model included all 8 factors as prediction variables resulting in Model 1. The regression model was then fitted by stepwise variable selection in both the forward and backward direction by using AIC as a model selection criterion. AIC rewards goodness of fit and penalizes the model's complexity (19). This was done by using the R package MASS version 7.3–54 (R Foundation for Statistical Computing) (22). The simplified model resulted in Model 2. At this point, a decision was made to remove the incidence rate from the data to better see how the other sociodemographic factors contributed to breast cancer outcomes, resulting in Model 3. Model 3 was then fitted with stepwise variable selection, yielding Model 4. To best compare the expected mortality to the observed mortality with all predictors available, Model 2 was chosen as the final model. Model 2 was then used to predict the expected mortality rates for the second standardized incidence ratio (SIR) that was computed (SIR_{LM}).

Standardized incidence ratio

A SIR was used to compare the spatial distribution of counties in terms of mortality rates due to breast cancer. In general, the SIR compares the expected value of deaths to the observed value of deaths in a county. This is calculated with $SIR_j = \frac{O_j}{E_j}$ where O_j is the observed value for county j, and E_j is the expected value for county j. A SIR of greater than 1 means that there were more deaths than expected, whereas a SIR of less than 1 means that there were fewer deaths than expected for that county.

Because age-specific mortality rates by county were not available, the observed age-adjusted counts per county were computed from the age-adjusted rates per county as follows: $O_j^* = \frac{O_j}{100{,}000} y_j$ and $E_j^* = \left(\frac{\sum_{j=1}^{66} O_j}{100{,}000} \cdot y_j\right) \cdot \frac{y_j}{\sum_{j=1}^{66} y_j}$, where O_j is the j^{th} county's observed age-adjusted mortality rate and y_j is the j^{th} county's population. The expected count was computed to be on the same scale as the observed count. These manipulations were used only to calculate an age-adjusted count SIR, referred to as SIR_{COUNT}. All other analysis of the data was completed with the original variable (O_j), as described in the data source.

A SIR was calculated on the mortality in each South Dakota county (N = 66). To account for the age adjustment of the data, 2 different SIRs were found: using the age-adjusted mortality count for South Dakota and using the expected rate obtained from the linear regression model. The first SIR accounted only for the age adjustment, and the second SIR accounted for more factors related to breast cancer. For example, the first SIR used the age-adjusted mortality count for a county for both observed and expected (SIR_{COUNT}). The second SIR (SIR_{LM}) used the mortality rates by county, which were per 100,000 persons and age-adjusted to the 2000 US standard population, and the South Dakota estimated population for observed and predicted mortality rate from the linear regression model for expected, which accounted for incidence rates and educational attainment. All statistical analyses used R version 4.1.2 and RStudio version 2021.09.2 build 382 (R Foundation for Statistical Computing) (23,24).

Results

Exploratory data analysis

To learn more about the data, we performed an Exploratory Data Analysis (EDA) using several techniques. We explored the geographic distribution of the breast cancer mortality rates by using a choropleth map of South Dakota with age-adjusted mortality and incidence rates (Figure 1). The eastern side of the state had lower and more consistent mortality rates, followed by the far western part of the state. The central west part of the state exhibited higher mortality rates.

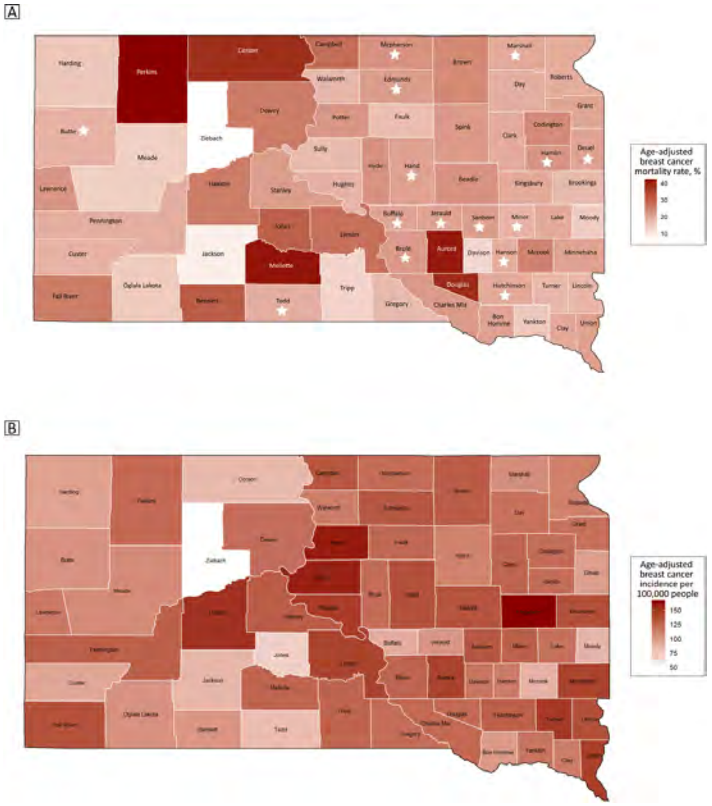


Figure 1. Map A shows the age-adjusted breast cancer mortality rates and Map B shows the age-adjusted breast cancer incidence rates, by county (N = 66), South Dakota, 2001–2015. Counties whose mortality rates have been imputed are marked with a star. Source: South Dakota State Cancer Registry, South Dakota Department of Health (12).

Five counties had higher mortality rates than the rest of the counties: Perkins, Mellette, Aurora, Douglas, and Corson. Two counties had lower mortality rates than the rest of the counties: Ziebach and Jackson. Eight counties (Bennet, Buffalo, Corson, Jackson, Mellette, Oglala, Todd, and Ziebach) had a poverty percentage greater than 30%, which was distinctly higher than the others; each county contains land on an American Indian Reservation. Minnehaha and Pennington counties had high screening rates; these counties are home to the first- and second-largest cities in South Dakota, respectively, so they also had the greatest number of screening providers.

All variables, besides the number of providers and screening rate, had a correlation greater than zero with mortality rate. Incidence, educational attainment, and uninsured percentage all had a low positive correlation with mortality rate. The highest correlation with mortality rate was education attainment, at a value of 0.26. The remaining variables had a negligible correlation with mortality rate.

Regression analysis

Model 2 had the highest adjusted R^2 value (0.10) and the lowest AIC (441.79), with 2 significant factors associated with mortality rate; Model 3 had the lowest adjusted R^2 (.004) and the highest AIC (453.32), with no significant predictors of mortality. In Model 2, breast cancer incidence and educational attainment were predictors of breast cancer mortality, indicating that as more people are diagnosed with breast cancer and as the percentage of people with less than a bachelor's degree increases, breast cancer mortality rate increases (Table 2). Educational attainment was a predictor of mortality in all models, and the educational attainment statistic had a $P < .001$.

Standardized incidence ratio

Thirty-five of the 66 counties had a SIR_{COUNT} greater than 1, meaning that more than half of the counties had more deaths than expected (Figure 2). The 5 counties with the highest SIR_{COUNT}, in decreasing order, were Perkins, Mellette, Aurora, Douglas, and Corson. Of those, Perkins, Mellette, and Aurora had more than twice the expected number of deaths. Ziebach, Jackson, Davison, Tripp, and Meade counties had the lowest SIR_{COUNT}, with Ziebach and Jackson both being less than half the expected number of deaths. The highest SIR_{COUNT} was 2.15 and the lowest was 0.31.

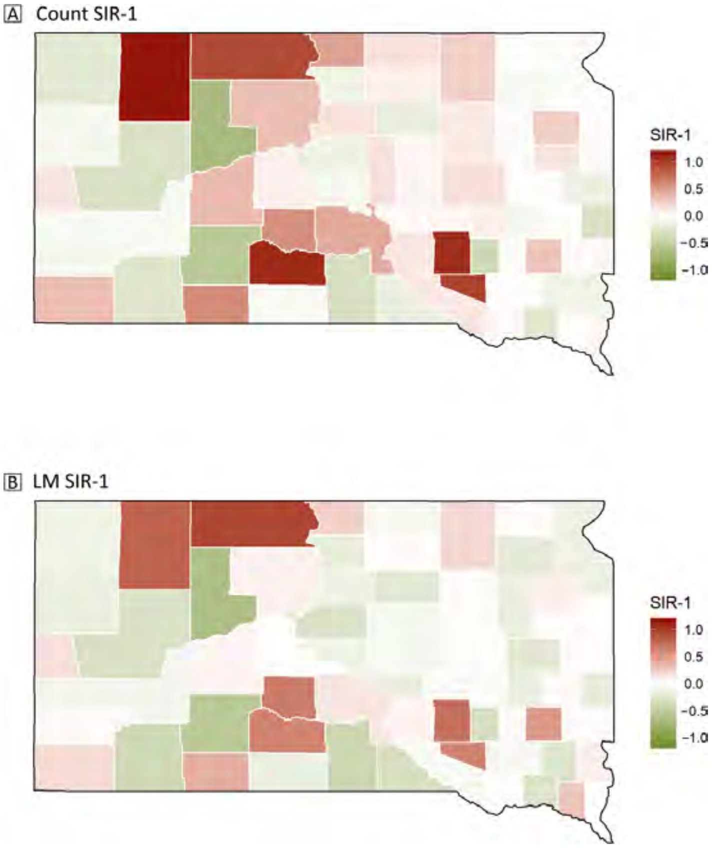


Figure 2. Map A shows the predicted breast cancer mortality rate of South Dakota counties, accounting for age-adjustment of the data, and Map B shows the predicted breast cancer mortality rate of South Dakota counties, accounting for age-adjustment, incidence rate, and educational attainment. Abbreviation: SIR, standardized incidence ratio. Map Sources: South Dakota State Cancer Registry, South Dakota Department of Health (12), Holzhauser et al (13), and the US Census Bureau (8).

On the other hand, for the SIR_{LM}, most counties had fewer deaths than expected, with 28 of 66 counties having a SIR_{LM} greater than 1. The 5 counties with the highest SIR_{LM} were similar to the SIR_{COUNT}: Corson, Perkins, Aurora, Jones, and Mellette in decreasing order. No county had more than twice the expected number of deaths. The 5 counties with the lowest SIR_{LM} were Ziebach, Jackson, Tripp, Oglala Lakota, and Davison. Ziebach and Jackson counties had less than half the number of expected deaths. The highest SIR_{LM} was 1.94 and the lowest SIR_{LM} was 0.35.

The results of the SIRs (SIR_{COUNT} and SIR_{LM}) are presented in Figure 2. The eastern side of the state showed similar SIR values while the western side of the state had more variation. Perkins, Mellette, Aurora, Douglas, and Corson counties had SIR_{COUNT} values that were much higher than those of the rest of the counties. Ziebach and Jackson counties had the lowest SIR_{COUNT} values. These counties with the highest SIR_{COUNT} made up 5 of the 6 counties with the highest SIR_{LM} values, with Jones County replacing Douglas County. The 2 counties with the lowest SIR_{COUNT} values were the same 2 counties with the lowest SIR_{LM} values.

Discussion

Overall, we found a significant association between incidence rate and educational level with respect to breast cancer mortality rates. Breast cancer incidence was positively associated with mortality rates in South Dakota, which suggests that more breast cancer cases are associated with more breast cancer deaths. In addition, educational attainment was repeatedly identified as a significant factor for mortality. Gadeyne et al found inconclusive results in their study of breast cancer mortality and education; however, Albano et al found a significant association between educational levels and cancer in general (7,25), specifically that lower educational attainment was related to higher cancer mortality rates, reflecting the findings of this study. Race, median age, and number of women screened were not selected in the feature selection during stepwise regression in our study; similarly, race, median age, and number of women screened were not significant in our full model.

The 2020–2021 South Dakota Department of Education yearly review stated that American Indians were the largest minority group in school. However, American Indians still have a 63% completion rate for high school graduation and 59% attendance rate, compared with Whites who have a 94% completion rate for high school graduation and 94% attendance rate (26). An interesting point to consider is that South Dakota has no set standards for sex education (27). Thus, students are not taught reproductive health in general, including the importance of breast examinations, Pap smears, or prostate examinations. We advocate that set and scientifically backed health standards in high school would expose students at an early age to the risks of breast cancer and their options for screening.

The western half of South Dakota had more variability in SIR values, and the state's demographics could be a possible explanation. The 4 counties with the highest SIRs for both count and linear model SIRs were Corson, Perkins, Aurora, and Mellette, which are either in an American Indian reservation or neighbor a county within an American Indian reservation. Research on 3 tribes in western South Dakota supported that trust is often a barrier for American Indians (as are remote location and approvals by Indian Health Service programs) (28). Research in New Mexico reported that even after in-depth implementation of screening programs that lowered the barriers of cost, availability, and access to Native American and Hispanic women, the screening rates remained low, under 40% of women annually (26). The high SIRs in or neighboring reservation counties may mean that trust is also an issue, and South Dakota has more to work on than accessibility to Native Americans.

After controlling for incidence rate and educational attainment, the SIR_{LM} values became less variable. The SIRs' decrease in range and mean closer to 1 indicate that the factors did affect mortality rate. This again agrees with findings from Albano et al that educational attainment affects mortality (7).

We found that some counties had a higher mortality rate than expected based on the age of the women in the county. Ziebach and Jackson counties had the highest mortality rates, and the counties with the lowest SIRs are not home to major medical centers. Haakon County is vertically between Ziebach and Jackson counties and is one of the counties that does not have a provider; however, Haakon County has a higher mortality rate than the average of counties of South Dakota and both SIRs greater than 1, which means there were more deaths than expected. The areas of the map where there are dark green counties next to dark red counties are either on an American Indian reservation or neighbor an American Indian reservation. The differences between counties do not come from any singular cause, but rather due to variations in race, poverty levels, and population size.

Our study has limitations, primarily in the absence of portions of mortality data. Because South Dakota is largely a rural state, several counties have small populations and see very few deaths from breast cancer. These numbers are then withheld from the public to protect the privacy of the patients. This suppression resulted in having to impute the mortality rates of 15 counties, possibly introducing errors. The assumptions of the regression model did not account for this error, which may confer bias on the results. The counties with American Indian reservations have another health system that could have resulted in the under-representation or over-representation of breast cancer deaths in those counties (29). In addition, a study found that the misclassification of the race of Native Americans caused an underestimation of mortality rates as well (30). These gaps in databases and their contents highlight research challenges that rural communities will continue to face when few data are collected, populations and incidence are sparse, and data are inconsistently collected by multiple sources. Sociodemographic data are also challenging to consistently collect throughout a state. More detailed data per county would help yield accurate and unbiased results. For example, considering education, Zajacova and Lawrence argue that education is not a single-generation factor (31). Having data on the educational attainment of a patient's parents or family, in addition to their own educational attainment, would allow us to assess the risk and see the relationship between education and breast cancer incidence or mortality. Thus, more research is needed to understand the effects education level has on financial security, stable employment, social success, and in turn, breast cancer mortality.

In conclusion, understanding the risk factors and geographic distribution of breast cancer mortality among women across the state will assist stakeholders with efforts at prevention and resource allocation guided by data.

Acknowledgments

This research was partially supported by the Joseph F. Nelson mentorship and undergraduate award and a Van D. and Barbara B. Fishback honors college undergraduate Research Award from South Dakota State University. The contents are solely the responsibility of the authors. No copyrighted figures, surveys, instruments, or tools were used in this study.

Author Information

Corresponding Author: Marisa Schulz, MSc, Mathematics and Statistics Department, South Dakota State University, AME Building, Box 2225, Brookings, South Dakota 57006. Telephone: 605-688-6316. Email: Marisa.schulz@jacks.sdstate.edu.

Author Affiliations: [1]Mathematics and Statistics Department, South Dakota State University, Brookings, South Dakota. [2]Sanford Research, Research Design and Biostatistics Core, Sioux Falls, South Dakota.

References

1. American Cancer Society. Cancer facts and statistics. Accessed March 21, 2022. http://cancerstatisticscenter.cancer.org/
2. Centers for Disease Control and Prevention. United States Cancer Statistics (USCS) data visualizations. Accessed September 13, 2022. https://gis.cdc.gov/Cancer/USCS/?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcancer%2Fdataviz%2Findex.htm#AtAGlance/
3. DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast cancer statistics, 2017, racial disparity in mortality by state. CA Cancer J Clin 2017;67(6):439–48.
4. Armstrong K, Eisen A, Weber B. Assessing the risk of breast cancer. N Engl J Med 2000;342(8):564–71.
5. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, et al. Breast cancer statistics, 2019. CA Cancer J Clin 2019;69(6):438–51.

6. Hussain SK, Altieri A, Sundquist J, Hemminki K. Influence of education level on breast cancer risk and survival in Sweden between 1990 and 2004. Int J Cancer 2008;122(1):165–9.
7. Albano JD, Ward E, Jemal A, Anderson R, Cokkinides VE, Murray T, et al. Cancer mortality in the United States by education level and race. J Natl Cancer Inst 2007;99(18):1384–94.
8. US Census Bureau. Office of Management and Budget in statistical policy directive 14. Accessed August 2, 2022. https://www.census.gov/topics/income-poverty/poverty/about/history-of-the-poverty-measure/omb-stat-policy-14.html#:~:text=RelatedSites-,OfficeofManagementandBudget(OMB)inStatisticalPolicyDirective,andestablishmentsforstatisticalpurposes
9. Olson J, Cawthra T, Beyer K, Frazer D, Ignace L, Maurana C, et al. Community and research perspectives on cancer disparities in Wisconsin. Prev Chronic Dis 2020;17:200183.
10. South Dakota Department of Health. South Dakota comprehensive cancer control state plan, 2015–2020.
11. Centers for Disease Control and Prevention. National Breast and Cervical Cancer Early Detection Program screening program summaries: South Dakota. Accessed March 22, 2022. https://www.cdc.gov/cancer/nbccedp/data/summaries/south_dakota.htm
12. South Dakota cancer incidence and mortality. Accessed March 22, 2022. https://www.sdcancerstats.org/
13. Holzhauser C, Da Rosa P, Michael S. Forecasting participants in the All Women Count! mammography program. Prev Chronic Dis 2018;15:180177.
14. South Dakota Department of Health. All Women Count! cancer control program. Accessed March 21, 2022. https://getscreened.sd.gov/count/
15. US Census Bureau. Small Area Health Insurance Estimates (SAHIE) program. Accessed August 2, 2022. http://www.census.gov/programs-surveys/sahie.html
16. Wasserman L. Classification. In: All of statistics: a concise course in statistical inference. New York (NY): Springer; 375 pages.
17. Peterson LE. K-nearest neighbor. Scholarpedia 2009;4(2):1883.
18. Kowarik A, Templ M. Imputation with the R Package VIM. J Stat Softw 2016;74(7):1–16.
19. Bruce PC, Bruce A, Gedeck P. Practical statistics for data scientists: 50+ essential concepts using R and Python. Second edition. Sebastopol (CA): O’Reilly Media, Inc; 2020.
20. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. Springer series in statistics. Second International Symposium on Information Theory; 1973. Pages 267–281.
21. Kutner MH, Nachtsheim C, Neter J. Applied linear regression models. New York (NY): McGraw-Hill/Irwin; 2004.
22. Venables WN, Ripley BD, Venables WN. Modern applied statistics with S. 4th edition. New York (NY): Springer; 2002.
23. R: the R project for statistical computing. Accessed March 22, 2022. https://www.r-project.org/
24. RStudio. Open source and professional software for data science teams. Accessed March 22, 2022. https://www.rstudio.com/
25. Gadeyne S, Menvielle G, Kulhanova I, Bopp M, Deboosere P, Eikemo TA, et al. The turn of the gradient? Educational differences in breast cancer mortality in 18 European populations during the 2000s. Int J Cancer 2017;141(1):33–44.
26. Paskett ED, Tatum C, Rushing J, Michielutte R, Bell R, Foley KL, et al. Racial differences in knowledge, attitudes, and cancer screening practices among a triracial rural population. Cancer 2004;101(11):2650–9.
27. South Dakota Department of Education. South Dakota health education standards; 2018. Accessed July 24, 2022. https://doe.sd.gov/board/packets/documents/112017/item5doc2.pdf
28. Rogers D, Petereit DG. Cancer disparities research partnership in Lakota Country: clinical trials, patient services, and community education for the Oglala, Rosebud, and Cheyenne River Sioux tribes. Am J Public Health 2005;95(12):2129–32.
29. Office of Minority Health. Profile: American Indian/Alaska Native. Published 2018. Accessed March 25, 2022. https://www.minorityhealth.hhs.gov/omh/browse.aspx?lvl=3&lvlid=62
30. Espey DK, Jim MA, Richards TB, Begay C, Haverkamp D, Roberts D. Methods for improving the quality and completeness of mortality data for American Indians and Alaska Natives. Am J Public Health 2014;104(Suppl 3):S286–94.
31. Zajacova A, Lawrence EM. The relationship between education and health: reducing disparities through a contextual approach. Annu Rev Public Health 2018;39(1):273–89.

Tables

Table 1. Factors for Regression Analysis, Study on Breast Cancer Mortality in South Dakota, 2001–2015

Factor	Description
Breast cancer incidence rates	Age-adjusted breast cancer incidence rate per 100,000 persons
Breast cancer screening rates	Estimated mammography screening rates
Number of screening providers	The number of medical providers per county that provide breast cancer screening
Poverty level	The percentage of residents living in poverty per county
Insurance status	The percentage of residents without any insurance
Median age	Median age of the county’s residents
Race	The percentage of residents that are non-White
Educational attainment	The percentage of the population aged 25 years or older with less than a bachelor’s degree

Table 2. Regression Models, Study on Breast Cancer Mortality in South Dakota, 2001–2015

Coefficient	Model 1	Model 2	Model 3	Model 4
	Estimate (95% CI)			
Intercept	−15.375 (−39.388 to 8.638)	−11.113 (−32.253 to 10.026)	−2.248 (−23.425 to 18.929)	4.114 (−12.158 to 20.386)
Median age	−0.059 (−0.481 to 0.363)	—	−0.043 (−0.478 to 0.391)	—
Non-white percentage	−15.119 (−37.725 to 7.486)	—	−11.266 (−34.233 to 11.700)	—
Poverty percentage	0.208 (−0.292 to 0.708)	—	0.030 (−0.455 to 0.515)	—
Uninsured percentage	0.532 (−0.275 to 1.340)	—	0.418 (−0.406 to 1.242)	—
Number of providers	0.210 (−1.318 to 1.738)	—	0.458 (−1.097 to 2.012)	—
Screened	−0.002 (−0.013 to 0.009)	—	−0.003 (−0.015 to 0.008)	—
Incidence	0.101[a] (0.006 to 0.197)	0.085[a] (0.007 to 0.163)	—	—
Educational attainment	0.286 (−0.100 to 0.672)	0.333[b] (0.088 to 0.578)	0.312 (−0.085 to 0.708)	0.258[a] (0.016 to 0.499)
AIC	450.30	441.79	453.32	444.55
Adjusted R^2	0.06	0.10	0.004	0.05

Abbreviation: — , not applicable; AIC, Akaike information criterion.
[a] $P = .01$.
[b] $P = .001$.

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
Volume 19, E72 NOVEMBER 2022

ORIGINAL RESEARCH

Differences in COVID-19 Hospitalizations by Self-Reported Race and Ethnicity in a Hospital in Honolulu, Hawaii

Brendan K. Seto, BA[1,2]; Laura Nishizaki, LSW[2]; Gerard Akaka, MD[2]; Jo Ann Kimura, MSW[2]; Todd B. Seto, MD, MPH[1,2]

Accessible Version: www.cdc.gov/pcd/issues/2022/22_0114.htm

Suggested citation for this article: Seto BK, Nishizaki L, Akaka G, Kimura JA, Seto TB. Differences in COVID-19 Hospitalizations by Self-Reported Race and Ethnicity in a Hospital in Honolulu, Hawaii. Prev Chronic Dis 2022;19:220114. DOI: https://doi.org/10.5888/pcd19.220114.

PEER REVIEWED

Summary

What is already known on this topic?

Racial and ethnic disparities in the number of COVID-19 hospitalizations exist, and data on race and ethnicity in hospital electronic medical records are known to be inaccurate for non-White populations.

What is added by this report?

We described the inaccuracy of race and ethnicity identification in a large multiracial population, then projected these findings onto publicly available COVID-19 hospitalization data to estimate disparities by self-identified, rather than hospital electronic medical record–based, race and ethnicity.

What are the implications for public health practice?

Accurate race and ethnicity data are essential for reliably measuring disparities. Race and ethnicity data, especially in multiracial populations, should be confirmed when possible, and reporting practices could be evaluated to promote reliable results.

Abstract

Introduction

The true extent of racial and ethnic disparities in COVID-19 hospitalizations may be hidden by misclassification of race and ethnicity. This study aimed to quantify this inaccuracy in a hospital's electronic medical record (EMR) against the gold standard of self-identification and then project data onto state-level COVID-19 hospitalizations by self-identified race and ethnicity.

Methods

To identify misclassification of race and ethnicity in the EMRs of a hospital in Honolulu, Hawaii, research and quality improvement staff members surveyed all available patients (N = 847) in 5 cohorts in 2007, 2008, 2010, 2013, and 2020 at randomly selected hospital and ambulatory units. The survey asked patients to self-identify up to 12 races and ethnicities. We compared these data with data from EMRs. We then estimated the number of COVID-19 hospitalizations by projecting racial misclassifications onto publicly available data. We determined significant differences via simulation-constructed medians and 95% CIs.

Results

EMR–based and self-identified race and ethnicity were the same in 86.5% of the sample. Native Hawaiians (79.2%) were significantly less likely than non–Native Hawaiians (89.4%) to be correctly classified on initial analysis; this difference was driven by Native Hawaiians being more likely than non–Native Hawaiians to be multiracial (93.4% vs 30.3%). When restricted to multiracial patients only, we found no significant difference in accuracy ($P = .32$). The number of COVID-19–related hospitalizations was 8.7% higher among Native Hawaiians and 3.9% higher among Pacific Islanders when we projected self-identified race and ethnicity rather than using EMR data.

Conclusion

Using self-identified rather than hospital EMR data on race and ethnicity may uncover further disparities in COVID-19 hospitalizations.

Introduction

Despite efforts to address health inequities, there are persistent — and sometimes substantial — disparities in health among some racial and ethnic groups in the US and worldwide (1). Efforts to understand the magnitude and causes of these disparities are often complicated by the lack of consensus on how one's race and ethnicity are defined. Although several approaches exist (2–4), the most common method and the current gold standard for determining race and ethnicity is self-identification, with federal best practices from the Centers for Medicare and Medicaid Services (CMS) available to guide standardized data collection, including information to address key challenges in collecting these data (5).

However, the accuracy of data on race and ethnicity in large administrative data sets may be lacking, especially in non-White patient populations; the accuracy of such data is estimated to be 88% among the US patient population overall and 66% in non-White patient populations (6–17). Accuracy may be even less reliable among the increasing number of people who identify as multiracial, which is especially common among young people. To our knowledge, only 2 studies have investigated the accuracy of racial and ethnic information of multiracial populations in hospitals (6,13). Both studies showed less accuracy in correctly identifying race and ethnicity among multiracial patients (21% accuracy) than among nonmultiracial patients (65% accuracy) in the electronic medical record (EMR), although the number of multiracial patients in both studies was small (0.4% and 4.3% of the patient population). US Census data show that people younger than 18 years are nearly twice as likely as people aged 18 years or older to identify as multiracial (15.1% vs 8.8%) and that the number of people identifying as multiracial increased by 276% from 2010 to 2020 (18). Thus, the challenges of identifying a person's race and ethnicity will continue to grow. It is important to explore the implications of these challenges by studying their potential impact in a highly diverse, multiracial, majority–minority population.

COVID-19 has presented a challenge to our modern lives, but many racial and ethnic minority groups have a disproportionate burden of cases, hospitalizations, long-term complications, and deaths (19,20). As of March 2021, Native Hawaiians and Pacific Islanders had the highest death rate of any racial or ethnic group in 18 of the 20 states that reported deaths among those 2 groups (21). In Hawaii, Pacific Islanders account for 5% of the population but 22% of COVID-19 cases and deaths (22). A fundamental requirement for understanding the magnitude and causes of these disparities is accurate data on race and ethnicity. For example, publicly available data from the Hawaii Department of Health show that Native Hawaiians and Pacific Islanders make up a disproportionate number of COVID-19 hospitalizations. But this number may have been underreported — and the true disparities underestimated — if the race and ethnicity of some patients have been misclassified.

The objective of our study was to determine the accuracy of race and ethnicity data in a hospital EMR system compared with self-identified data and then use this information to determine how the magnitude of COVID-19 disparities among racial and ethnic groups would change if patients were correctly classified.

Methods

Accuracy of EMR-based data on race and ethnicity vs self-identification

The study population consisted of patients at The Queen's Medical Center (QMC), a 500-bed university-affiliated tertiary care hospital in Honolulu, Hawaii. QMC is the largest health care system in the state and serves as the primary referral center for the Pacific Basin.

We obtained survey data from QMC. Both inpatients and outpatients were recruited to participate in a survey administered by trained data collectors who visited randomly selected hospital and ambulatory units and asked all available patients if they would participate. Data were collected in 5 cohorts as part of an ongoing quality assurance project conducted by hospital staff during 5 years from 2007 through 2020 (2007, 2008, 2010, 2013, 2020). The major inclusion criterion was QMC patients who were provided care on the day of data collection. Non–English-speaking patients were included if a friend or family member was able to interpret. We excluded patients who were in intensive care units, unable to respond verbally, declined participation, or lacked an accessible EMR at QMC at the time of data collection. The QMC institutional review board approved the study protocol.

Patients were first asked to list all their races and ethnicities. Twelve spaces were provided for entries, but no patient listed more than 10 races or ethnicities. They were then asked to select the one that they identified with the most; this was defined as the self-identified race and ethnicity. Patient responses were aggregated to modified 1997 OMB minimal reporting guidelines that disaggregated Native Hawaiians and Pacific Islanders into 2 groups to mirror COVID-19 reporting practices in the state. OMB guidelines encourage this additional granularity when possible and relevant to the population (3). A separate multiracial indicator was created to identify participants who reported at least 2 racial and ethnic categories per 1997 OMB guidelines (3).

The process used at QMC to identify patient race and ethnicity for the EMR was developed in 2010 as part of a statewide collaborative among all acute care hospitals in Hawaii. The process was based on a framework developed by the Health Research & Educational Trust and implemented statewide via standardized tools and training (23). All patients were asked to identify the one race or ethnicity that they identified with the most. For patients who selected multiple races and ethnicities, hospital staff members were instructed to follow a hierarchical algorithm used by the Hawaii Department of Health (23). By hospital policy, once a patient identified a race and ethnicity, the hospital did not ask the question again during future visits; thus, this framework applies only to first-time entries since 2010 (23).

Survey responses on race and ethnicity for each patient were then compared with the race and ethnicity noted in their EMR (Epic Systems Corporation), the gold-standard for self-identification. Patients provided their hospital medical record number in their survey response, which study team members used to link surveys to the EMR. We defined accuracy as the sensitivity of the EMR in predicting a patient's self-identification. EMR–based data were considered accurate if they matched the self-identification and inaccurate if they did not. We calculated accuracy as the total number of hospital EMR entries that matched self-identification divided by the total number of surveys. We also calculated positive predictive values for each self-identification and determined significance for both measures via analysis of variance followed by pairwise Welch *t* tests. Reasons for lack of agreement were grouped into 3 categories: 1) the race and ethnicity listed in the EMR differed from self-identification but was listed in the patient's original list of self-reported races and ethnicities, 2) the race and ethnicity listed in the EMR was not included in the patient's original list of self-reported races and ethnicities, and 3) no entry for race was found in the EMR. When we found no entry for race (n = 6), we categorized patients as belonging in the "Other" group and set their default status as "inaccurate EMR entry." We generated a confusion matrix (a table used to define the performance of a classification algorithm) to explore patterns in disagreement by race and ethnicity.

We conducted additional analyses to compare accuracy among subpopulations, such as multiracial versus single-racial, patients with different self-identifications, and patients in different cohorts. Subsequent analyses consisted of χ^2 tests, paired Welch *t* tests, or analysis of variance followed by paired Welch *t* tests, as appropriate.

Impact on statewide COVID-19 racial and ethnic health disparities

Our data on the number of COVID-19–related hospitalizations were publicly available from the Hawaii Department of Health, current as of January 12, 2022, when the state stopped publicly reporting the number of hospitalizations by race and ethnicity. During the COVID-19 pandemic, hospitals and laboratories were required to submit detailed data on hospitalizations, including data on race and ethnicity, age, and other demographic characteristics of patients. Thus, statewide, publicly reported data on race and ethnicity data were derived directly from hospital EMR data. All data are available at https://health.hawaii.gov/coronavirusdisease2019/current-situation-in-hawaii (24). QMC accounted for 45% of all COVID-19 hospitalizations in the state, with patient demographics closely matching the characteristics of all COVID-19–hospitalized patients statewide.

We explored projected COVID-19 hospitalization rates by self-identification vs EMR-based data by using a simulation (25) in 3 steps. In the first step, we created a pseudo-population matrix based on publicly available COVID-19 hospitalization data from the Hawaii Department of Health. This pseudo-population matrix had a row for each hospitalization and a column containing a hospital-reported race and ethnicity in proportion to the state's racial breakdown. Second, each entry (a patient's EMR-based race and ethnicity) was then randomly assigned a projected self-identified race and ethnicity, with probabilities based on the race and ethnicity confusion matrix derived from the QMC surveys. This value was added as a second column. In the third step, we tallied total self-identified race and ethnicity estimates for the projected population. Steps 2 and 3 were repeated 1,000 times to generate a distribution, from which a median and 95% CIs for each self-identified race and ethnicity were derived. An initial χ^2 test was conducted to determine whether the median projected distribution of self-identification differed significantly from the state's estimates based on hospital EMR data. Projected self-identified population proportions by race and ethnicity were considered significantly different from the proportions among the hospital-derived race and ethnicity if the hospital-derived number fell below the 2.5th or above the 97.5th percentile (ie, outside a 95% CI); *P* values were determined by assuming simulation results were normally distributed. We created a density plot to compare the distribution of self-identified race and ethnicity with the distribution of EMR–based race and ethnicity for each racial and ethnic group.

All analyses were conducted by using R statistical software for Mac, version 4.0.5 (R Foundation for Statistical Computing). We also used tidyverse and Mosaic for data manipulation and general utility (26,27).

Results

Accuracy of EMR-based race and ethnicity vs self-identification

A total of 847 surveys were obtained from QMC. Participants were evenly distributed among the 5 cohorts, with each cohort consisting of more than 100 responses.

Our study population was majority–minority, with no single self-identified race and ethnicity reported by more than 50% of the survey participants. The largest self-identified groups reported by survey were Native Hawaiian (21.7%) and Pacific Islander (18.1%), and Asian (33.3%), followed by non-Hispanic White (21.2%), Hispanic (2.7%), non-Hispanic Black (1.9%), and Other (1.1%). This distribution closely matched the distribution of EMR–based data, which was the following: Native Hawaiian (18.8%) and Pacific Islander (17.2%), and Asian (33.0%), followed by non-Hispanic White (22.5%), Hispanic (2.5%), non-Hispanic Black (2.2%), and unknown or missing (3.8%). Our survey sample was similar to the statewide population of people reported to have been hospitalized with COVID-19, but the sample had more Native Hawaiians and Pacific Islanders relative to the state's general population and fewer Asians and non-Hispanic White people.

Forty-four percent (373 of 847) of survey participants listed more than 1 race and ethnicity. Among these multiracial participants, the average number of races and ethnicities reported was 3.1, with a maximum of 10. Participants who self-identified as Native Hawaiian were more likely than all other groups to be multiracial (92.4%), while Pacific Islanders were the least likely to be multiracial (15.0%). Approximately one-quarter (26.3%) of self-identified Asian participants, 43.6% of self-identified non-Hispanic White participants (Table 1), 47.8% of self-identified Hispanic participants, and 43.8% of self-identified non-Hispanic Black participants identified as multiracial. Self-identified Native Hawaiians who were multiracial also listed significantly more races and ethnicities than other groups (Native Hawaiian, 3.3; non–Native Hawaiian, 2.8; $t = -3.39$; $P < .001$).

The overall agreement between self-identified race and ethnicity and EMR-based race and ethnicity was 86.5% (733 accurate, 114 inaccurate). Of the nonagreements, 43 (37.7%) of the EMR-based races and ethnicities matched their self-reported options, but it was not the race and ethnicity the patient identified with the most; 65 (57.0%) were complete mismatches, where the EMR-based race and ethnicity were not listed by the patient at all; and 6 (5.3%) were the result of the EMR lacking any entry for race (Table 2).

The accuracy of EMR-based race and ethnicity was significantly lower for Native Hawaiian patients (79.2%) than for Asian (92.2%), non-Hispanic White (90.5%), and Pacific Islander patients (90.8%) (Figure 1). (Non-Hispanic Black and Hispanic patient populations were considered too small for reliable and comparable analysis.) This disparity was driven largely by Native Hawaiians being more likely to be multiracial, as patients who were multiracial were significantly less likely to be categorized correctly in the EMR (78.0% vs 94.5%; $P < .001$). When we restricted our analysis to multiracial patients only, we found no significant differences ($P = .32$) in accuracy among Native Hawaiian (78.5%), Asian (78.4%), non-Hispanic White (80.8%), and Pacific Islander patients (82.6%) (Figure 1). We found no significant difference in accuracy among the 162 patients who reported 2 races and/or ethnicities and the 210 patients who reported more than 2 (80.9% vs 75.7%, $P = .23$).

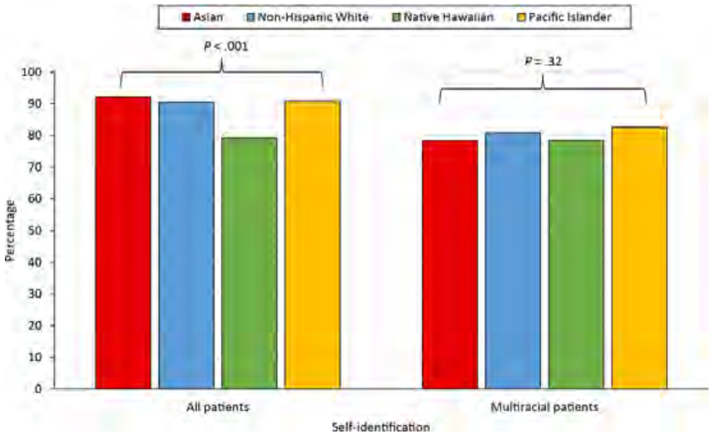


Figure 1. Overall accuracy of race and ethnicity in the electronic medical records of patients in a hospital in Honolulu, Hawaii. Overall accuracy was defined as the total number of hospital electronic medical record entries that matched the self-identified description divided by the total number of surveys.

Accuracy varied by year with no apparent trend (2007, 82.4%; 2008, 87.5%; 2010, 92.1%; 2013, 90.2%; 2020, 85.6%). We found no differences in accuracy by patient age ($t = 1.59$; $P = .11$).

Impact on statewide COVID-19 racial and ethnic health disparities

As of January 12, 2022, Hawaii had 4,041 COVID-19–related hospitalizations. Asian patients accounted for the largest percentage of COVID-19–related hospitalizations (37.8%), followed by Native Hawaiian (22.8%), Pacific Islander (17.8%), non-Hispanic White (15.4%), and patients of other races and ethnicities (6.2%). Native Hawaiians and Pacific Islanders were overrepresented relative to their share of the state's population, whereas Asian and non-Hispanic White patients were underrepresented. However, the racial and ethnic distribution of COVID-19–related hospitalizations in Hawaii was similar to the distribution in our EMR data, indicating that the demographic characteristics of patients with COVID-19–related hospitalizations were similar to the demographic characteristics of patients who were hospitalized before COVID-19.

The overall differences between the adjusted and original COVID-19–related hospitalizations were significant ($\chi^2 = 22.0$, $P < .001$). The simulated distributions showed that projected COVID-19–related hospitalizations among self-identified Native Hawaiians and Pacific Islanders were significantly higher than state estimates, whereas projected COVID-19–related hospitalizations among the “other” races were significantly lower (Figure 2). The median number of COVID-19–related hospitalizations among Native Hawaiian patients was 8.7% higher (1,003 vs 923 hospitalizations) when self-identification rather than EMR-based data were used, and the overall increase in population share was 2.0 percentage points (from 22.8% of the population to 24.8%). The number of COVID-19–related hospitalizations also was higher when self-identification was used among Pacific Islander patients in both total numbers (+3.8%, from 728 hospitalizations to 756) and population share (+0.7 percentage points, from 18.0% to 18.7%); we found lower median numbers of COVID-19–related hospitalizations among all other races and ethnicities when we used self-identification rather than EMR-based data (Figure 2). While most of the newly identified Pacific Islander patients were reclassified from the EMR-based “other” race category, patients newly identified as Native Hawaiian came from many different EMR-based categories (Table 2).

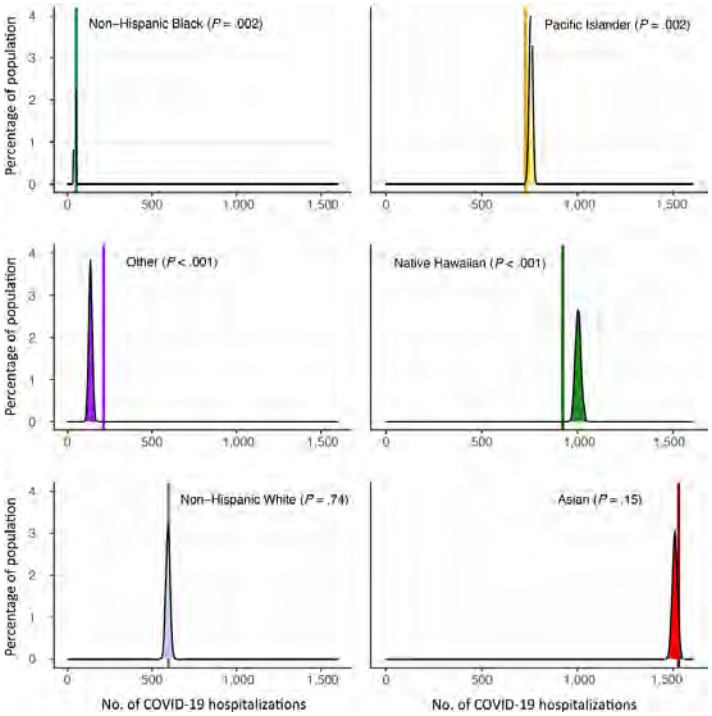


Figure 2. Results for a simulation of COVID-19–related hospitalizations that compared the distribution of adjusted self-identified race and ethnicity (simulated distribution) with the distribution of state-reported race and ethnicity (solid vertical lines). “Other” refers to any patient whose self-reported race did not match predefined categories (eg, “metropolitan,” “mixed,” blank response).

Discussion

The accuracy of race and ethnicity in the EMR system of our study hospital, which has a diverse and multiracial population, was similar to the accuracy of the gold-standard of self-identification (86.5% accuracy for both). Yet even the slight disagreement in categorization was enough to affect health disparities in COVID-19–related hospitalization rates in Hawaii: the median number of COVID-19–related hospitalizations was 8.7% higher among Native Hawaiians and 3.8% higher among Pacific Islanders when we used self-identified data on race and ethnicity instead of EMR-based data.

The accuracy of our EMR race and ethnicity data is similar to the accuracy found in other reports in the literature (88%) (6–17), although our analysis included a diverse, majority–minority population and data that were collected over a longer period. The accuracy of racial and ethnic classification in our hospital’s EMR system for non-White populations (86%) was greater than the accuracy reported in the literature (66% for non-White) (6–17).

Several factors may explain the discordance between our survey results and hospital EMR-based racial and ethnic categories. First, despite standardized hospital procedures to identify race and ethnicity, inconsistencies in data collection may exist. These inconsistencies may apply particularly to patients who are critically ill, cannot speak English, or have difficulty communicating (6–17). Second, the assumption that the race with which a person most identifies is fixed over time, and thus does not need to be confirmed at subsequent visits, may not be appropriate. However, this discrepancy does not account for the 5.1% (43 of 847) of patients who indicated an EMR-based race or ethnicity that was not among any of the self-identified races or ethnicities reported during the survey. Third, it may not be reasonable to expect a single race and ethnicity to fully describe a person’s identity, especially a person who is multiracial. The US Department of Health and Human Services acknowledges that the most recent data standards published in 2011 might not work in “other contexts,” such as administrative records that allow for a single entry only (4). This shortcoming is especially important for groups such as Native Hawaiians, who have access to resources (such as special programs and funding) devoted to any person who has Native Hawaiian lineage. Thus, estimates involving Native Hawaiian people often intend to capture a broader audience than estimates comprising people who identify as Native Hawaiian only. Other strategies, such as allowing for a separate indicator variable for populations of interest, may be necessary to ensure a complete census and appropriate allocation of resources.

Our results address a major gap in the literature by determining the accuracy of EMR-based data on race and ethnicity in a highly diverse population, including people who are most likely to experience health inequities, and they demonstrate the potential impact of misclassification of race and ethnicity in health research. Our findings have broad implications for public health. First, with 86.5% congruence between EMR-based and self-identified race and ethnicity, our findings suggest that EMR-based data are generally accurate. However, the lower accuracy among multiracial patients than among nonmultiracial patients highlights the need to reinforce or modify the standardized approach to collection of data on race and ethnicity. Second, the number of people who self-identify as multiracial is rapidly growing, and this population is becoming increasingly diverse. Although our results suggest that the accuracy of EMR-based data may be lower among multiracial populations than among nonmultiracial populations, it is not clear what the correct approach should be for collecting data on multiracial populations. What is the preferred approach if multiple categories of race and ethnicity are allowed in the collection of data for public health purposes? What if a multiracial person does not want to choose a single race or ethnicity? We emphasize that accuracy is diminished not because people are multiracial but because systems are not set up to capture data on race and ethnicity for this population.

Limitations

Our study is subject to several limitations. First, our projections were based on self-identification of race and ethnicity in a sample of patients seeking care in a single hospital with a strong commitment to the health of Native Hawaiians and Pacific Islanders, and our results may be less generalizable to other hospitals within and outside Hawaii. However, QMC is the largest hospital in the state and accounted for 45% of all COVID-19–related hospitalizations. Moreover, the demographic characteristics of COVID-19 patients at QMC were similar to the demographics of the population in Hawaii. Second, we aggregated race and ethnicity categories to broader categories to explore the impact of our findings on state-reported data. This process resulted in the loss of specificity and, thus, may weaken our claim that our reference group was the gold standard. Third, it is possible that our projections would be less accurate for other COVID-19–related indicators, such as COVID-19 vaccinations, cases, and deaths, where race- and ethnicity-stratified data may be collected in a variety of ways rather than solely through the hospital's EMR system. Fourth, there may be confounders in our self-identification projections. Multiracial people are more likely to be younger than 18 years, but younger people are also less likely to be hospitalized with COVID-19. We demonstrated that age did not affect the misclassification rate. However, our data set did not permit us to assess the confounding potential of other risk factors of COVID-19 hospitalizations such as comorbidities, occupation, or likelihood of vaccination. Individual-level COVID-19 hospitalization data with more covariates may have led to more accurate projections. Finally, all single-category racial and ethnic classification systems assume that racial identity is a static and singular entity and make no distinctions among the wide range of experiences, backgrounds, and needs of individuals.

Conclusion

In a multicultural, majority–minority population of patients in a hospital in Hawaii, the accuracy of race and ethnicity in the hospital EMR system was 86.5% when compared to the gold-standard of self-reported race and ethnicity. Multiracial patients were significantly more likely than nonmultiracial patients to be miscategorized. When we projected this misclassification onto state-level COVID-19 hospitalization data in Hawaii, we found larger health disparities by race and ethnicity among Native Hawaiians and Pacific Islanders. Thus, race and ethnicity misclassification in hospital EMR records may mask the true burden of disease among Native Hawaiians and Pacific Islanders. Further research is needed to determine whether these findings are generalizable to other racial and ethnic groups in other geographical areas.

Acknowledgments

The authors received no external financial support for the research, authorship, and/or publication of this article. The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. No copyrighted materials were used in this article.

Author Information

Corresponding Author: Brendan K. Seto, BA, John A. Burns School of Medicine, University of Hawaii, 651 Ilaalo St, Honolulu, HI 96813. Telephone: 808-469-6767. Email: bkseto@hawaii.edu.

Author Affiliations: [1]John A. Burns School of Medicine, University of Hawaii, Honolulu, Hawaii. [2]The Queen's Medical Center, Honolulu, Hawaii.

References

1. Lavizzo-Mourey RJ, Besser RE, Williams DR. Understanding and mitigating health inequities — past, current, and future directions. N Engl J Med 2021;384(18):1681–4.

2. Office of Management and Budget. OMB directive 15: race and ethnic standards for federal statistics and administrative reporting. 1977. Accessed March 27, 2022. https://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html
3. Office of Management and Budget. Revisions to the standards for the classification of federal data on race and ethnicity. 1997. Accessed March 27, 2022. https://obamawhitehouse.archives.gov/omb/fedreg_1997standards
4. US Department of Health and Human Resources. HHS implementation guidance on data collection standards for race, ethnicity, sex, primary language, and disability status. 2011. Accessed March 27, 2022. https://aspe.hhs.gov/reports/hhs-implementation-guidance-data-collection-standards-race-ethnicity-sex-primary-language-disability-0
5. Centers for Medicare & Medicaid Services. Inventory of resources for standardized demographic and language data collection. 2021. Accessed March 27, 2022. https://www.cms.gov/About-CMS/Agency-Information/OMH/Downloads/Data-Collection-Resources.pdf
6. Saunders CL, Abel GA, El Turabi A, Ahmed F, Lyratzopoulos G. Accuracy of routinely recorded ethnic group information compared with self-reported ethnicity: evidence from the English Cancer Patient Experience survey. BMJ Open 2013;3(6):e002882.
7. West CN, Geiger AM, Greene SM, Harris EL, Liu IL, Barton MB, et al. Race and ethnicity: comparing medical records to self-reports. J Natl Cancer Inst Monogr 2005;2005(35):72–4.
8. Klinger EV, Carlini SV, Gonzalez I, Hubert SS, Linder JA, Rigotti NA, et al. Accuracy of race, ethnicity, and language preference in an electronic health record. J Gen Intern Med 2015;30(6):719–23.
9. Hamilton NS, Edelman D, Weinberger M, Jackson GL. Concordance between self-reported race/ethnicity and that recorded in a Veteran Affairs electronic medical record. N C Med J 2009;70(4):296–300.
10. Pan CX, Glynn RJ, Mogun H, Choodnovskiy I, Avorn J. Definition of race and ethnicity in older people in Medicare and Medicaid. J Am Geriatr Soc 1999;47(6):730–3.
11. Gomez SL, Kelsey JL, Glaser SL, Lee MM, Sidney S. Inconsistencies between self-reported ethnicity and ethnicity recorded in a health maintenance organization. Ann Epidemiol 2005;15(1):71–9.
12. Gomez SL, Glaser SL. Misclassification of race/ethnicity in a population-based cancer registry (United States). Cancer Causes Control 2006;17(6):771–81.
13. Kressin NR, Chang BH, Hendricks A, Kazis LE. Agreement between administrative data and patients' self-reports of race/ethnicity. Am J Public Health 2003;93(10):1734–9.
14. Boehmer U, Kressin NR, Berlowitz DR, Christiansen CL, Kazis LE, Jones JA. Self-reported vs administrative race/ethnicity data and study results. Am J Public Health 2002;92(9):1471–2.
15. Blustein J. The reliability of racial classifications in hospital discharge abstract data. Am J Public Health 1994;84(6):1018–21.
16. Stewart SL, Swallen KC, Glaser SL, Horn-Ross PL, West DW. Comparison of methods for classifying Hispanic ethnicity in a population-based cancer registry. Am J Epidemiol 1999;149(11):1063–71.
17. Wey A, Davis J, Juarez DT, Sentell T. Distinguishing between primary and secondary racial identification in analyses of health disparities of a multiracial population in Hawaii. Ethn Health 2018;23(3):233–48.
18. US Census Bureau. Hawaii 2020 census. 2020. Accessed March 27, 2022. https://www.census.gov/library/stories/state-by-state/hawaii-population-change-between-census-decade.html
19. Tai DBG, Shah A, Doubeni CA, Sia IG, Wieland ML. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. Clin Infect Dis 2021;72(4):703–6.
20. Kaholokula JK, Samoa RA, Miyamoto RES, Palafox N, Daniels SA. COVID-19 special column: COVID-19 hits Native Hawaiian and Pacific Islander communities the hardest. Hawaii J Health Soc Welf 2020;79(5):144–6.
21. Penaia CS, Morey BN, Thomas KB, Chang RC, Tran VD, Pierson N, et al. Disparities in Native Hawaiian and Pacific Islander COVID-19 mortality: a community-driven data response. Am J Public Health 2021;111(S2):S49–52.
22. Quint JJ, Van Dyke ME, Maeda H, Worthington JK, Dela Cruz MR, Kaholokula JK, et al. Disaggregating data to measure racial disparities in COVID-19 outcomes and guide community response — Hawaii, March 1, 2020–February 28, 2021. MMWR Morb Mortal Wkly Rep 2021;70(37):1267–73.
23. Pellegrin KL, Miyamura JB, Ma C, Taniguchi R. Improving accuracy and relevance of race/ethnicity data: results of a statewide collaboration in Hawaii. J Healthc Qual 2016;38(5):314–21.
24. Quint J. Hawaii COVID-19 data. Accessed March 25, 2021. https://health.hawaii.gov/coronavirusdisease2019/what-you-should-know/current-situation-in-hawaii/
25. Heath A, Strong M, Glynn D, Kunst N, Welton NJ, Goldhaber-Fiebert JD. Simulating study data to support expected value of sample information calculations: a tutorial. Med Decis Making 2022;42(2):143–55.

26. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the tidyverse. J Open Source Softw 2019;4(43):1686.
27. Pruim R, Kaplan DT, Horton NJ. The mosaic package: helping students to "think with data" using R. R J 2017;9(1):77–102.

Tables

Table 1. Demographic Characteristics of Patient Respondents to a Hospital Race Validation Survey,[a] Patients in a Hospital Electronic Medical Record System,[b] and Patients Included in State-Reported COVID-19–Related Hospitalizations[c]

Item	Asian	Native Hawaiian	Pacific Islander	Non-Hispanic White
Patient survey population				
Overall, %	33.3	21.7	18.1	21.2
Multiracial, %	26.3	92.4	15.0	43.6
No. of races listed, mean (SD)	2.8 (0.9)	3.3 (1.5)	3.0 (1.1)	2.5 (1.2)
Age, mean (SD), y	53.2 (19.3)	50.5 (13.5)	49.1 (17.9)	56.1 (16.3)
Hospital electronic medical records,[b] %	33.0	18.8	17.2	22.5
State-reported COVID-19–related hospitalizations,[c] %	37.8	22.8	17.8	15.4

[a] Survey data were collected in 5 cohorts as part of an ongoing quality assurance project conducted by The Queen's Medical Center, Honolulu, Hawaii, during 5 years from 2007 through 2020 (2007, 2008, 2010, 2013, 2020).
[b] Extracted from The Queen's Medical Center's electronic medical records.
[c] Extracted from publicly available data on 4,041 COVID-19–related hospitalizations from the Hawaii Department of Health, current as of January 12, 2022 (24).

Table 2. Confusion Matrix Showing Agreement Between Hospital Electronic Medical Records and Self-Identified Race and Ethnicity Among 847 Patients in a Hospital in Honolulu, Hawaii[a]

Self-identified race and ethnicity	Hospital EMR-based race and ethnicity						
	Asian	Non-Hispanic Black	Non-Hispanic White	Native Hawaiian	Pacific Islander	Other	Positive predictive value
Asian	259	0	8	5	0	9	0.92
Non-Hispanic Black	0	16	0	0	0	0	1.00
Non-Hispanic White	3	0	162	6	1	7	0.91
Native Hawaiian	13	2	12	145	3	8	0.79
Pacific Islander	2	0	1	0	139	11	0.91
Other	2	1	7	3	2	18	0.55
Sensitivity[b]	.93	.84	.85	.91	.96	.34	—

[a] Survey data were collected in 5 cohorts as part of an ongoing quality assurance project conducted by The Queen’s Medical Center, Honolulu, Hawaii, during 5 years from 2007 through 2020 (2007, 2008, 2010, 2013, 2020).
[b] Overall accuracy of data (sensitivity) = 0.88.

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
Volume 19, E51 AUGUST 2022

ORIGINAL RESEARCH

Food Insecurity in a Sample of Informal Caregivers in 4 Southern US States

Swarnali Goswami, MPharm[1]; Siddhi Korgaonkar, MS[1]; Kaustuv Bhattacharya, PhD[1]; Meagen Rosenthal, PhD[1]

Accessible Version: www.cdc.gov/pcd/issues/2022/22_0069.htm

Suggested citation for this article: Goswami S, Korgaonkar S, Bhattacharya K, Rosenthal M. Food Insecurity in a Sample of Informal Caregivers in 4 Southern US States. Prev Chronic Dis 2022;19:220069. DOI: https://doi.org/10.5888/pcd19.220069.

PEER REVIEWED

Summary

What is already known on this topic?

Southern US states bear a disproportionate burden of food insecurity and also have a higher prevalence of informal caregivers compared with other US states.

What is added by this report?

The characteristics of caregivers associated with food insecurity have not been examined previously, so we assessed caregiving-related predictors of food insecurity among caregivers in Alabama, Mississippi, Tennessee, and Louisiana.

What are the implications for public health practice?

Screening of caregivers for food insecurity in health care settings and subsequent linkage to appropriate food and caregiving support resources should be a priority of future policies targeting food insecurity.

Abstract

Introduction

Given the disproportionate burden of food insecurity in the southern US states and the high prevalence of caregiving in this area, we assessed caregiving-related predictors of food insecurity among caregivers in 4 southern US states.

Methods

We used data from the 2015 Behavioral Risk Factor Surveillance System (BRFSS) for individuals aged 18 years or older who resided in Alabama, Louisiana, Mississippi, and Tennessee to assess the association between caregiving status and food insecurity, accounting for the complex survey design of BRFSS. Caregiving-related predictors of food insecurity were identified by using multivariable logistic regression.

Results

Weighted counts of caregivers and noncaregivers were 356,198 and 652,737, respectively. Prevalence of food insecurity was higher among caregivers than noncaregivers (35.9% vs 25.9%). Adjusting for sociodemographic predictors, caregivers had 56% (95% CI, 1.30–1.87; $P < .001$) higher odds of food insecurity than noncaregivers. Among caregivers, those caring for a spouse or a partner (adjusted odds ratio [aOR] = 1.7; 95% CI, 1.02–2.85; $P = .04$) had significantly higher odds of food insecurity compared with those caring for parents or parents-in-law. Caregivers who had been caregiving for 6 months to 2 years had higher odds of food insecurity compared with those who had been caregiving for less than 6 months (aOR = 1.88; 95% CI, 1.12–3.16; $P = .02$). Caregivers who reported a need for support services had higher odds of food insecurity compared with those who did not (aOR = 3.38; 95% CI, 2.19–5.21; $P < .001$). Caregivers caring for people with musculoskeletal conditions, compared with people with neurologic conditions, had higher odds of food insecurity (aOR = 3.47; 95% CI, 1.52–7.91; $P = .003$).

Conclusion

Caregiver screening for food insecurity in health care settings and linkage to appropriate food and caregiving support resources should be prioritized by future health policies.

Introduction

Informal caregivers provide unpaid assistance or supervision with personal tasks not including childcare to a relative or friend who cannot perform these tasks because of cognitive, physical, or psychological impairments (1). Southern US states were reported to have the highest prevalence of informal caregivers (≥25%) during 2015 to 2017 (2). Caregiving is demanding and is associated with poor health outcomes such as chronic stress, obesity, diabetes, and mental health problems (3–5). Additionally, age-adjusted rates of informal caregivers reporting fair or poor health in Alabama, Louisiana, Mississippi, and Tennessee during 2015 to 2017 were reported to be 20% or more (2). Caregiving has been recognized as a public health issue, and its burden is likely to worsen with the rapidly growing aging population in the US (6).

In addition to adversely affecting caregiver health, caregiving also creates financial strain on the caregiver, which could affect their ability to afford food. Approximately 20% of a caregiver's income is reportedly spent on caregiving expenses, with household and medical expenses being the biggest drivers of caregiving-related expenses (7). A 2012 study reported that caregivers were twice as likely to report food insecurity compared with noncaregivers (8). Between 2017 and 2019, the household food insecurity rate in southern US states was higher than that of the rest of the country (9).

Despite southern US states bearing a disproportionate burden of food insecurity and caregiving, the characteristics of caregivers associated with food insecurity has not been examined. We assessed the prevalence of food insecurity among adult caregivers and the association of food insecurity and caregiving status in 4 southern US states: Alabama, Louisiana, Mississippi, and Tennessee. We also sought to identify caregiving characteristics associated with food insecurity among caregivers. Our findings will help plan appropriate policies for assisting caregivers most at risk of food insecurity.

Methods

This was a cross-sectional study of data from the 2015 Behavioral Risk Factor Surveillance System (BRFSS) that included adult informal caregivers in Alabama, Louisiana, Mississippi, and Tennessee.

The BRFSS is a collaborative project of the Centers for Disease Control and Prevention (CDC) and the US states and territories. BRFSS data are collected annually from noninstitutionalized US adults by state-based surveillance systems following a population density–based strata sampling design and random-digit–dialing telephone survey. Each respondent is assigned weights calculated through iterative proportional fitting for each stratum to be considered nationally representative. Because BRFSS data are publicly available, this project was deemed exempt from institutional review board review. We used data from 2015, the latest year for which the information on both caregiving and food insecurity was available in the BRFSS data set.

Measures

Adult informal caregivers were identified from the survey item, "People may provide regular care or assistance to a friend or family member who has a health problem or disability. During the past 30 days, did you provide regular care or assistance to a friend or family member who has a health problem or disability?" Those who replied yes or no were included in the analysis, and their caregiver status was designated as such. Those who refused to answer the caregiving question were excluded from analysis.

Food insecurity, which is the lack of reliable access to affordable and nutritious food, was assessed via the item, "How often in the past 12 months would you say you were worried or stressed about having enough money to buy nutritious meals?" Those who responded "rarely" or "never" were considered food secure, and those who responded "always," "usually," or "sometimes" were considered food insecure. The food insecurity variable was dichotomized following the methodology of previous studies (10). Respondents who had missing values for the food insecurity question were excluded from the analysis.

Caregiving characteristics considered for examining predictors of food insecurity among adult informal caregivers were relationship with caregiver (parent or parent-in-law, child or grandchild, spouse or partner, other), care recipient condition (mental or neurologic, metabolic or cardiovascular, musculoskeletal, cancer, or other), caregiving 40 hours or more per week (yes or no), months spent caregiving (less than 6 months, 6 months to up to 2 years, 2 years or more), need for support services (yes or no), helping with activities of daily living (ADLs) and instrumental activities of daily living (IADLs) (yes or no). ADLs are personal activities that a care recipient might need an informal caregiver's help with and were identified from the BRFSS item, "In the past 30 days, did you provide care for this person by managing personal care such as giving medications, feeding, dressing, or bathing?" IADLs are activities that are broader in scope, requiring coordination and planning, with which the care recipient might need an informal caregiver's help (11), and were identified from the BRFSS item, "In the past 30 days, did you provide care for this person by managing household tasks such as cleaning, managing money, or preparing meals?" Informal caregivers often benefit from support services such as classes about caregiving activities (eg, giving medications, help in getting access to services, support groups, individual counseling to help cope with giving care, respite care) that enable them to take better care of their care recipients. Therefore, need for support services was assessed from the BRFSS item, "Of the following support services, which one do you MOST need, that you are not currently getting?" Responses were dichotomized into needed any type of support services or did not need any.

Sociodemographic variables included age, sex, race and ethnicity, education, marital status, employment status, annual household income, health insurance, and Metropolitan Statistical Area (MSA) indicator. We controlled for MSA in our analysis because it had been reported that the prevalence of food insecurity in urban areas is higher than in suburban or rural areas (9).

Statistical analysis

We described the overall and state-based prevalence estimates of food insecurity among caregivers and noncaregivers using weighted percentages, and we described the sociodemographic characteristics as proportions, by caregiving status. We also reported the caregiving-related characteristics of the caregivers in the sample and the prevalence of food insecurity among them. The association between caregiving status and food insecurity was determined by using a multivariable logistic regression model, adjusted for sociodemographic characteristics (Figure 1). Sociodemographic and caregiving-related correlates (relationship with caregiver, care recipient condition, hours per week, months since caregiving began, need for support services, assisting with ADLs, and assisting with IADLs) of food insecurity were assessed by using multivariable logistic regression (Figure 2). We reported odds ratios, 95% CIs, and the associated *P* values. All analyses accounted for the complex sampling design of the BRFSS, and appropriate subsample procedures and survey weights were used. Analysis was conducted by using SAS version 9.4 (SAS Institute, Inc), and complete case analysis was done for all analyses. We set significance at $P < .05$.

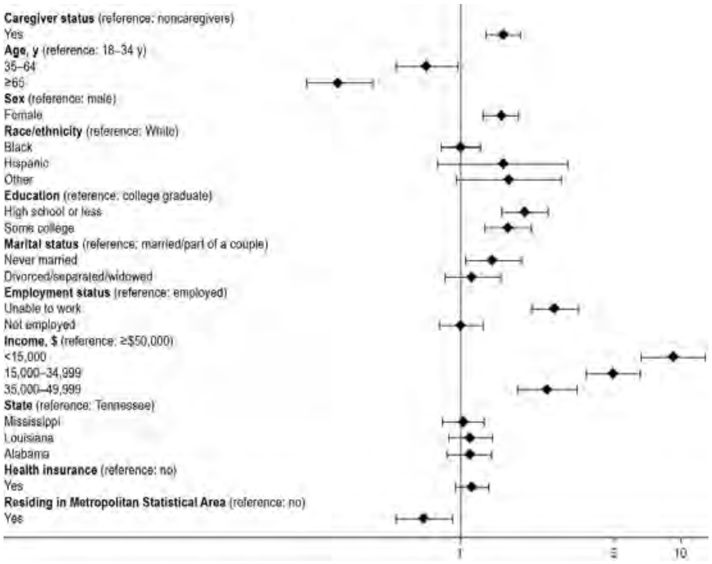


Figure 1. Association of caregiver status with food insecurity, adjusting for sociodemographic covariates, Behavioral Risk Factor Surveillance System, 2015.

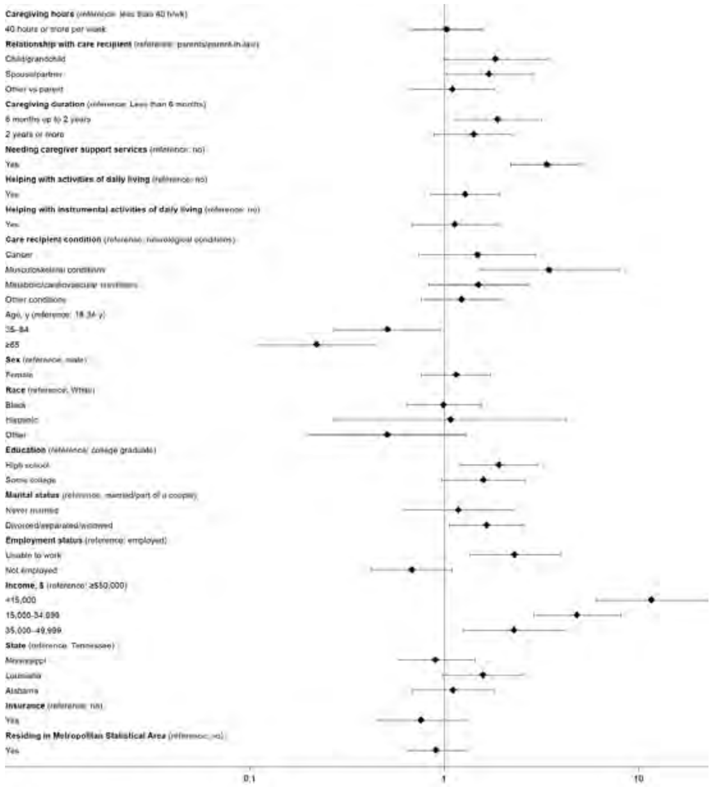


Figure 2. Association of caregiver characteristics with food insecurity, adjusting for sociodemographic covariates, Behavioral Risk Factor Surveillance System, 2015.

Results

Overall, 35.9% (95% CI, 33.9%–37.9%) of caregivers and 25.9% (95% CI, 24.8%–26.9%) of noncaregivers in Alabama, Louisiana, Mississippi, and Tennessee reported food insecurity in 2015 (Table 1). In these 4 states, the prevalence of food insecurity among caregivers was highest in Louisiana (38.2%; 95% CI, 34.2%–42.3%). Caregivers aged 18 to 34 years and 35 to 64 years had a higher proportion of food insecurity than their noncaregiving counterparts (18.1% vs 13.2% and 66.6% vs 60.3%, respectively), and they also had a higher proportion of food insecurity than caregivers aged 65 years or older (15.3%) (Table 2). Most caregivers who experienced food insecurity were aged 35 to 64 years (66.6%; 95% CI, 61.0%–72.1%). Among both caregivers and noncaregivers, most food-insecure individuals were White and female. Most of the food-insecure adults in the caregiver sample were either married or part of a couple (46.1%; 95% CI, 40.5%–51.6%) followed by those who were divorced, separated, or widowed (31.8%; 95% CI, 26.8%–36.8%). Unemployed adults reported a higher prevalence of food insecurity (38.0%; 95% CI, 32.4%–43.6%) than employed adults (36.1%; 95% CI, 30.7–41.5). Among the food-insecure caregivers in the sample, the highest proportion had an annual household income between $15,000 and $35,000 (46.7%; 95% CI, 41.2%–52.3%), had health insurance (74.9%; 95% CI, 69.5%–80.4%), and lived in an MSA (67.9%; 95% CI, 63.1%–72.7%).

Most caregivers were caring for their parents or parents-in-law (42.9%; 95% CI, 39.5%–46.3%), were caring for less than 40 hours each week (73.6%; 95% CI, 70.6%–76.7%), and had been providing care for 2 years or more (56.4%; 95% CI, 53.1%–59.7%) (Table 3). In terms of caregiving conditions, 18.9% (95% CI, 16.4%–21.4%) were caring for people with mental or neurologic conditions, 13.9% (95% CI, 11.8%–16.0%) for metabolic or cardiovascular conditions, 6.8% (95% CI, 4.8%–8.8%) for musculoskeletal conditions, and 7.8% (95% CI, 5.8%–9.7%) for cancer. Approximately two-thirds of caregivers (62.0%; 95% CI, 58.8%–65.2%) helped their care recipient with ADLs and 84.5% (95% CI, 82.3%–86.8%) with IADLs. Approximately one-fifth of caregivers (18.7%; 95% CI, 16.1%–21.3%) expressed a need for caregiver support services (Table 3). Prevalence of food insecurity was highest among caregivers who were caring for children or grandchildren who had a health condition (37.2%; 95% CI, 27.7%–46.7%) and those with musculoskeletal conditions (48.3%; 95% CI, 32.6%–63.9%). Food insecurity was also highest among those who were caregiving for 40 hours per week or more (39.2%; 95% CI, 32.3%–46.1%), had been caregiving for 6 months up to 2 years (34.0%; 95% CI, 25.5%–42.6%), expressed a need for caregiver support services (49.5%; 95% CI, 41.8%–57.3%), and were helping with ADLs (34.6%; 95% CI, 30.2%–39.1%) or IADLs (32.7%; 95% CI, 28.9%–36.3%).

After accounting for age, sex, race and ethnicity, state, income, education level, marital status, insurance status, and MSA, caregivers had higher odds of reporting food insecurity (adjusted odds ratio [aOR] = 1.56; 95% CI, 1.30–1.87; $P < .01$) than noncaregivers (Figure 1). Caregivers who cared for a spouse or partner (aOR = 1.7; 95% CI, 1.02–2.85; $P = .04$) had significantly higher odds of food insecurity compared with those who cared for parents or parents-in-law (Figure 2). In terms of caregiving duration, compared with caregivers who had been caregiving for less than 6 months, those caregiving for 6 months to up to 2 years had significantly higher odds of food insecurity (aOR = 1.88; 95% CI, 1.12–3.16; $P = .02$) (Figure 2). Compared with those who did not express a need for caregiver support services, caregivers who expressed a need for such support services had higher odds of food insecurity (aOR = 3.38; 95% CI, 2.19–5.21; $P < .001$). Finally, in terms of caregiving conditions, compared with caregiving for neurologic conditions, caregivers caring for people with musculoskeletal conditions, especially arthritis (aOR = 3.47; 95% CI, 1.52–7.91; $P = .003$), had higher odds of food insecurity.

Discussion

In 2015, food insecurity was higher among caregivers, both overall and in the 4 southern US states we assessed — Alabama, Mississippi, Louisiana, and Tennessee — with the highest prevalence reported in Louisiana. Furthermore, caregivers in these states had higher odds of food insecurity even after accounting for sociodemographic characteristics. We found that caregivers who care for their spouses or partners and those who care for their children or grandchildren had higher odds of food insecurity than those who cared for their parents or parents-in-law. This finding could be due to care recipients such as spouses, partners, and children sharing the same household, which could result in increased health care spending, increased financial strain, and resultant decrease in resources to afford nutritious food (7). Food insecurity issues among caregivers of children with certain health conditions has been well documented (12,13). Although literature on food insecurity among those who provide care for spouses and partners is scarce, evidence exists of a substantial caregiving burden among caregivers for spouses with chronic or terminal diseases and its association with health conditions such as depression and anxiety (14,15). Therefore, caring for spouses or partners could result in worse physical and mental health of the caregivers, which in turn could increase health care spending and predispose them to food insecurity. Thus, screening for food insecurity should be made available for those caring for young children and for spouses or partners. Future food insecurity interventions should also prioritize such caregivers and their households.

According to our study, people caregiving for 6 months to less than 2 years were more likely to experience food insecurity than those caregiving for less than 6 months. This finding indicates that food-related stress may be more intense from 6 months to less than 2 years of the caregiving, a time that the caregivers would most require food-related support. Thus, new caregivers should be screened in the health care setting using validated food insecurity questionnaires and connected to appropriate food access programs such as the Supplemental Nutrition Assistance Program (SNAP), the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), and the Child and Adult Care Food Program (CACFP), and others (16). Validated food insecurity questionnaires have been the most widely implemented and evaluated method of screening in health care settings and have been reported to be effective (16). These resources will improve newer caregivers' awareness of food access programs and make them better equipped to manage their food-related needs as they progress toward the more intense caregiving periods (17). Moreover, our study indicated that those who expressed a need for caregiver support services had a higher likelihood of reporting food insecurity. Informal caregivers tend to support their care recipients in managing symptoms, administering medications, changing bandages, and other medical and nursing tasks for which they often do not receive necessary training. Literature suggests that informal caregivers do not receive adequate support in medical care training and access to health care facilities, counseling, and support groups to cope with caregiving stress as well as respite care services (18,19), emphasizing the need for interventions that include provisions for connecting caregivers to appropriate channels where they can access such services (20). Addressing the unmet needs of these caregivers can help alleviate their financial strain, reduce caregiving-related mental and physical burden, and ultimately improve their food security. As of 2022, several federal and state-based health insurance programs such as Medicare Advantage and Medicaid cover a variety of in-home care services and nonskilled needs, such as help with daily activities, to promote aging in place (21,22). Alabama, Louisiana, Mississippi, and Tennessee have Medicaid programs that aid informal caregivers. In Alabama, Medicaid programs such as the Elderly and Disabled Waiver, the State of Alabama Independent Living (SAIL) Waiver, and the Personal Choices Program and Alabama Community Transition (ACT) Medicaid Waiver aid with home care. Louisiana has several Medicaid programs such as Long-term Personal Care Services Waiver, the Adult Day Health Care Waiver, and the Community Choices Waiver, which provide similar assistance. Mississippi Medicaid's Elderly and Disabled Waiver provides a variety of in-home support and care services to individuals, including personal care and adult day care. Tennessee covers home care with the CHOICES in Long-term Care program, which provides benefits such as personal care and homemaker services, assistive technology, personal emergency response systems, and home modifications (23). According to a recent policy analysis, 13% of Medicare Advantage plans have been reported to offer family caregiver supports such as respite care, counseling, and skills training (24).

More than 80% of food-insecure caregivers in the sample were younger than 65 years. Approximately 64% were either unemployed or not able to work, 28% had an annual household income of $15,000 or less, and approximately 25% of informal caregivers were uninsured. Health insurance also plays an important role in food insecurity, as demonstrated by a recent study that showed a positive association between Medicaid expansion and improvement in food insecurity as a spillover effect of reducing poverty (25). Aforementioned sociodemographic characteristics of food-insecure caregivers underscore the importance of Medicaid coverage in this sample. Although the insurance programs discussed provide waivers to support caregivers, lack of health insurance or being ineligible for Medicaid in states that have not yet expanded their Medicaid programs may limit the eligibility of several caregivers to access these programs and, in turn, predispose them to food insecurity. None of the 4 states studied had expanded their Medicaid programs as of 2015. Even to date, Alabama, Mississippi, and Tennessee have not adopted Medicaid expansion. Hence, along with increasing awareness on the availability of facilities for caregivers, future health policies should focus on the development of better strategies for improving access to such services (eg, clinic-to-community models for addressing food insecurity, increased collaborations between health care systems and food assistance providers) (17). Additionally, regional variations among SNAP eligibility requirements should be streamlined to improve access to food services for caregivers who are most in need of those services (17).

Caregiving conditions often dictate the care intensity and involvement of the informal caregiver. Our study highlighted that caring for people with musculoskeletal conditions (eg, arthritis) is associated with food insecurity among caregivers. The impact of helping to manage a care recipient's arthritis condition on a caregiver's health-related quality of life, physical health, and mental health has been noted because of the chronic nature of the disease, which typically requires more than 20 hours of care per week (26). The demanding nature of care, along with high health care costs, could predispose caregivers to food insecurity, which underscores the need for better care coordination for patients and involvement of the caregiver in the care plan so that caregivers are aware of their financial responsibility in the situation. Further research should be conducted to understand the impact of caregiving burden on food insecurity, specifically among those caring for people with arthritis. Moreover, screening for food insecurity among caregivers of those with arthritis and other musculoskeletal conditions should be made a priority in clinical settings.

Strengths and limitations

Our study fills a gap in the literature on food insecurity among caregivers in southern US states, where food insecurity prevalence is higher. Moreover, no previous study has examined the caregiving-related predictors of food insecurity. We used 2015 BRFSS data, and other data sets, such as the Medical Expenditure Panel Survey, have the latest food insecurity data, although their caregiver modules date back to 1998 (27). The American Community Survey also has questions on food insecurity but only asks about grandparents as caregivers and not about family caregiving, thus providing an incomplete picture of caregivers in the US (28). The National Health and Aging Trends Study (NHATS) only includes Medicare beneficiaries aged 65 years or older, and its companion, the National Study of Caregiving, is a survey of the informal caregivers of respondents in NHATS, thus limiting the generalizability of the data to the general US population (29). Noting the limitations of other data sets, we chose the BRFSS 2015 data set because of the availability of data on food insecurity and caregiving.

Our study has limitations. First, food insecurity and caregiving data were only available for the 4 southern US states of Alabama, Louisiana, Mississippi, and Tennessee, so our results may not be generalizable to the US caregiver population. Future studies should explore the caregiving-related predictors of food insecurity in a more generalizable sample. Second, we used data from 2015, and it is possible that caregiving and food insecurity prevalence have changed since then. However, food insecurity in the US has recently only minimally decreased, from 12.7% in 2015 to 10.5% in 2020, indicating that the results of this study are relevant. The COVID-19 pandemic–related unemployment (with the loss of employer-sponsored health insurance for many), income loss, and health care disruptions (increased caregiving burden if the care recipient condition worsened), coupled with the shelter-in-place policies that might have further limited access to affordable nutritious food, could have increased the risk of food insecurity among caregivers. Third, our study used self-reported questions to ascertain respondents' food insecurity status and may be subject to social desirability and recall bias. However, this item has been validated by the US Department of Agriculture as part of their 10-item food security scale and reported to have a reliability of 0.71 ($P < .001$), helping to mitigate the risk of such biases (30). Lastly, because of the cross-sectional nature of the study, we were unable to control for temporality or to make any causal inferences. Future studies should further examine the temporal relationship between food insecurity and caregiving status.

Conclusion

We found that food insecurity was more prevalent among caregivers compared with noncaregivers in the southern US states of Alabama, Louisiana, Mississippi, and Tennessee. These states should be considered a priority group for future food insecurity–related interventions. Our study provides insights for planning future policies focused on alleviating food insecurity among caregivers. Key strategies included timely screening in health care settings using validated food insecurity questionnaires, involvement of caregivers in care planning, helping caregivers access support services, and local food-related resources. Appropriate training, education, and support for caregivers could be incorporated into routine care settings such as physicians' offices, hospitals, and pharmacies. Results from this study can help public health practitioners develop effective policies and direct public funds to alleviate food insecurity among caregivers.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. No copyrighted materials or tools were used in this research.

Author Information

Corresponding Author: Swarnali Goswami, MPharm, Department of Pharmacy Administration, School of Pharmacy, University of Mississippi, Faser Hall 129, University, MS 38677. Telephone: 662-202-4867. Email: sgoswami@go.olemiss.edu.

Author Affiliations: [1]Department of Pharmacy Administration, School of Pharmacy, University of Mississippi, University, Mississippi.

References

1. Fredman L, Cauley JA, Hochberg M, Ensrud KE, Doros G; Study of Osteoporotic Fractures. Mortality associated with caregiving, general stress, and caregiving-related stress in elderly women: results of caregiver-study of osteoporotic fractures. J Am Geriatr Soc 2010;58(5):937–43.
2. Edwards VJ, Bouldin ED, Taylor CA, Olivari BS, McGuire LC. Characteristics and health status of informal unpaid caregivers — 44 states, District of Columbia, and Puerto Rico, 2015–2017. MMWR Morb Mortal Wkly Rep 2020;69(7):183–8.
3. Schulz R, Sherwood PR. Physical and mental health effects of family caregiving. Am J Nurs 2008;108(9Suppl):23–7, quiz 27.
4. Trivedi R, Beaver K, Bouldin ED, Eugenio E, Zeliadt SB, Nelson K, et al. Characteristics and well-being of informal caregivers: results from a nationally-representative US survey. Chronic Illn 2014;10(3):167–79.
5. Giovannetti ER, Wolff JL, Xue QL, Weiss CO, Leff B, Boult C, et al. Difficulty assisting with health care tasks among caregivers of multimorbid older adults. J Gen Intern Med 2012;27(1):37–44.
6. Centers for Disease Control and Prevention. Caregiving for family and friends — a public health issue. Accessed December 20, 2021. https://www.cdc.gov/aging/caregiving/caregiver-brief.html
7. Rainville C, Skufca L, Mehegan L. Family caregiving and out-of-pocket costs: 2016 report: fact sheet. AARP Research; 2016.

8. Horner-Johnson W, Dobbertin K, Kulkarni-Rajasekhara S, Beilstein-Wedel E, Andresen EM. Food insecurity, hunger, and obesity among informal caregivers. Prev Chronic Dis 2015; 12:150129.
9. US Department of Agriculture, Economic Research Service. Food security and nutrition assistance. Accessed June 19, 2022. https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/food-security-and-nutrition-assistance/
10. Pan L, Sherry B, Njai R, Blanck HM. Food insecurity is associated with obesity among US adults in 12 states. J Acad Nutr Diet 2012;112(9):1403–9.
11. Kuluski K, Peckham A, Gill A, Arneja J, Morton-Chang F, Parsons J, et al. "You've got to look after yourself, to be able to look after them": a qualitative study of the unmet needs of caregivers of community based primary health care patients. BMC Geriatr 2018;18(1):275.
12. Burke MP, Martini LH, Çayır E, Hartline-Grafton HL, Meade RL. Severity of household food insecurity is positively associated with mental disorders among children and adolescents in the United States. J Nutr 2016;146(10):2019–26.
13. Lee A, Bocchini C, Lopez M, Clegg K, Greeley C. Food insecurity in the caregivers of hospitalized pediatric patients. Pediatrics 2019;144(2_MeetingAbstract):481.
14. Happe S, Berger K; FAQT Study Investigators. The association between caregiver burden and sleep disturbances in partners of patients with Parkinson's disease. Age Ageing 2002;31(5):349–54.
15. Mishra S, Gulia A, Satapathy S, Gogia A, Sharma A, Bhatnagar S. Caregiver burden and quality of life among family caregivers of cancer patients on chemotherapy: a prospective observational study. Indian J Palliat Care 2021; 27(1):109–12.
16. De Marchis EH, Torres JM, Fichtenberg C, Gottlieb LM. Identifying food insecurity in health care settings: a systematic scoping review of the evidence. Fam Community Health 2019; 42(1):20–9.
17. Barnidge E, Stenmark S, Seligman H. Clinic-to-community models to address food insecurity. JAMA Pediatr 2017; 171(6):507–8.
18. Mollica MA, Smith AW, Kent EE. Caregiving tasks and unmet supportive care needs of family caregivers: a U.S. population-based study. Patient Educ Couns 2020;103(3):626–34.
19. Denham AMJ, Wynne O, Baker AL, Spratt NJ, Turner A, Magin P, et al. An online survey of informal caregivers' unmet needs and associated factors. PLoS One 2020; 15(12):e0243502.
20. Administration for Community Living. Finding local services. Accessed June 19, 2022. https://acl.gov/ltc/basic-needs/finding-local-services
21. American Council on Aging. Medicaid and in-home care: eligibility, benefits and state rules. Accessed June 19, 2022. https://www.medicaidplanningassistance.org/in-home-care/
22. Health Affairs Blog. Shifting the burden? Consequences of postacute care payment reform on informal caregivers. Accessed June 19, 2022. https://www.healthaffairs.org/do/10.1377/hblog20190828.894278/full/
23. Medicaid and home care: state by state benefits and eligibility. Accessed June 19, 2022. https://www.payingforseniorcare.com/medicaid-waivers/home-care
24. AARP. Supplemental benefits in Medicare advantage: what's changing in 2019and what's not. Accessed June 19, 2022. https://blog.aarp.org/thinking-policy/supplemental-benefits-in-medicare-advantage-whats-changing-in-2019-and-whats-not
25. Torres ME, Capistrant BD, Karpman H. The effect of Medicaid expansion on caregiver's quality of life. Soc Work Public Health 2020;35(6):473–82.
26. Brouwer WBF, van Exel NJA, van de Berg B, Dinant HJ, Koopmanschap MA, van den Bos GAM. Burden of caregiving: evidence of objective burden, subjective burden, and quality of life impacts on informal caregivers of patients with rheumatoid arthritis. Arthritis Rheum 2004;51(4):570–7.
27. Agency for Healthcare Research and Quality. Medical Expenditure Panel Survey questionnaire sections. Accessed June 19, 2022. https://meps.ahrq.gov/survey_comp/survey_results_ques_sections.jsp?Section=CG&Year1=AllYear
28. Cohen S. Why are family caregiving questions missing from the 2020census? Forbes. Accessed June 19, 2022. https://www.forbes.com/sites/nextavenue/2019/04/10/why-are-family-caregiving-questions-missing-from-the-2020-census/
29. National Health and Aging Trends Study. National Study of Caregiving (NSOC). Accessed June 19, 2022. https://nhats.org/researcher/nsoc
30. Irving SM, Njai RS, Siegel PZ. Food insecurity and self-reported hypertension among Hispanic, Black, and White adults in 12 states, Behavioral Risk Factor Surveillance System, 2009. Prev Chronic Dis 2014;11:140190.

Tables

Table 1. Prevalence of Food Insecurity in 4 Southern US States, by Caregiving Status, Behavioral Risk Factor Surveillance System, 2015

State	Prevalence of food insecurity, weighted % (95% CI)		
	Caregivers (weighted n = 356,198)	Noncaregivers (weighted n = 652,737)	*P* value[a]
Overall	35.9 (33.9–37.9)	25.9 (24.8–26.9)	<.001
Alabama	35.3 (32.1–38.5)	25.9 (24.1–27.6)	<.001
Louisiana	38.2 (34.2–42.3)	25.9 (23.7–28.2)	<.001
Mississippi	34.9 (30.8–38.9)	29.3 (27.2–31.4)	.01
Tennessee	35.3 (31.0–39.6)	24.0 (21.8–26.2)	<.001

[a] Calculated by using χ^2 test.

Table 2. Sociodemographic Characteristics of Food Insecure Adults in Alabama, Louisiana, Mississippi, and Tennessee, by Caregiving Status, Behavioral Risk Factor Surveillance System, 2015

Characteristic	Proportion of food insecure adults, weighted % (95% CI)		
	Caregivers (weighted n = 356,198)	Noncaregivers (weighted n = 652,737)	*P* value[a]
Age, y			
18–34	18.1 (12.7–23.6)	13.2 (10.2–16.3)	<.001
35–64	66.6 (61.0–72.1)	60.3 (56.6–63.9)	
≥65	15.3 (12.2–18.4)	26.5 (23.6–29.4)	
Sex			
Female	66.4 (60.8–71.9)	66.6 (62.9–70.4)	.94
Male	33.6 (28.0–39.2)	33.3 (29.6–37.1)	
Race and ethnicity			
White	64.8 (59.2–70.3)	63.8 (60.2–67.4)	.20
Black	31.1 (25.6–36.5)	29.9 (26.5–33.2)	
Hispanic	1.9 (0.4–3.4)	1.4 (0.5–2.4)	
Other[b]	2.3 (0.9–3.7)	4.9 (2.9–6.8)	
Education			
High school or less	63.3 (58.2–68.4)	65.3 (61.8–68.8)	.38
Some college	28.8 (24.1–33.6)	25.5 (22.2–28.8)	
College graduate	7.9 (5.8–9.9)	9.2 (7.5–10.9)	
Marital status			
Married/couple	46.1 (40.5–51.6)	46.3 (42.4–50.2)	.07
Divorced/separated/widowed	31.8 (26.8–36.8)	37.4 (33.9–40.8)	
Never married	22.2 (16.6–27.7)	16.3 (13.4–19.3)	
Employment			
Employed	36.1 (30.7–41.5)	28.4 (24.8–31.9)	.04
Unemployed	38.0 (32.4–43.6)	39.7 (35.9–43.6)	
Unable to work	25.9 (21.3–30.6)	31.9 (28.5–35.3)	
Annual household income, $			
<15,000	28.4 (23.2–33.6)	28.2 (25.1–31.4)	.83
15,000–34,999	46.7 (41.2–52.3)	45.7 (41.8–49.6)	
35,000–49,999	10.4 (6.6–14.2)	12.5 (9.8–15.2)	
≥50,000	14.5 (10.2–18.8)	13.6 (10.6–16.6)	
Health insurance			
No	25.0 (19.6–30.5)	13.8 (10.8–16.8)	<.001
Yes	74.9 (69.5–80.4)	86.2 (83.2–89.2)	
Metropolitan statistical area			
No	32.1 (27.3–36.9)	31.8 (28.6–34.9)	.91
Yes	67.9 (63.1–72.7)	68.3 (65.1–71.4)	

[a] Calculated by using χ^2 test.
[b] "Other" includes American Indian, Alaska Native, Asian, Native Hawaiian, Pacific Islander, or multiracial individuals.

Table 3. Caregiving-Related Characteristics and Food Insecurity Prevalence Among Caregivers in Alabama, Louisiana, Mississippi, and Tennessee (Weighted N = 356,198), Behavioral Risk Factor Surveillance System, 2015

Caregiver characteristic	Proportion of caregivers, weighted % (95% CI)	Prevalence of food insecurity, weighted % (95% CI)
Relationship with care recipient		
Parent/parent-in-law	42.9 (39.5–46.3)	28.9 (23.6–34.3)
Child or grandchild	10.7 (8.7–12.6)	37.2 (27.7–46.7)
Spouse/partner	19.5 (17.1–21.9)	29.6 (23.5–35.8)
Other[a]	26.9 (23.8–30.0)	33.6 (26.9–40.3)
Care recipient health problems		
Mental or neurologic	18.9 (16.4–21.4)	29.6 (22.8–36.4)
Metabolic or cardiovascular	13.9 (11.8–16.0)	33.4 (25.6–41.3)
Musculoskeletal	6.8 (4.8–8.8)	48.3 (32.6–63.9)
Cancer	7.8 (5.8–9.7)	33.2 (18.8–47.6)
Other[b]	52.6 (49.2–55.9)	28.7 (24.5–32.9)
Caregiving 40 hours or more per week		
No	73.6 (70.6–76.7)	28.3 (24.6–32.0)
Yes	26.4 (23.3–29.4)	39.2 (32.3–46.1)
Months spent caregiving		
Less than 6 mos	25.4 (22.6–28.2)	24.0 (18.8–29.3)
6 mos to 2 y	18.2 (15.5–20.9)	34.0 (25.5–42.6)
2 y or more	56.4 (53.1–59.7)	33.5 (29.0–37.9)
Need for caregiver support services		
No	81.3 (78.7–83.9)	26.9 (23.3–30.6)
Yes	18.7 (16.1–21.3)	49.5 (41.8–57.3)
Helping with ADL		
No	38.0 (34.8–41.2)	25.6 (20.9–30.3)
Yes	62.0 (58.8–65.2)	34.6 (30.2–39.1)
Helping with IADL		
No	15.5 (13.2–17.7)	23.3 (17.2–29.4)
Yes	84.5 (82.3–86.8)	32.7 (28.9–36.3)

Abbreviations: ADL, activities of daily living; IADL, instrumental activities of daily living.

[a] Other relationships include grandparents, siblings-in-law, other relatives, and nonrelatives/friends.

[b] Other diseases include asthma, chronic obstructive pulmonary disease, substance abuse, addiction, HIV infection, or organ failure.

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
Volume 19, E64 OCTOBER 2022

ORIGINAL RESEARCH

Forgone Health Care for Non–COVID-19–Related Needs Among Medicare Beneficiaries During the COVID-19 Pandemic, Summer 2020–Winter 2021

Kara Tsuzaki, PharmD[1,2]; Deborah Taira, ScD[1]

Accessible Version: www.cdc.gov/pcd/issues/2022/22_0110.htm

Suggested citation for this article: Tsuzaki K, Taira D. Forgone Health Care for Non–COVID-19–Related Needs Among Medicare Beneficiaries During the COVID-19 Pandemic, Summer 2020–Winter 2021. Prev Chronic Dis 2022;19:220110. DOI: https://doi.org/10.5888/pcd19.220110.

PEER REVIEWED

Summary

What is already known on this topic?

The pandemic has affected access to health care even for conditions unrelated to COVID-19.

What is added by this report?

This study identified factors related to forgone care during the COVID-19 pandemic, including COVID-19 vaccination status, age, sex, race and ethnicity, US region, availability of primary care telehealth appointments, and chronic conditions (heart disease, arthritis, depression, osteoporosis or a broken hip, and diabetes or high blood glucose).

What are the implications for public health practice?

Our findings highlight the need for health care reform and changes in policy to address the issue of access to care for people with chronic conditions during a pandemic or other public health emergency.

Abstract

Introduction

Forgone health care, defined as not using health care despite perceiving a need for it, is associated with poor health outcomes, especially among people with chronic conditions. The objective of our study was to examine how the pandemic affected forgone health care during 3 stages of the pandemic.

Methods

We used the Medicare Current Beneficiary Survey COVID-19 Rapid Response Questionnaire administered in summer 2020, fall 2020, and winter 2021 to examine sociodemographic characteristics, chronic diseases, COVID-19 vaccination status, and telehealth availability in relation to beneficiary reports of forgone health care.

Results

Of the 3 periods studied, the overall rate of forgone health care was highest in summer 2020 (20.8%), followed by fall 2020 (7.8%) and winter 2021 (6.5%). COVID-19 vaccination status, age, sex, race and ethnicity, US region, availability of primary care telehealth appointments, and chronic conditions (heart disease, arthritis, depression, osteoporosis or a broken hip, and diabetes or high blood glucose) were significantly related to forgone care.

Conclusion

High rates of forgone care among Medicare participants varied over time and were significantly related to beneficiary characteristics. Our findings highlight the need for health care reform and changes in policy to address the issue of access to care for people with chronic conditions during a pandemic or other public health emergency.

Introduction

The COVID-19 pandemic has led to widespread changes in the US health care system (1). Despite increases in the number of hospitalizations and related care for COVID-19 infections, health care use overall declined because people stopped accessing care for non–COVID-19–related conditions (2). Even with increased availability of telehealth options, forgone health care, defined as not using health care despite perceiving a need for it, increased during the pandemic (3–5).

People aged 65 years or older with comorbidities such as heart disease, lung disease, and diabetes are at higher risk than younger populations for severe health outcomes from COVID-19 (6). Additionally, people with chronic conditions may be more likely than those without to forgo care because of a lack of coordinated care between practitioners, hospitals, and clinics (7).

The problem of forgone health care encompasses multiple factors such as government and provider policy, perceived risk, and financial costs (8–10). The COVID-19 pandemic has directly affected all 3 factors. At the start of the pandemic, many states initiated stay-at-home orders that remained in place for months; elective surgeries were canceled, regular checkups were delivered via telehealth, clinics were voluntarily closed, and people elected to forgo needed care. Finally, the pandemic had vast financial effects: an economic downturn was directly linked to the decline in use of health care services (4).

Failure to make use of health care services has been associated with poor health outcomes (7). Previous studies identified disparities in health care access associated with sociodemographic factors such as sex, race and ethnicity, and income (11,12). Research has begun to identify similar predictors of forgone care and barriers to care during the COVID-19 pandemic (13–16). The level of forgone care was expected to be highest in the first stages of the pandemic because of the perceived risk of exposure to COVID-19 as well as operational and procedural barriers such as clinic closures and drug shortages. Our objective was to examine how COVID-19 affected forgone care during 3 stages of the pandemic — summer 2020, fall 2020, and winter 2021.

Methods

We analyzed data from the Medicare Current Beneficiary Survey (MCBS) COVID-19 supplemental public use files for summer 2020 (17), fall 2020 (18), and winter 2021 (19). MCBS is a continuous and longitudinal survey that provides a representative national sample of the Medicare population. Medicare beneficiaries were contacted via telephone in June and July 2020 for the summer survey, October and November 2020 for the fall survey, and February, March, and April 2021 for the winter survey. The sample consisted of Medicare beneficiaries aged 65 years or older and people younger than 65 years with a Medicare-qualifying disability (eg, end-stage renal disease); 11,114 beneficiaries were interviewed for the summer survey, 9,686 for the fall survey, and 11,107 for the winter survey. All participants were continuously enrolled in Medicare and living in the community. Weighting was adjusted according to preliminary weights, eligibility, and completion of the survey. Because MCBS consists of de-identified, publicly available data, institutional review board approval was not sought for this study.

The primary dependent variable was forgone health care because of the pandemic. This information was self-reported in response to the question “Since [reference date], did you need medical care for something other than coronavirus, but not get it because of the coronavirus pandemic?” Forgone care included the need for urgent care, surgery, diagnostic tests, regular checkups, treatment of previous conditions, prescription drugs, and dental, vision, and hearing care. Information about the availability of telemedicine was obtained by asking, “Does your usual provider offer telephone or video appointments, so that you don’t need to physically visit their office or facility?” For information on vaccination status during winter 2021, beneficiaries were asked, “Since the [date of COVID-19 vaccine available] have you had a coronavirus vaccination?”

Sociodemographic characteristics assessed were age (<65, 65–74, or ≥75 y), sex (male or female), race and ethnicity (non-Hispanic Black, Hispanic, non-Hispanic White, or Other [any person that did not identify as non-Hispanic Black, Hispanic, or non-Hispanic White]), language spoken at home (only English spoken at home or other language besides English spoken at home), whether the respondent resided in a metropolitan (>50,000 people) or a nonmetropolitan area (≤50,000 people), US region of residence (Northeast, Midwest, South, or West), and annual household income (<$25,000 or ≥$25,000 per year).

Data on chronic conditions were self-reported by the beneficiary. For diabetes or high blood glucose, beneficiaries were asked, “Has a doctor or other health professional ever told [you] that [you/he/she] had any type of diabetes, including: sugar diabetes, high blood sugar, [borderline diabetes, pre-diabetes, or pregnancy-related diabetes/borderline diabetes, or pre-diabetes]?” The survey then asked about the beneficiary’s overall health and chronic conditions. We chose chronic conditions for this study on the basis of preliminary prevalence estimates of comorbidities associated with COVID-19 in the US as well as conditions that may present a physical obstacle to accessing care (20).

We weighted all frequencies to appropriately represent the national population. We used the χ^2 test to conduct a cross-tabulation analysis of differences between each demographic characteristic and forgone care. $P \leq .05$ indicates significance. We examined forgone care in relation to sociodemographic factors, vaccination status, telehealth availability, and our 5 selected chronic conditions. We used RStudio version 4.1.0 (RStudio Team) and Stata version 17 (StataCorp LLC) to perform all statistical analyses.

Results

Overall, survey response rates were 78.9% for summer 2020, 72.6% for fall 2020, and 79.6% for winter 2021. The response rate for the primary variable of forgone care was 99.98% for all 3 surveys. The percentage of beneficiaries with forgone care was highest in summer 2020 (20.8%), followed by fall 2020 (7.8%) and winter 2021 (6.5%).

Demographic characteristics were similar in all 3 periods (Table 1). By age, the largest group of survey respondents was aged 65–74 years. Most were female (range, 54.6%–54.9%), approximately 75% were non-Hispanic White, almost 90% spoke only English at home, and approximately 80% lived in metropolitan areas. By region, the largest group lived in the South (range, 38.1%–39.5%), and approximately two-thirds had an annual household income of $25,000 or more.

In winter 2021, beneficiaries who were vaccinated for COVID-19 were significantly less likely than unvaccinated beneficiaries to forgo health care (6.5% vs 6.7%; $P < .001$).

Forgone care was significantly related to many demographic characteristics (Table 2). Beneficiaries aged younger than 65 years had the highest rates of forgone care, by age, in all 3 periods (21.7%, 9.5%, and 10.2%, respectively; $P < .001$). Other significant predictors were sex, race and ethnicity, US region, and annual household income. Women were more likely than men to forgo care in summer 2020 (21.6% vs 19.7%; $P = .05$). By race and ethnicity, non-Hispanic White beneficiaries (22.5%) were most likely and non-Hispanic Black beneficiaries (12.6%) were least likely to forgo care in summer 2020 ($P = .002$); however, we found no significant differences by race and ethnicity in fall 2020 or winter 2021. By US region, the Midwest was most likely (23.5%) and the South least likely (16.6%) in summer 2020 to forgo care ($P = .001$). In winter 2021, however, the West was most likely (8.2%) and the South least likely (5.3%) ($P = .03$). In summer 2020, an annual household income of $25,000 or more was associated with an increase in forgone care. We found no significant differences by income in fall 2020, but in winter 2021, an annual household income less than $25,000 was associated with more forgone care (7.0%) than an income of $25,000 or more (6.3%).

The proportion of primary care providers (PCPs) that beneficiaries said offered telehealth was similar in summer 2020, fall 2020, and winter 2021 at 60.1%, 62.8%, and 63.4%, respectively. The highest proportion of forgone care related to telehealth access was reported in summer 2020: 22.8% of beneficiaries whose PCP offered telehealth reported forgoing care, while 18.1% whose PCPs did not offer telehealth reported forgoing care (Figure). In all 3 periods, beneficiaries whose PCP offered telehealth appointments were more likely to forgo care.

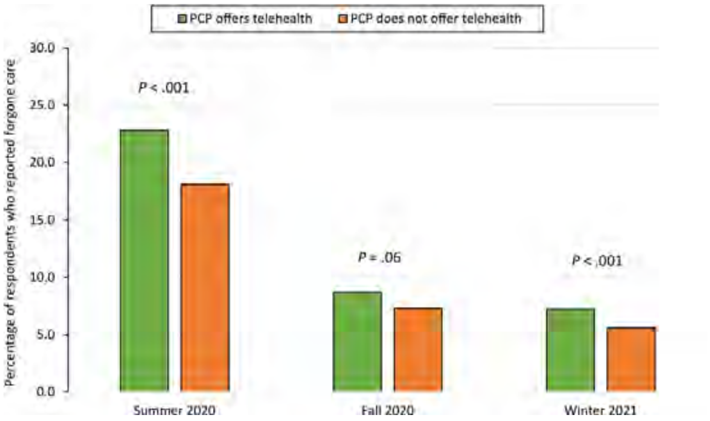


Figure. Percentage of Medicare beneficiaries who indicated forgoing health care, by whether primary care providers offered telehealth, summer 2020, fall 2020, and winter 2021. Availability of telemedicine was measured by asking Medicare beneficiaries, “Does your usual provider offer telephone or video appointments, so that you don’t need to physically visit their office or facility?” Abbreviation: PCP, primary care provider. Data source: Centers for Medicare & Medicaid Services (17–19).

Chronic conditions that were significantly associated with forgone care in all 3 periods were heart disease, arthritis, and depression (Table 3). For example, 22.4% ($P = .01$) of beneficiaries with heart disease reported forgoing care in summer 2020, 8.6% ($P = .05$) in fall 2020, and 7.6% ($P = .01$) in winter 2021. Additionally, osteoporosis or a broken hip was significantly associated with forgoing care in summer 2020 and winter 2021. Finally, diabetes was significantly associated with forgone care in fall 2020.

Discussion

In this study of Medicare survey respondents at 3 time points during the COVID-19 pandemic, we found a significant relationship between high rates of beneficiaries forgoing health care and beneficiary characteristics. Age was significantly associated at all 3 time points. By age group, beneficiaries aged 65 years or younger were the most likely to forgo care. These beneficiaries have disabilities such as end-stage renal disease, amyotrophic lateral sclerosis (ALS), and other clinical conditions that severely diminish their ability to work and they qualify for Medicare through a rigorous, consistent review process. Previous studies showed that people in this group are likely to identify as a racial or ethnic minority, be of low socioeconomic status, report mental health conditions, have comorbidities, and have high health care expenditures (21).

Chronic conditions were also a strong predictor of forgone care, especially heart disease, arthritis, and depression. According to the Centers for Disease Control and Prevention, the risk of severe illness from COVID-19 increases with age and number of underlying medical conditions. Diseases such as chronic kidney disease (any stage), dementia and other neurologic conditions, and disabilities are linked to a high risk for severe illness from COVID-19 (22). These conditions may have limited the type of care and risk of accessing care. Mental health status is also associated with disparities in health care access (16,23,24) and may be related to gaps in Medicare coverage for mental health disorders (24).

In our study, women were significantly more likely than men to forgo care in summer 2020. This mirrors a national trend in which more women than men forgo care (25). The decision to forgo care might result from gender-related differences in thoughts about COVID-19, perceived and real risks of COVID-19, and economic concerns (25,26). Our study showed that higher income (≥$25,000 per year) predicted not accessing care in summer 2020, but the opposite was true in winter 2021. Socioeconomic disparities in access to care are not as extreme in the Medicare population as in the general US population. However, differences still exist, potentially from gaps in coverage, high out-of-pocket costs, and lack of knowledge about the availability of telehealth and the ability to use it (27).

Race and ethnicity significantly predicted forgoing care in summer 2020, with the highest rates among non-Hispanic White beneficiaries. Although racial and ethnic disparities exist in US health care, such disparities are smaller among Medicare beneficiaries in terms of insurance coverage, access to care, and self-reported health (28). Our study showed that race and ethnicity were not significant predictors of not accessing care in fall 2020 or winter 2021.

In comparisons of US regions, the South is commonly associated with low levels of access to quality care (28,29). However, in our study of Medicare beneficiaries, by US region, the South had the lowest percentage of beneficiaries who did not have access needed health care in summer 2020 and winter 2021. It is possible the South has access-to-care issues unrelated to COVID-19 or access to care.

Overall, our findings may have been confounded by socioeconomic status, education, and overall feelings about COVID-19. Studies have associated race and ethnicity with perceptions of COVID-19 (10,12,26). Racial and ethnic minority populations were less likely than the White population to be vaccinated and more likely to disagree that COVID-19 is more severe than influenza during the early phases of the COVID-19 pandemic (26,30,31).

We found that in winter 2021, COVID-19–vaccinated beneficiaries were less likely than unvaccinated beneficiaries to forgo care. The availability and receipt of vaccines may have affected the overall reduction in forgone care from fall 2020 to winter 2021. However, the difference in the percentage of beneficiaries who reported forgoing care was small: 6.7% among unvaccinated beneficiaries and 6.5% among vaccinated beneficiaries. Potential confounders might include the passage of time or a sense of safety resulting from the availability and receipt of the COVID-19 vaccination.

Although the overall number of PCPs who reportedly offered telehealth services remained consistent over time, access to care may have been affected by logistical barriers. With the announcement of clinic closures and surgery cancellations, physicians may not have been prepared to offer telehealth at the scale needed or may have been unable to offer in-person visits. This factor may have affected the initial reduction in forgoing care from summer 2020 to fall 2020 as health care providers became better equipped to handle telehealth appointments.

Limitations

Our study has several limitations. First, all data were self-reported, which may have introduced recall bias, sampling bias, social desirability bias, and varying levels of introspection. The period of forgone care could have been recalled incorrectly. The survey also asked beneficiaries to self-report chronic diseases; we had no information on whether they were being treated for them. Additionally, although data were reported by US region, COVID-19 guidelines and mandates varied by state. Each supplemental survey covered a short period (2 or 3 months), which may not have allowed certain mandates or perceived risks to be implemented and understood. Although the MCBS sampling frame allows for estimates of national averages, it does not allow for an in-depth analysis of racial or ethnic minority groups, undocumented workers, or other groups of people who are socially and economically marginalized. Because of small sample sizes, analyses of non-Hispanic Black and Hispanic beneficiaries were especially susceptible to sampling errors (32). We could not identify additional racial and ethnic groups and included these under the general term “other.” These groups are affected by small sample sizes and also by the limitation of the survey being offered only in English and Spanish. Furthermore, because MCBS is conducted by the US government, people who have experienced historical trauma may not be inclined to participate.

Conclusion

Our study identified associations between various factors and access to health care among Medicare beneficiaries during the COVID-19 pandemic. Our results are corroborated by Park and Stimpson in regard to the significance found for age, chronic conditions, sex, income, race and ethnicity, and region (16). Their study highlighted the effect of physician-driven factors, such as availability of telehealth consultations, and mental health status on beneficiaries' decision to forgo health care. Our study includes an analysis of vaccination status and emphasizes multiple chronic conditions and beneficiary characteristics in relation to nonuse of care. It highlights the need for further research and policy change for Medicare beneficiaries in the US, especially populations with multiple chronic health conditions and low socioeconomic status. Additionally, future research is needed to fully understand the extent of forgone care among communities that are socially and economically marginalized.

Forgone or postponed care can have long-term health consequences. The lack of preventive care, including screenings and vaccinations, can lead to delayed treatment or not receiving proper care. Examples of interventions include the Coronavirus Aid, Relief, and Economic Security Act (33), which provided financial support to patients and medical providers, thereby decreasing barriers to care. Further research pertaining to access to care will continue to change with the introduction of booster shots, relaxation of mask and social distancing mandates, and perceived risks and attitudes toward COVID-19.

Acknowledgments

No financial support was received for this study. No copyrighted material was used in this article. The authors have no conflicts of interest to disclose.

Author Information

Corresponding Author: Kara Tsuzaki, PharmD, The Daniel K. Inouye College of Pharmacy, 200 W Kawili St, Hilo, HI 96720. Telephone: 808-391-5586. Email: ktsuzaki@hawaii.edu.

Author Affiliations: [1]The Daniel K. Inouye College of Pharmacy, Honolulu, Hawaii. [2]Now with Cynosure Consulting, LLC, Seattle, Washington.

References

1. Blumenthal D, Fowler EJ, Abrams M, Collins SR. COVID-19 — implications for the health care system. N Engl J Med 2020; 383(15):1483–8.
2. Heist T, Schwartz K. Trends in overall and non-COVID-19 hospital admissions. Kaiser Family Foundation; 2021.
3. Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly T, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic — United States, January–March 2020. MMWR Morb Mortal Wkly Rep 2020;69(43):1595–9.
4. Anderson KE, McGinty EE, Presskreischer R, Barry CL. Reports of forgone medical care among US adults during the initial phase of the COVID-19 pandemic. JAMA Netw Open 2021;4(1):e2034882.
5. Jaffe DH, Lee L, Huynh S, Haskell TP. Health inequalities in the use of telehealth in the United States in the lens of COVID-19. Popul Health Manag 2020;23(5):368–77.
6. Alam MR, Kabir MR, Reza S. Comorbidities might be a risk factor for the incidence of COVID-19: evidence from a web-based survey. Prev Med Rep 2021;21:101319.
7. McDonald KM, Sundaram V, Bravata DM, Lewis R, Lin N, Kraft SA, et al.Background: ongoing efforts in care coordination and gaps in the evidence. In: Closing the quality gap: a critical analysis of quality improvement strategies. Rockville (MD): Agency for Healthcare Research and Quality; 2007.
8. Ksinan Jiskrova G, Bobák M, Pikhart H, Ksinan AJ. Job loss and lower healthcare utilisation due to COVID-19 among older adults across 27 European countries. J Epidemiol Community Health 2021;75(11):1078–83.
9. Roadevin C, Hill H. How can we decide a fair allocation of healthcare resources during a pandemic? J Med Ethics 2021; 47(12):medethics-2020-106815.
10. Burgard SA, Hawkins JM. Race/ethnicity, educational attainment, and foregone health care in the United States in the 2007–2009 recession. Am J Public Health 2014; 104(2):e134–40.
11. Kurichi JE, Pezzin L, Streim JE, Kwong PL, Na L, Bogner HR, et al. Perceived barriers to healthcare and receipt of recommended medical care among elderly Medicare beneficiaries. Arch Gerontol Geriatr 2017;72:45–51.
12. Stepanikova I, Oates GR. Perceived discrimination and privilege in health care: the role of socioeconomic status and race. Am J Prev Med 2017;52(1S1):S86–94.
13. Alobuia WM, Dalva-Baird NP, Forrester JD, Bendavid E, Bhattacharya J, Kebebew E. Racial disparities in knowledge, attitudes and practices related to COVID-19 in the USA. J Public Health (Oxf) 2020;42(3):470–8.
14. Caraballo C, Massey D, Mahajan S, Lu Y, Annapureddy AR, Roy B, et al. Racial and ethnic disparities in access to health care among adults in the United States: a 20-year National Health Interview Survey analysis, 1999–2018. MedRxiv. Published online November 4, 2020.
15. Quan D, Luna Wong L, Shallal A, Madan R, Hamdan A, Ahdi H, et al. Impact of race and socioeconomic status on outcomes in patients hospitalized with COVID-19. J Gen Intern Med 2021;36(5):1302–9.
16. Park S, Stimpson JP. Trends in self-reported forgone medical care among Medicare beneficiaries during the COVID-19 pandemic. JAMA Health Forum 2021;2(12):e214299.

17. Centers for Medicare & Medicaid Services. 2020Medicare Current Beneficiary Survey COVID-19 summer supplement public use file [data files and code books]. Accessed March 18, 2022. https://www.cms.gov/research-statistics-data-and-systems/downloadable-public-use-files/mcbs-public-use-file
18. Centers for Medicare & Medicaid Services. 2020Medicare Current Beneficiary Survey COVID-19 fall supplement public use file [data files and code books]. Accessed March 18, 2022. https://www.cms.gov/research-statistics-data-and-systems/downloadable-public-use-files/mcbs-public-use-file
19. Centers for Medicare & Medicaid Services. 2021Medicare Current Beneficiary Survey COVID-19 winter supplement public use file [data files and code books]. Accessed March 18, 2022. https://www.cms.gov/research-statistics-data-and-systems/downloadable-public-use-files/mcbs-public-use-file
20. Chow N, Fleming-Dutra K, Gierke R, Hall A, Hughes M, Pilishvili T, et al.; CDC COVID-19 Response Team. COVID-19 Response Team. Preliminary estimates of the prevalence of selected underlying health conditions among patients with coronavirus disease 2019 — United States, February 12–March 28, 2020. MMWR Morb Mortal Wkly Rep 2020; 69(13):382–6.
21. Iezzoni LI. Quality of care for Medicare beneficiaries with disabilities under the age of 65 years. Expert Rev Pharmacoecon Outcomes Res 2006;6(3):261–73.
22. Centers for Disease Control and Prevention. People with certain medical conditions. Updated May 2, 2020. Accessed March 27, 2022. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html
23. Saltzman LY, Lesen AE, Henry V, Hansel TC, Bordnick PS. COVID-19 mental health disparities. Health Secur 2021; 19(S1):S5–13.
24. Balasuriya L, Quinton JK, Canavan ME, Holland ML, Edelman EJ, Druss BG, et al. The association between history of depression and access to care among Medicare beneficiaries during the COVID-19 pandemic. J Gen Intern Med 2021; 36(12):3778–85.
25. Cameron KA, Song J, Manheim LM, Dunlop DD. Gender disparities in health and healthcare use among older adults. J Womens Health (Larchmt) 2010;19(9):1643–50.
26. Alqahtani MMJ, Arnout BA, Fadhel FH, Sufyan NSS. Risk perceptions of COVID-19 and its impact on precautionary behavior: a qualitative study. Patient Educ Couns 2021; 104(8):1860–7.
27. McBride TD. Disparities in access to Medicare managed care plans and their benefits. Health Aff (Millwood) 1998; 17(6):170–80.
28. National Academy of Social Insurance Study Panel on Medicare and Disparities. Strengthening Medicare’s role in reducing racial and ethnic health disparities. Medicare Brief 2006;(16):1–4.
29. Jha AK, Orav EJ, Epstein AM. Low-quality, high-cost hospitals, mainly in South, care for sharply higher shares of elderly Black, Hispanic, and Medicaid patients. Health Aff (Millwood) 2011;30(10):1904–11.
30. Vai B, Cazzetta S, Ghiglino D, Parenti L, Saibene G, Toti M, et al. Risk perception and media in shaping protective behaviors: insights from the early phase of COVID-19 Italian outbreak. Front Psychol 2020;11:563426.
31. Padamsee TJ, Bond RM, Dixon GN, Hovick SR, Na K, Nisbet EC, et al. Changes in COVID-19 vaccine hesitancy among Black and White individuals in the US. JAMA Netw Open 2022;5(1):e2144470.
32. Adler GS. A profile of the Medicare Current Beneficiary Survey. Health Care Financ Rev 1994;15(4):153–63.
33. Relief and Economic Security Act or the CARES Act. Public Law No. 116–136/ HR 748 (2020). Accessed July 6, 2022. https://www.congress.gov/bill/116th-congress/house-bill/748/all-info

Tables

Table 1. Demographic Characteristics of Medicare Beneficiaries During 3 Periods of the COVID-19 Pandemic, Summer 2020, Fall 2020, and Winter 2021[a]

Characteristic	Summer (June–July) 2020, %	Fall (October–November) 2020, %	Winter (February–April) 2021, %
Age, y			
<65	14.6	20.5	14.6
65–74	53.3	47.9	51.5
≥75	32.1	31.6	33.9
Sex			
Male	45.1	45.1	45.4
Female	54.9	54.9	54.6
Race and ethnicity			
Black, non-Hispanic	9.9	9.7	9.6
Hispanic	8.0	8.4	8.4
Other[b]	6.4	6.3	6.2
White, non-Hispanic	75.7	75.7	75.8
Language			
Other language besides English spoken at home	11.0	11.4	11.1
Only English spoken at home	88.9	88.6	88.9
Residence			
Metropolitan	79.8	80.1	80.0
Nonmetropolitan	20.2	19.9	20.0
US region			
Northeast	17.9	17.9	17.5
Midwest	22.1	22.0	21.5
South	38.1	38.3	39.5
West	21.8	21.9	21.6
Annual household income, $			
<25,000	30.6	30.8	29.3
≥25,000	65.7	65.6	67.4

[a] Data source: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey COVID-19 supplements (17–19). Not all survey participants answered all questions; percentages may not sum to 100%.
[b] “Other” includes any person that did not identify as non-Hispanic Black, Hispanic, or non-Hispanic White.

Table 2. Demographic Characteristics of Medicare Beneficiaries Who Reported Forgoing Non–COVID-19–Related Health Care During 3 Periods, Summer 2020, Fall 2020, and Winter 2021[a]

Characteristic	Summer 2020		Fall (October–November) 2020		Winter (February–April) 2021	
	%	*P* value[b]	%	*P* value[b]	%	*P* value[b]
Reported forgone care due to COVID-19[c]	20.8	—	7.8	—	6.5	—
Age, y						
<65	21.7	<.001	9.5	.003	10.2	<.001
65–74	22.3		8.1		6.6	
≥75	17.8		6.3		4.9	
Sex						
Male	19.7	.05	7.5	.57	5.8	.07
Female	21.6		8.0		7.1	
Race and ethnicity						
Black non-Hispanic	12.6	.002	7.0	.68	6.8	.30
Hispanic	16.3		8.5		8.0	
Other[d]	18.4		7.6		6.9	
White non-Hispanic	22.5		7.8		6.3	
Language						
Other language besides English spoken at Home	15.6	.11	7.3	.06	7.2	.22
Only English spoken at Home	21.4		7.9		6.5	
Residence						
Metropolitan	20.9	.81	7.8	.62	6.7	.06
Nonmetropolitan	20.1		7.9		5.8	
US Region						
Northeast	23.2	.001	8.4	.29	8.0	.03
Midwest	23.5		7.7		6.0	
South	16.6		6.9		5.3	
West	23.4		9.1		8.2	
Annual household income, $						
<25,000	15.5	<.001	7.2	.20	7.0	.008
≥25,000	23.6		8.2		6.3	

Abbreviation: — , does not apply.

[a] Data source: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey COVID-19 supplements (17–19).

[b] Determined by χ^2 test.

[c] Question on survey was "Since [reference date], did you need medical care for something other than coronavirus, but not get it because of the coronavirus pandemic?"

[d] "Other" includes any person that did not identify as non-Hispanic Black, Hispanic, or non-Hispanic White.

Table 3. Selected Chronic Conditions Among Medicare Beneficiaries Who Responded to a Survey on Forgone Health Care During 3 Periods, Summer 2020, Fall 2020, and Winter 2021[a]

	Summer 2020		Fall 2020		Winter 2021	
Characteristic	**% of Respondents who reported forgoing care**	***P* value[b]**	**% of Respondents who reported forgoing care**	***P* value[b]**	**% of Respondents who reported forgoing care**	***P* value[b]**
Heart disease	22.4	.01	8.6	.05	7.6	.01
Diabetes or high blood glucose	22.0	.17	6.9	.001	7.3	.71
Arthritis	22.5	<.001	9.4	.001	6.2	.05
Osteoporosis or broken hip	24.2	<.001	9.4	.14	8.0	.04
Depression	24.8	.002	10.4	<.001	9.2	<.001

[a] Data source: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey COVID-19 supplements (17–19).
[b] Difference between percentage of people who self-reported having a chronic disease and percentage who did not; determined by χ^2 test.

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
Volume 19, E56 SEPTEMBER 2022

ORIGINAL RESEARCH

Exploring Residents' Perceptions of Neighborhood Development and Revitalization for Active Living Opportunities

Nishita Dsouza, PhD[1]; Natalicio Serrano, PhD[2]; Kathleen B. Watson, PhD[3]; Jean McMahon, PhD[4]; Heather M. Devlin, MA[3]; Stephenie C. Lemon, PhD[5]; Amy A. Eyler, PhD[2]; Jeanette Gustat, PhD[6]; Jana A. Hirsch, PhD[7]

Accessible Version: www.cdc.gov/pcd/issues/2022/22_0033.htm

Suggested citation for this article: Dsouza N, Serrano N, Watson KB, McMahon J, Devlin HM, Lemon SC, et al. Exploring Residents' Perceptions of Neighborhood Development and Revitalization for Active Living Opportunities. Prev Chronic Dis 2022;19:220033. DOI: https://doi.org/10.5888/pcd19.220033.

PEER REVIEWED

Summary

What is already known about this topic?

New or improved infrastructure in neighborhoods aiming to increase physical activity might lead to unintended social, development, or economic pressures; however, little empirical research exists about residents' opinions and perceptions of these tradeoffs.

What is added by this report?

Our report sheds light on the complexities of public opinions on neighborhood improvements and highlights the need for community engagement on new built environment projects.

What are the implications for public health practice?

Participatory planning holds potential for community engagement in decisions about active living infrastructure. Public health and planning professionals can also partner with communities to select the most appropriate measures for each context and evaluate each project for equity and effectiveness.

Abstract

Introduction

Community fears of gentrification have created concerns about building active living infrastructure in neighborhoods with low-income populations. However, little empirical research exists related to these concerns. This work describes characteristics of residents who reported 1) concerns about increased cost of living caused by neighborhood development and 2) support for infrastructural improvements even if the changes lead to a higher cost of living.

Methods

Data on concerns about or support for transportation-related and land use–related improvements and sociodemographic characteristics were obtained from the 2018 SummerStyles survey, an online panel survey conducted on a nationwide sample of US adults (n = 3,782). Descriptive statistics characterized the sample, and χ^2 tests examined associations among variables.

Results

Overall, 19.1% of study respondents agreed that development had caused concerns about higher cost of living. Approximately half (50.7%) supported neighborhood changes for active living opportunities even if they lead to higher costs of living. Prevalences of both concern and support were higher among respondents who were younger and who had higher levels of education than their counterparts. Support did not differ between racial or ethnic groups, but concern was reported more often by Hispanic/Latino (28.9%) and other non-Hispanic (including multiracial) respondents (25.5%) than by non-Hispanic White respondents (15.6%). Respondents who reported concerns were more likely to express support (65.3%) than respondents who did not report concerns (47.3%).

Conclusion

The study showed that that low-income, racial, or ethnic minority populations support environmental changes to improve active living despite cost of living concerns associated with community revitalization.

Introduction

Regular physical activity positively influences 7 of the 10 most common chronic conditions diagnosed in the US (1). Physical activity improves cognition, decreases depression, and is associated with a reduction in early death and risk for chronic diseases such as coronary heart disease, stroke, type 2 diabetes, obesity, depression, and many forms of cancer (1). Despite the benefits to health, less than half (46%) of US adults engage in enough aerobic physical activity to achieve substantial health benefits (2).

Making physical activity a part of everyday living makes it easier to achieve the benefits of regular physical activity (1). Changes in community design can create opportunities for physical activity and make neighborhoods more supportive of active living (3). The Community Preventive Services Task Force recommends built environment approaches that combine improvements in transportation such as sidewalks, bicycle lanes, and expanded public transit, with changes in land use and community design such as improved parks and recreation facilities and mixed-use development that enable housing in proximity to destinations such as businesses and schools (3).

Communities with low-income populations often have minimal resources for physical activity (4). Furthermore, racial and ethnic minorities, who disproportionately reside in communities with low-income populations, tend to have high rates of leisure-based physical inactivity and chronic diseases such as cardiovascular disease (5). Even when facilities such as parks exist in these communities, they tend to have few amenities, often show neglect, and project a violent or unsafe environment to community residents (6). Geographic health disparities by race, ethnicity, and income have created interest in addressing health equity by improving neighborhood environments through community development (7).

Community development and revitalization strategies that improve neighborhood environments can promote physical activity and improve health in disinvested communities (8). These broad initiatives benefit from multidisciplinary collaboration among health, planning, housing, transportation, and government for the ultimate goal of improving the lives of the community residents (9).

Widespread concerns about gentrification associated with robust community development and revitalization exist (10). However, mixed findings show how active living infrastructure may contribute to these concerns. For example, a study in 3 cities described concerns in communities with Black poulations about the installation of bike lanes and their possible contribution to gentrification (11). Empirical research has found that property values rise as proximity to bicycle facilities increases (12). However, a recent longitudinal analysis of bike lanes showed that these investments were not associated with changing demographics in a neighborhood, and that bike lanes were more commonly installed in neighborhoods with low-income White populations than in neighborhoods with low-income Black populations (13). Contrasting this research are findings that showed that intensive community development such as greenways near downtown areas may increase gentrification (14). Higher property values may encourage gentrification by encouraging long-time homeowners to sell their homes to capture their increased wealth; higher property values can also lead to higher rents, potentially displacing current renters who may no longer be able to afford living in that neighborhood (15). Through this displacement, the remaining residents may lose a sense of belonging in their own neighborhood as their surrounding demographics change, which can negatively affect health and quality of life (10).

Concerns about neighborhood change and potential displacement may present barriers to community support for changes in the built environment to improve access to physical activity. However, limited empirical research exists about residents' opinions of these potential tradeoffs. Thus, there is a need to understand residents' perceptions of neighborhood development and revitalization, and their concerns about the potential for displacement of current residents. Our study aimed to describe perceptions and characteristics of residents who reported 1) concerns about increased cost of living from neighborhood development and revitalization and 2) support for neighborhood changes to make it easier to walk or bike even if the changes could lead to a higher cost of living. Because increased cost of living disproportionately affects low-income residents and because racial and ethnic minority populations are overrepresented in disinvested neighborhoods, we hypothesized that low-income and racial and ethnic minority populations would be more likely than non-Hispanic White populations to report concerns about the increased cost of living from neighborhood development and revitalization and less likely to express support for neighborhood changes to make it easier to walk or bike even if the changes could lead to higher cost of living (16).

Methods

SummerStyles sample

The 2018 Porter Novelli ConsumerStyles' database is built from a series of web-based surveys via the GfK KnowledgePanel that gathers insights about US consumers, including information about their lifestyle, health, knowledge, and behaviors (17). Panel members are randomly recruited by using probability-based sampling by address. The panel is continuously replenished and maintains approximately 55,000 panelists. The initial SpringStyles survey was sent from March 21, 2018, to April 11, 2018, to 10,904 panelists. SpringStyles respondents included 6,427 adults who completed the survey for a response rate of 58.9%. Those who completed the SpringStyles survey received reward points worth approximately $5.

Our study used data from the subsequent SummerStyles survey, which was sent to 5,584 respondents that completed SpringStyles, from June 12, 2018, to July 7, 2018. The subsequent SummerStyles survey included survey questions that were not in the initial SpringStyles survey. The final sample had 4,088 adults (response rate = 73.2%). Those who completed the SummerStyles survey also received reward points worth approximately $5. The data were then weighted to match the 2018 US Current Population Survey proportions for sex, age, annual household income, race and ethnicity, household size, education level, census region, and metropolitian statistical area status (18).

Of the 4,088 respondents, we excluded data from 306 respondents (7.4%) who were missing information on concerns and support for active living development (n = 40), physical activity (n = 71), body mass index (n = 64), smoking (n = 98), air pollution (n = 12), and neighborhood features of concern (n = 21). The final analytic sample had data from 3,782 respondents.

Concerns for neighborhood revitalization and support for walking and biking infrastructure

Respondents were asked about their agreement (using a 5-point Likert scale) with 2 statements: 1) "My neighborhood is experiencing development or revitalization that has caused concerns about higher cost of living"; 2) "I would support changes to my neighborhood to make it easier to walk or bike even if the changes lead to a higher cost of living for me." Survey questions were developed de novo and were not cognitively tested.

To compare those who agreed there were concerns with those who did not, concern was dichotomized into "concerned" (by grouping "somewhat agree" and "strongly agree") and "not concerned" (by grouping "strongly disagree," "somewhat disagree," and "neither agree nor disagree"). To better distinguish between those who did not support and those who were neutral because the latter group may need different strategies for change, support was categorized into 3 groups: "supporters" (by grouping "somewhat agree" and "strongly agree"), "nonsupporters" (by grouping "strongly disagree" and "somewhat disagree"), and neither ("neither agree nor disagree"). To assess specific concerns, we asked respondents "Which of the following changes to your neighborhood or community would cause the most concern about higher cost of living?" Respondents selected 1 of the following options: new sidewalks or stop signs, new bicycle lanes or paths, expanded public transportation, new businesses with condos above, improved parks and recreational facilities, or none of these would cause concern.

Sociodemographic and health characteristics

Respondents self-reported sociodemographic and economic characteristics including sex (male, female), age category (18–34, 35–49, 50–64, ≥65 years), education (high school graduate or less, some college, college graduate or more), race and ethnicity (Hispanic/Latino, non-Hispanic Black, non-Hispanic White, other non-Hispanic [including multiracial]), annual household income (<$50,000, $50,000–$99,999, ≥$100,000), current employment status (working, retired or not working), and housing type (one family house, apartment or other). ConsumerStyles' database provided geographic information on US Census region (Northeast, Midwest, South, West) and metropolitan statistical area status (nonmetropolitan, metropolitan) (19).

Respondents also self-reported health behaviors (aerobic physical activity and smoking status), anthropometry (height and weight), and a health behavior–related decision about air pollution. To assess physical activity, we used modified versions of the National Health Interview Survey physical activity questions (20). We asked respondents how often in a usual week and, if applicable, the amount of time during leisure time that they participated for at least 10 minutes at a time in 1) vigorous-intensity activities (ie, heavy sweating or large increases in breathing or heart rate) and 2) moderate-intensity activities (ie, medium sweating or moderate increase in breathing or heart rate). To classify adults into levels of physical activity, we calculated minutes of moderate-intensity equivalent activity by counting 1 minute of vigorous-intensity activity as 2 minutes of moderate-intensity activity (1). We then classified respondents into 3 activity levels by using the current adult aerobic guideline (1): 1) active, reporting at least 150 minutes per week of moderate-intensity equivalent physical activity; 2) insufficiently active, reporting some moderate-intensity equivalent physical activity but not enough to meet active definition; and 3) inactive, reporting no moderate-intensity equivalent physical activity that lasted at least 10 minutes. We assessed smoking status by using 2 questions, one about lifetime cigarette use and one about current cigarette use. We combined these and classified respondents into 3 categories: 1) current smoker (respondents who self-reported having smoked at least 100 cigarettes in their lifetime and currently smoked some days or every day); 2) former smoker (respondents who reported having smoked at least 100 cigarettes in their lifetime and currently smoked not at all); and 3) never smoker (respondents who reported having smoked fewer than 100 cigarettes in their lifetime). We used self-reported anthropometry to calculate body mass index (BMI, calculated as weight in kilograms divided by the square of height in meters) and categorized respondents by using standard cut points (21): 1) underweight/normal (<25.0); 2), overweight (25.0–29.9), and 3) obesity (≥30.0). Finally, we asked respondents about decisions related to air pollution exposure by using the question, “When walking, biking, or exercising outdoors, how often do you avoid busy roads to reduce your exposure to air pollution?” (always, usually, sometimes, rarely, never, don’t know).

Statistical analyses

We calculated descriptive statistics (weighted and unweighted) for all sociodemographic and health characteristics. We calculated prevalence and 95% CIs for the following: 1) agreement that neighborhood development or revitalization has caused concerns about higher cost of living, 2) support for active living improvements even if they lead to a higher cost of living, and 3) specific changes in neighborhood transportation-related and land use–related features. We stratified prevalences by respondent characteristics. We tested assocations between concern, support, and neighborhood features and respondent characteristics by using adjusted Wald χ^2 tests. Where appropriate, we used pairwise *t* tests with a Bonferroni correction and orthogonal polynomial contrasts to identify significant pairwise differences and trends by participant characteristics. We considered tests significant at $P < .05$, Bonferroni adjusted. All analyses were conducted in 2020 by using SUDAAN release 11 (RTI International) to account for survey weights.

Results

The largest unweighted percentages of respondents for each demographic group were women, non-Hispanic White, 50 to 64 years old, currently employed, living in a 1-family house, and living in a nonmetropolitan area (Table 1). Slightly more than half were active and never smoked, although more than half were overweight or had obesity. Almost half attempted to reduce air pollution exposure when walking, biking, or exercising outdoors by avoiding busy roads.

Almost 1 in 5 respondents reported that development or revitalization had caused concerns about higher cost of living in their neighborhood (19.1%; 95% CI, 17.7%–20.6%) (Table 2). Concern decreased with increasing age and increased with increasing education and physical activity levels. Concern was more prevalent among respondents who were Hispanic/Latino or other non-Hispanic (including multiracial) versus non-Hispanic White; were currently employed versus retired or not working; lived in nonmetropolitan versus metropolitan areas; and lived in the West versus other regions.

Overall, approximately half of respondents (50.7%; 95% CI, 48.9%–52.6%) supported changes to make it easier to walk or bike even if they lead to a higher cost of living (Table 3). Respondents who reported concerns about higher cost of living in their neighborhood were more likely to express support (65.3%) than respondents who did not report concerns (47.3%). Similar to the prevalence of concerns about neighborhood development, the prevalence of support decreased with increasing age and increased with increasing education and physical activity level. The prevalence of support also increased with increasing income and decreased with increasing BMI. We found no association between race or ethnicity and support for changes to make it easier to walk or bike even if they lead to a higher cost of living. Unlike concern, support did not vary by employment status, housing type, or region.

Of the specific changes in neighborhood features that could cause concern about higher cost of living, new businesses with condos above was the greatest land use–related concern (21.7%; 95% CI, 20.2%–23.3%), followed by improved parks and recreation facilities (8.2%; 95% CI, 7.2%–9.3%) (Figure). Expanded public transportation was the greatest transportation-related concern (8.0%; 95% CI, 7.1%–9.0%). Respondents who reported concern about higher cost of living from neighborhood development overall also had higher prevalence of concern than respondents who did not report concern across all infrastructure types, except new sidewalks or stop signs and improved parks and recreation facilities. Supporters (versus neither) of active living changes reported greatest concern about new businesses and improved parks, while nonsupporters (versus neither) reported greatest concern for new bicycle lanes or paths (Figure). Physical activity and BMI were the only respondent characteristics that were not associated with concerns about specific changes in neighborhood features.

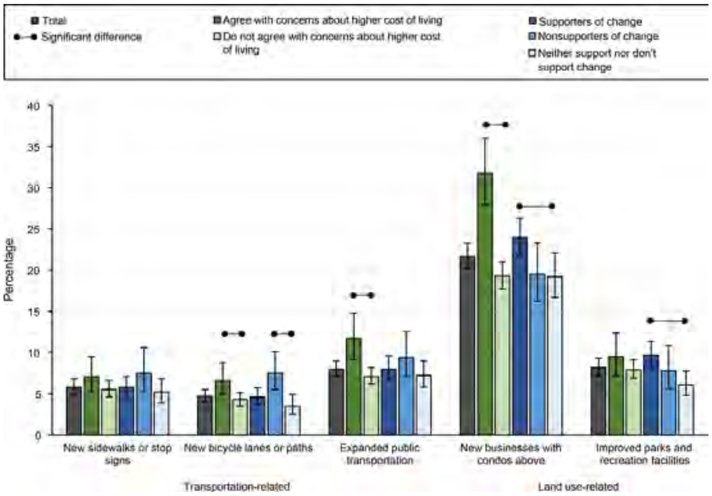


Figure. Prevalence of residents reporting specific changes in neighborhood features as causing concern, stratified by agreement with concerns about higher cost of living caused by changes and by support for changes to their neighborhoods even if the changes lead to higher cost of living, SummerStyles survey, 2018 (N = 3,782).

Discussion

This study leveraged data from a nationwide consumer panel survey to better understand perceptions of community development and revitalization strategies and whether residents were concerned about cost of living increases resulting from built environment changes to improve health. Results showed substantial levels of support for health-promoting neighborhood improvements in potential contradiction with concerns about increased costs among some of the same demographic groups. Respondents who reported concerns about development raising costs were more likely to support changes to make it easier to walk or bike even if they led to increased costs (65.3%) than respondents who did not report such concerns (47.3%). However, given that a small percentage of the population reported concerns to begin with, these inconsistencies are not highly prevalent in the overall population. Certain types of built environment changes, such as new businesses with condos above, were associated with more concern than other types of built environment changes, such as new sidewalks or stop signs. Understanding that residents may be both supportive and concerned, as well as understanding the sources of concern, may be useful for decision makers as they seek to build community support for built environmental changes to improve active living.

Cost of living is a concern among US residents (22). However, these results offer only limited support for our hypothesis that populations disproportionately affected by increases in cost would express more concern about and be less likely to support built environment changes to increase physical activity. In particular, the finding that support for changes to make it easier to walk or bike despite increasing costs was more prevalent among respondents reporting concerns compared with those not reporting concerns about development raising costs ran counter to our hypothesis. Survey results depicted greater concern among respondents living in the West, potentially because that area is developing and urbanizing more rapidly than other parts of the US (23). Consistent with previous SummerStyles analyses, our analysis found that respondents with higher education levels expressed increased concern about cost of living changes resulting from neighborhood development and revitalization (24).

More than half of study respondents supported built environment or infrastructural changes to promote active living even if the changes could lead to a higher cost of living, despite nearly 1 in 5 reporting concerns about such changes leading to increases in cost of living. This is consistent with another national survey examining support for policies that promote physical activity in neighborhood environments even when these policies are associated with tax increases (25). However, the present survey failed to confirm our hypothesis that low-income or racial or ethnic minority populations would not support environmental changes to improve active living.

Although this support from low-income or racial or ethnic minority populations may seem counterintuitive, other studies have discussed the complexity and nuances of perceptions of neighborhood revitalization and development, with residents expressing support while acknowledging a “not for us” sentiment (26). Additional investigation of the complexities of resident concerns and support for changes in the active living environment could guide implementation of active living improvements and help communities avoid unintended consequences. Involving community members meaningfully in neighborhood revitalization processes could also enhance understanding of these complexities. Further research could clarify whether inclusive policies could help ensure that economic development benefits existing residents who have often been historically disinvested and excluded (27).

When examining specific transportation-related and land use–related neighborhood changes that could cause concern about a higher cost of living, respondents were more likely to be concerned about changes representing larger investments, such as new businesses with condos above, improved parks and recreation facilities, and expanded urban transit. These sizable urban developments also have considerable economic implications and have been documented to draw businesses and new residents that can change neighborhood character, potentially resulting in gentrification and displacement of previous businesses and residents (28). In contrast, environmental changes such as new sidewalks and stop signs or new bicycle lanes and paths may be perceived as smaller investments, potentially driven by safety versus economic concerns and thus less likely to indicate changes in an area's economic opportunities.

Limitations and strengths

This study has some important limitations. One is potential sample selection bias associated with data from a volunteer-based panel survey. Although the sample had nationwide representation, people who agreed to participate may be different than those who did not, which might potentially bias results (29). Sample selection bias could also result from address-based sampling that does not reach homeless or institutionalized populations, which could potentially skew results toward higher-resource populations (29). In addition, data were self-reported and may be affected by recall bias (29). Survey responses may also be affected by social desirability bias, or the inclination to frame behaviors or attitudes in a positive manner, especially regarding physical activity behaviors and support for active living infrastructural changes. Survey questions were also not cognitively tested or psychometrically assessed before administration, so insight is lacking about respondents' interpretation of survey question phrasing, such as "new businesses with condos above." In particular, the survey question about concerns did not specify perceived concerns for the individual but concerns overall, possibly at the neighborhood level, leading to possibility that the individual is supportive despite neighborhood concerns. In addition, respondents may have disagreed with the question about concern if they did not perceive their neighborhood as experiencing development. However, an interpretation focused on cost of living seems more likely, because the question on concern was immediately preceded on the survey by the question on support for changes to make it easier to walk or bike despite higher cost of living. The SummerStyles survey also did not collect data about active or vehicular commutes. This could affect perceptions, especially among those with active commutes who may be more supportive of active transportation infrastructure than others. Lastly, we were unable to differentiate between new residents and long-term residents who may have been aware of previous or planned built environment changes to improve active living, because the survey did not ask the length of time that respondents have lived in their neighborhoods.

This study also has several strengths. Data about perceptions of neighborhood infrastructure to support physical activity with a nationwide sample are rare and a substantial advantage of the survey. Furthermore, no previous study has examined the association between demographic characteristics and perceptions of neighborhood infrastructure to support physical activity. Survey questions parallel those from other widely used built environment assessment tools, such as the Neighborhood Environmental Walkability Scale (30), which facilitates comparison of results across studies. Lastly, the sample size was large, allowing us to examine differences across many different demographic characteristics.

Conclusion

While support for built environment changes to promote active living differed between demographic groups, this study found support for active living infrastructural changes, despite concerns over increased cost of living. It is important to understand community perceptions about the built environment or infrastructural changes to facilitate active living, because community buy-in and meaningful participation are important for implementation (31). Communities that comprise people with low incomes and other traditionally marginalized demographic groups are historically and currently oppressed by government entities, resulting in a lack of trust in engagement opportunities during the development process (13). As a result, community input may not accurately represent the views of these populations. Future research should aim to articulate best practices for equitable community engagement during the development and implementation processes. Solutions such as participatory planning hold potential for community engagement in decisions about active living infrastructure (32). Policy recommendations abound for mitigating displacement risk, though few have been evaluated (33). Public health and planning professionals can partner with communities to select the most appropriate measures for each context and evaluate each carefully for equity and effectiveness. Studies that further explore complexity in resident perceptions of neighborhood improvements to support active living would help communities respond to concerns when planning changes to promote health. This study has implications for engaging residents in decisions by addressing potential barriers to support for transportation and land use changes focused on increasing active living and physical activity.

Acknowledgments

This work is a product of the Physical Activity Policy Research and Evaluation Network (PAPREN). PAPREN is supported by cooperative agreement no. U48DP006381 from the Centers for Disease Control and Prevention and is a product of the Physical Activity Policy Evaluation and Research Network, a thematic network of the Prevention Research Center network. Natalicio Serrano is supported by the National Cancer Institute training grant (no. T32CA057699). The findings and conclusions in this article are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Analyses of these data were exempt from the Centers for Disease Control and Prevention's institutional review board approval because personal identifiers were not included in the data file.

Author Information

Corresponding Author: Nishita Dsouza, PhD, Nesbitt Hall, 3215 Market St NW, Philadelphia, PA 19104. Telephone: 605-310-7902. Email: nd595@drexel.edu.

Author Affiliations: [1]Department of Community Health and Prevention, Dornsife School of Public Health, Drexel University, Philadelphia, Pennsylvania. [2]Institute for Health Research and Policy, School of Public Health, University of Illinois at Chicago, Chicago, Illinois. [3]Division of Nutrition, Physical Activity, and Obesity, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, Georgia. [4]Division of Emergency Operations, Center for Preparedness and Response, Centers for Disease Control and Prevention, Atlanta, Georgia. [5]UMass Worcester Prevention Research Center, Division of Preventive and Behavioral Medicine, Department of Population and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, Massachusetts. [6]Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, Louisiana. [7]Urban Health Collaborative and Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Drexel University, Philadelphia, Pennsylvania.

References

1. US Department of Health and Human Services. Physical activity guidelines for Americans. Second edition. Washington (DC): US Department of Health and Human Services; 2018.
2. Villarroel M, Blackwell D, Jen A. Tables of summary health statistics for US adults: 2018 National Health Interview Survey. National Center for Health Statistics; 2019. Accessed November 6, 2021. https://ftp.cdc.gov/pub/Health_Statistics/NCHS/NHIS/SHS/2018_SHS_Table_A-14.pdf
3. Community Preventive Services Task Force. Physical activity: built environment approaches combining transportation system interventions with land use and environmental design. 2017. Accesssed November 6, 2021. https://www.thecommunityguide.org/sites/default/files/assets/PA-Built-Environments.pdf
4. Hanson S, Cross J, Jones A. Promoting physical activity interventions in communities with poor health and socio-economic profiles: a process evaluation of the implementation of a new walking group scheme. Soc Sci Med 2016;169:77–85.
5. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al.; American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics — 2020 update: a report from the American Heart Association. Circulation 2020;141(9):e139–596.
6. Sallis JF, Slymen DJ, Conway TL, Frank LD, Saelens BE, Cain K, et al. Income disparities in perceived neighborhood built and social environment attributes. Health Place 2011;17(6):1274–83.
7. Rogerson B, Lindberg R, Givens M, Wernham A. A simplified framework for incorporating health into community development initiatives. Health Aff (Millwood) 2014;33(11):1939–47.
8. Durand CP, Andalib M, Dunton GF, Wolch J, Pentz MA. A systematic review of built environment factors related to physical activity and obesity risk: implications for smart growth urban planning. Obes Rev 2011;12(5):e173–82.
9. Litt JS, Reed HL, Tabak RG, Zieff SG, Eyler AA, Lyn R, et al. Active living collaboratives in the United States: understanding characteristics, activities, and achievement of environmental and policy change. Prev Chronic Dis 2013;10:E19.
10. Heidelberg BM. Arts and intra-community strength. 2015. Accesssed November 6, 2021. https://www.americansforthearts.org/sites/default/files/Arts%26America_IntraCommunity.pdf
11. Hoffman ML. Bike lanes are white lanes: bicycle advocacy and urban planning. Lincoln (NE): University of Nebraska Press; 2016.
12. Liu JH, Shi W. Impact of bike facilities on residential property prices. Transp Res Rec 2017;2662(1):50–8.
13. Ferenchak NN, Marshall WE. Bicycling facility inequalities and the causality dilemma with socioeconomic/sociodemographic change. Transp Res Part D Transp Environ 2021;97:102920.
14. Rigolon A, Németh J. Green gentrification or 'just green enough': do park location, size and function affect whether a place gentrifies or not? Urban Stud 2020;57(2):402–20.
15. Aboelata MJ, Bennett R, Yañez E, Bonilla A, Akhavan N. Healthy development without displacement: realizing the vision of healthy communities for all. Oakland (CA): Prevention Institute; 2017.
16. Sullivan DM. Assessing residents' opinions on changes in a gentrifying neighborhood: a case study of the Alberta neighborhood in Portland, Oregon. Hous Policy Debate 2006;17(3):595–624.
17. Porter Novelli. ConsumerStyles and YouthStyles. Washington (DC): Porter Novelli; 2021. Accessed November 6, 2021. https://styles.porternovelli.com/consumer-youthstyles/

18. Porter Novelli. Styles 2018 methods. Washington (DC): Porter Novelli; 2018.
19. US Census Bureau. Metropolitan and micropolitan. 2020. Accessed November 6, 2021. https://www.census.gov/programs-surveys/metro-micro.html
20. Centers for Disease Control and Prevention. National Health Interview Survey. 2017 data release. Washington (DC): National Center for Health Statistics, Centers for Disease Control and Prevention; 2017. Accessed December 4, 2020. https://www.cdc.gov/nchs/nhis/nhis_2017_data_release.htm
21. National Heart, Lung, and Blood Institute, National Institute of Diabetes and Digestive and Kidney Diseases. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence report. Bethesda (MD): National Heart, Lung, and Blood Institute; 1998.
22. Hudson DL, Gehlert S, Pandey S. Tackling the social determinants of ill health. In: Rank M, ed. Toward a livable life: a 21st century agenda for social work. Oxford University Press; 2020:16–45.
23. Parker K, Horowitz JM, Brown A, Fry R, Cohn D, Igielnik R. What unites and divides urban, suburban and rural communities. Washington (DC): Pew Research Center; 2018.
24. Carlson SA, Watson KB, Paul P, Schmid TL, Fulton JE. Understanding the demographic differences in neighborhood walking supports. J Phys Act Health 2017;14(4):253–64.
25. Gustat J, O'Malley K, Hu T, Tabak RG, Goins KV, Valko C, et al.. Support for physical activity policies and perceptions of work and neighborhood environments: variance by BMI and activity status at the county and individual levels. Am J Health Promot 2014;28(3 Suppl):S33–S43.
26. Doucet B. Living through gentrification: subjective experiences of local, non-gentrifying residents in Leith, Edinburgh. J Housing Built Environ 2009;24(3):299–315.
27. Zuk M, Bierbaum AH, Chapple K, Gorska K, Loukaitou-Sideris A. Gentrification, displacement, and the role of public investment. J Plann Lit 2018;33(1):31–44.
28. Dawkins C, Moeckel R. Transit-induced gentrification: who will stay, and who will go? Hous Policy Debate 2016; 26(4–5):801–18.
29. Olson K. Survey participation, nonresponse bias, measurement error bias, and total bias. Public Opin Q 2006;70(5):737–58.
30. Adams MA, Ryan S, Kerr J, Sallis JF, Patrick K, Frank LD, et al. Validation of the Neighborhood Environment Walkability Scale (NEWS) items using geographic information systems. J Phys Act Health 2009;6(Suppl 1):S113–23.
31. Bors P, Dessauer M, Bell R, Wilkerson R, Lee J, Strunk SL. The Active Living by Design national program: community initiatives and lessons learned. Am J Prev Med 2009;37(6 Suppl 2):S313–21.
32. Levine MA. Government policy, the local state, and gentrification: the case of Prenzlauer Berg (Berlin), Germany. J Urban Aff 2004;26(1):89–108.
33. Ghaffari L, Klein JL, Angulo Baudin W. Toward a socially acceptable gentrification: a review of strategies and practices against displacement. Geogr Compass 2018;12(2):e12355.

Tables

Table 1. Characteristics of Analytic Sample of US Adults (N = 3,782), SummerStyles Survey, 2018

Characteristics	No. (%)	Weighted % (95% CI)[a]
Sex		
Men	1,887 (49.9)	48.6 (46.7–50.4)
Women	1,895 (50.1)	51.4 (49.6–53.3)
Age, y		
18–34	690 (18.2)	29.4 (27.5–31.4)
35–49	995 (26.3)	24.1 (22.6–25.6)
50–64	1,242 (32.8)	26.5 (25.0–28.0)
≥65	855 (22.6)	20.1 (18.8–21.4)
Education level		
High school graduate or less	1,239 (32.8)	38.7 (36.9–40.6)
Some college	1,101 (29.1)	28.8 (27.2–30.5)
College graduate or more	1,442 (38.1)	32.4 (30.8–34.1)
Race and ethnicity		
Black, non-Hispanic	318 (8.4)	10.9 (9.8–12.2)
Hispanic/Latino	353 (9.3)	15.6 (14.0–17.2)
Other, non-Hispanic (including multiracial)	281 (7.4)	8.2 (7.1–9.4)
White, non-Hispanic	2,830 (74.8)	65.3 (63.4–67.2)
Annual household income, $		
<50,000	1,171 (31.0)	33.8 (32.0–35.6)
50,000–99,999	1,238 (32.7)	32.6 (30.9–34.4)
≥100,000	1,373 (36.3)	33.6 (32.0–35.3)
Current employment status		
Working	2,390 (63.2)	61.5 (59.7–63.3)
Retired or not working	1,392 (36.8)	38.5 (36.7–40.3)
Housing type		
One family house	3,146 (83.2)	80.5 (78.9–82.0)

[a] Weighted to the total US population as estimated by the 2018 Current Population Survey by sex, age, annual household income, race and ethnicity, household size, education, census region, and MSA status.

[b] Regions are defined as the following: Northeast: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, New Jersey, New York, Pennsylvania, and Vermont; Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; South: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, Maryland, North Carolina, Oklahoma, South Carolina, Virginia, Tennessee, Texas, West Virginia, and District of Columbia; West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

[c] An MSA was categorized as metropolitan if it was associated with at least 1 urbanized area that has a population of at least 50,000.

[d] Respondents were classified into 3 activity levels by using the current adult aerobic guideline (1): 1) active, reporting at least 150 min/week of moderate-intensity equivalent physical activity; 2) insufficiently active, reporting some moderate-intensity equivalent physical activity but not enough to meet active definition; and 3) inactive, reporting no moderate-intensity equivalent physical activity that lasted at least 10 min.

[e] Current smoker: respondents who self-reported having smoked at least 100 cigarettes in their lifetime and currently smoked some days or every day; former smoker: respondents who reported having smoked at least 100 cigarettes in their lifetime, and currently smoked not at all; and never smoker: respondents who reported having smoked fewer than 100 cigarettes in their lifetime.

[f] Calculated as weight in kilograms divided by the square of height in meters. Underweight/normal: <25.0; overweight: 25.0–29.9; and obesity: ≥30.0.

(continued on next page)

(continued)

Table 1. Characteristics of Analytic Sample of US Adults (N = 3,782), SummerStyles Survey, 2018

Characteristics	No. (%)	Weighted % (95% CI)[a]
Apartment or other	636 (16.8)	19.5 (18.0–21.1)
Census region[b]		
Northeast	709 (18.7)	17.9 (16.6–19.4)
Midwest	841 (22.2)	20.9 (19.5–22.4)
South	1,380 (36.5)	37.5 (35.7–39.3)
West	852 (22.5)	23.6 (22.1–25.3)
Metropolitan statistical area (MSA) status[c]		
Nonmetropolitan	3,225 (85.3)	86.0 (84.7–87.2)
Metropolitan	557 (14.7)	14.0 (12.8–15.3)
Aerobic physical activity level[d]		
Inactive	562 (14.9)	15.5 (14.2–16.9)
Insufficiently active	1,079 (28.5)	28.5 (26.9–30.2)
Active	2,141 (56.6)	56.0 (54.2–57.8)
Smoking status[e]		
Current smoker	426 (11.3)	11.6 (10.5–12.8)
Former smoker	1,143 (30.2)	27.1 (25.6–28.6)
Never smoker	2,213 (58.5)	61.3 (59.6–63.1)
Body mass index[f]		
Underweight/normal	1,184 (31.3)	34.0 (32.3–35.9)
Overweight	1,321 (34.9)	32.8 (31.1–34.5)
Obesity	1,277 (33.8)	33.2 (31.5–34.9)
How often respondent avoids busy roads to reduce exposure to air pollution exposure when walking, biking, or exercising outdoors		
Always, usually, sometimes	1,885 (49.8)	49.3 (47.1–51.1)
Rarely, never	1,518 (40.2)	40.0 (38.2–41.8)
Don’t know	379 (10.0)	10.7 (9.6–12.0)

[a] Weighted to the total US population as estimated by the 2018 Current Population Survey by sex, age, annual household income, race and ethnicity, household size, education, census region, and MSA status.

[b] Regions are defined as the following: Northeast: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, New Jersey, New York, Pennsylvania, and Vermont; Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; South: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, Maryland, North Carolina, Oklahoma, South Carolina, Virginia, Tennessee, Texas, West Virginia, and District of Columbia; West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

[c] An MSA was categorized as metropolitian if it was associated with at least 1 urbanized area that has a population of at least 50,000.

[d] Respondents were classified into 3 activity levels by using the current adult aerobic guideline (1): 1) active, reporting at least 150 min/week of moderate-intensity equivalent physical activity; 2) insufficiently active, reporting some moderate-intensity equivalent physical activity but not enough to meet active definition; and 3) inactive, reporting no moderate-intensity equivalent physical activity that lasted at least 10 min.

[e] Current smoker: respondents who self-reported having smoked at least 100 cigarettes in their lifetime and currently smoked some days or every day; former smoker: respondents who reported having smoked at least 100 cigarettes in their lifetime, and currently smoked not at all; and never smoker: respondents who reported having smoked fewer than 100 cigarettes in their lifetime.

[f] Calculated as weight in kilograms divided by the square of height in meters. Underweight/normal: <25.0; overweight: 25.0–29.9; and obesity: ≥30.0.

Table 2. Characteristics of Analytic Sample of US Adults (N = 3,782), by Agreement That Neighborhood Development or Revitalization Has Caused Concerns About Higher Cost of Living, SummerStyles Survey, 2018[a]

Characteristics	Agree, % (95% CI)	Do not agree, % (95% CI)	χ^2 *P* value
Total	19.1 (17.7–20.6)	80.9 (79.4–82.3)	NA
Sex			
Men	19.6 (17.6–21.8)	80.4 (78.2–82.4)	.53
Women	18.6 (16.6–20.8)	81.4 (79.2–83.4)	
Age, y			
18–34	21.9 (18.6–25.6)[b]	78.1 (74.4–81.4)	<.001
35–49	20.7 (18.0–23.7)	79.3 (76.3–82.0)	
50–64	18.5 (16.2–21.0)	81.5 (79.0–83.8)	
≥65	13.8 (11.5–16.6)	86.2 (83.4–88.5)	
Education level			
High school graduate or less	15.7 (13.5–18.3)[b]	84.3 (81.7–86.5)	<.001
Some college	19.9 (17.3–22.8)	80.1 (77.2–82.7)	
College graduate or more	22.4 (20.0–25.0)	77.6 (75.0–80.0)	
Race and ethnicity			
Black, non-Hispanic	21.4 (16.8–26.7)[c,d]	78.6 (73.3–83.2)	<.001
Hispanic/Latino	28.9 (23.9–34.4)[d]	71.1 (65.6–76.1)	
Other, non-Hispanic (including multiracial)	25.5 (20.0–31.9)[d]	74.5 (68.1–80.0)	
White, non-Hispanic	15.6 (14.2–17.1)[c]	84.4 (82.9–85.8)	
Annual household income, $			
<50,000	19.4 (16.8–22.3)	80.6 (77.7–83.2)	.76
50,000–99,999	19.6 (17.1–22.3)	80.4 (77.7–82.9)	
≥100,000	18.4 (16.1–20.8)	81.6 (79.2–83.9)	
Current employment status			
Working	20.6 (18.8–22.6)	79.4 (77.4–81.2)	.01

Abbreviation: NA, not applicable.

[a] Weighted to the total US population as estimated by the annual Current Population Survey by sex, age, annual household income, race and ethnicity, household size, education, census region, and MSA status.

[b] Signifiant linear trend, using orthogonal polynomial contrasts for trends test.

[c, d, e] Values within a column and in the same category that do not share a common superscripted letter are significantly different (Bonferroni corrected *P* < .05), whereas values that do share a common superscripted letter are not significantly different, using pairwise *t* tests.

[f] Regions are defined as the following: Northeast: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, New Jersey, New York, Pennsylvania, and Vermont; Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; South: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, Maryland, North Carolina, Oklahoma, South Carolina, Virginia, Tennessee, Texas, West Virginia, and District of Columbia; West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

[g] An MSA was categorized as metropolitian if it was associated with at least 1 urbanized area that has a population of at least 50,000.

[h] Respondents were classified into 3 activity levels by using the current adult aerobic guideline (1): 1) active, reporting at least 150 min/week of moderate-intensity equivalent physical activity; 2) insufficiently active, reporting some moderate-intensity equivalent physical activity but not enough to meet active definition; and 3) inactive, reporting no moderate-intensity equivalent physical activity that lasted at least 10 min.

[i] Current smoker: respondents who self-reported having smoked at least 100 cigarettes in their lifetime and currently smoked some days or every day; former smoker: respondents who reported having smoked at least 100 cigarettes in their lifetime, and currently smoked not at all; and never smoker: respondents who reported having smoked fewer than 100 cigarettes in their lifetime.

[j] Calculated as weight in kilograms divided by the square of height in meters. Underweight/normal: <25.0; overweight: 25.0–29.9; and obesity: ≥30.0.

(continued on next page)

(continued)

Table 2. Characteristics of Analytic Sample of US Adults (N = 3,782), by Agreement That Neighborhood Development or Revitalization Has Caused Concerns About Higher Cost of Living, SummerStyles Survey, 2018[a]

Characteristics	Agree, % (95% CI)	Do not agree, % (95% CI)	χ^2 *P* value
Retired or not working	16.7 (14.5–19.1)	83.3 (80.9–85.5)	
Housing type			
One family house	17.1 (15.6–18.7)	82.9 (81.3–84.4)	<.001
Apartment or other	27.4 (23.6–31.7)	72.6 (68.3–76.4)	
Region[f]			
Northeast	20.6 (17.2–24.4)[c]	79.4 (75.6–82.8)	<.001
Midwest	12.8 (10.5–15.7)[d]	87.2 (84.3–89.5)	
South	15.7 (13.7–18.0)[c]	84.3 (82.0–86.3)	
West	28.9 (25.4–32.7)[e]	71.1 (67.3–74.6)	
Metropolitan statistical area (MSA) status[g]			
Nonmetropolitan	20.6 (19.0–22.3)	79.4 (77.7–81.0)	<.001
Metropolitan	9.8 (7.3–12.9)	90.2 (87.1–92.7)	
Aerobic physical activity level[h]			
Inactive	14.1 (11.1–17.7)[b]	85.9 (82.3–88.9)	.003
Insufficiently active	18.4 (15.8–21.2)	81.6 (78.8–84.2)	
Active	20.9 (18.9–23.0)	79.1 (77.0–81.1)	
Smoking status[i]			
Current smoker	18.5 (14.7–23.1)	81.5 (76.9–85.3)	.14
Former smoker	17.0 (14.7–19.6)	83.0 (80.4–85.3)	
Never smoker	20.2 (18.2–22.2)	79.8 (77.8–81.8)	
Body mass index[j]			
Underweight/normal	19.4 (16.9–22.2)	80.6 (77.8–83.1)	.89
Overweight	18.6 (16.3–21.2)	81.4 (78.8–83.7)	
Obesity	19.3 (16.8–22.0)	80.7 (78.0–83.2)	
How often respondent avoids busy roads to reduce exposure to air pollution when walking, biking, or exercising outdoors			

Abbreviation: NA, not applicable.

[a] Weighted to the total US population as estimated by the annual Current Population Survey by sex, age, annual household income, race and ethnicity, household size, education, census region, and MSA status.

[b] Signifiant linear trend, using orthogonal polynomial contrasts for trends test.

[c, d, e] Values within a column and in the same category that do not share a common superscripted letter are significantly different (Bonferroni corrected $P < .05$), whereas values that do share a common superscripted letter are not significantly different, using pairwise *t* tests.

[f] Regions are defined as the following: Northeast: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, New Jersey, New York, Pennsylvania, and Vermont; Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; South: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, Maryland, North Carolina, Oklahoma, South Carolina, Virginia, Tennessee, Texas, West Virginia, and District of Columbia; West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

[g] An MSA was categorized as metropolitian if it was associated with at least 1 urbanized area that has a population of at least 50,000.

[h] Respondents were classified into 3 activity levels by using the current adult aerobic guideline (1): 1) active, reporting at least 150 min/week of moderate-intensity equivalent physical activity; 2) insufficiently active, reporting some moderate-intensity equivalent physical activity but not enough to meet active definition; and 3) inactive, reporting no moderate-intensity equivalent physical activity that lasted at least 10 min.

[i] Current smoker: respondents who self-reported having smoked at least 100 cigarettes in their lifetime and currently smoked some days or every day; former smoker: respondents who reported having smoked at least 100 cigarettes in their lifetime, and currently smoked not at all; and never smoker: respondents who reported having smoked fewer than 100 cigarettes in their lifetime.

[j] Calculated as weight in kilograms divided by the square of height in meters. Underweight/normal: <25.0; overweight: 25.0–29.9; and obesity: ≥30.0.

(continued on next page)

(continued)

Table 2. Characteristics of Analytic Sample of US Adults (N = 3,782), by Agreement That Neighborhood Development or Revitalization Has Caused Concerns About Higher Cost of Living, SummerStyles Survey, 2018[a]

Characteristics	Agree, % (95% CI)	Do not agree, % (95% CI)	χ^2 *P* value
Always, usually, sometimes	22.0 (19.9–24.2)[c]	78.0 (75.8–80.1)	<.001
Rarely, never	17.0 (14.9–19.5)[d]	83.0 (80.5–85.1)	
Don't know	13.5 (9.8–18.3)[e]	86.5 (81.7–90.2)	
Support changes to make it easier to walk or bike even if they lead to a higher cost of living			
Supporters	24.6 (22.4–26.9)[c]	75.4 (73.1–77.6)	<.001
Nonsupporters	19.7 (16.3–23.5)[c]	80.3 (76.5–83.7)	
Neither	10.5 (8.6–12.6)[d]	89.5 (87.4–91.4)	

Abbreviation: NA, not applicable.

[a] Weighted to the total US population as estimated by the annual Current Population Survey by sex, age, annual household income, race and ethnicity, household size, education, census region, and MSA status.

[b] Signifiant linear trend, using orthogonal polynomial contrasts for trends test.

[c, d, e] Values within a column and in the same category that do not share a common superscripted letter are significantly different (Bonferroni corrected $P < .05$), whereas values that do share a common superscripted letter are not significantly different, using pairwise *t* tests.

[f] Regions are defined as the following: Northeast: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, New Jersey, New York, Pennsylvania, and Vermont; Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; South: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, Maryland, North Carolina, Oklahoma, South Carolina, Virginia, Tennessee, Texas, West Virginia, and District of Columbia; West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

[g] An MSA was categorized as metropolitian if it was associated with at least 1 urbanized area that has a population of at least 50,000.

[h] Respondents were classified into 3 activity levels by using the current adult aerobic guideline (1): 1) active, reporting at least 150 min/week of moderate-intensity equivalent physical activity; 2) insufficiently active, reporting some moderate-intensity equivalent physical activity but not enough to meet active definition; and 3) inactive, reporting no moderate-intensity equivalent physical activity that lasted at least 10 min.

[i] Current smoker: respondents who self-reported having smoked at least 100 cigarettes in their lifetime and currently smoked some days or every day; former smoker: respondents who reported having smoked at least 100 cigarettes in their lifetime, and currently smoked not at all; and never smoker: respondents who reported having smoked fewer than 100 cigarettes in their lifetime.

[j] Calculated as weight in kilograms divided by the square of height in meters. Underweight/normal: <25.0; overweight: 25.0–29.9; and obesity: ≥30.0.

Table 3. Characteristics of Analytic Sample of US Adults (N = 3,782), by Support for Changes to Make It Easier to Walk or Bike Even if They Lead to a Higher Cost of Living, SummerStyles Survey, 2018[a]

Characteristics	Supporters, % (95% CI)	Nonsupporters, % (95% CI)	Neither, % (95% CI)	χ^2 *P* value
Total	50.7 (48.9–52.6)	16.1 (14.8–17.4)	33.2 (31.5–35.0)	NA
Sex				
Men	50.6 (48.0–53.2)	17.1 (15.2–19.1)	32.3 (29.9–34.8)	.32
Women	50.8 (48.2–53.4)	15.1 (13.4–17.1)	34.0 (31.6–36.6)	
Age, y				
18–34	51.0 (46.7–55.2)[b]	15.0 (12.2–18.2)	34.1 (30.1–38.2)	.01
35–49	54.2 (50.8–57.6)	16.4 (14.1–19.1)	29.4 (26.3–32.6)	
50–64	51.2 (48.2–54.2)	16.7 (14.6–19.0)	32.1 (29.4–35.0)	
≥65	45.5 (42.0–49.0)	16.5 (14.0–19.3)	38.0 (34.6–41.6)	
Education level				
High school graduate or less	40.6 (37.5–43.9)[b]	16.8 (14.5–19.3)[b]	42.6 (39.4–45.8)[b]	<.001
Some college	51.5 (48.2–54.9)	18.2 (15.8–20.9)	30.3 (27.3–33.4)	
College graduate or more	62.1 (59.2–64.8)	13.3 (11.6–15.3)	24.6 (22.2–27.2)	
Race and ethnicity				
Black, non-Hispanic	55.4 (49.3–61.3)	13.5 (9.7–18.4)	31.2 (25.8–37.1)	.03
Hispanic/Latino	56.2 (50.4–61.8)	12.9 (9.5–17.2)	30.9 (25.8–36.5)	
Other, non-Hispanic (including multiracial)	47.3 (40.3–54.4)	13.5 (9.4–19.1)	39.2 (32.3–46.6)	
White, non-Hispanic	49.1 (47.0–51.1)	17.6 (16.1–19.2)	33.4 (31.4–35.3)	
Annual household income, $				
<50,000	42.2 (38.9–45.5)[b]	17.6 (15.2–20.2)	40.3 (37.0–43.6)[b]	<.001
50,000–99,999	52.2 (49.0–55.4)	15.8 (13.6–18.2)	32.0 (29.1–35.1)	
≥100,000	57.9 (54.8–60.8)	14.8 (12.9–17.1)	27.3 (24.6–30.2)	
Current employment status				
Working	55.3 (53.0–57.6)[c]	15.8 (14.3–17.5)	28.8 (26.8–31.0)[c]	<.001

Abbreviation: NA, not applicable.

[a] Weighted to the total US population as estimated by the annual Current Population Survey by sex, age, annual household income, race and ethnicity, household size, education, census region, and MSA status.

[b] Signifiant linear trend, using orthogonal polynomial contrasts for trends test.

[c, d, e] Values within a column and in the same category that do not share a common superscripted letter are significantly different (Bonferroni corrected $P < .05$), whereas values that do share a common superscripted letter are not significantly different, using pairwise *t* tests.

[f] Regions are defined as the following: Northeast: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, New Jersey, New York, Pennsylvania, and Vermont; Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; South: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, Maryland, North Carolina, Oklahoma, South Carolina, Virginia, Tennessee, Texas, West Virginia, and District of Columbia; West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

[g] An MSA was categorized as metropolitian if it was associated with at least 1 urbanized area that has a population of at least 50,000.

[h] Respondents were classified into 3 activity levels by using the current adult aerobic guideline (1): 1) active, reporting at least 150 min/week of moderate-intensity equivalent physical activity; 2) insufficiently active, reporting some moderate-intensity equivalent physical activity but not enough to meet active definition; and 3) inactive, reporting no moderate-intensity equivalent physical activity that lasted at least 10 min.

[i] Current smoker: respondents who self-reported having smoked at least 100 cigarettes in their lifetime and currently smoked some days or every day; former smoker: respondents who reported having smoked at least 100 cigarettes in their lifetime, and currently smoked not at all; and never smoker: respondents who reported having smoked fewer than 100 cigarettes in their lifetime.

[j] Calculated as weight in kilograms divided by the square of height in meters. Underweight/normal: <25.0; overweight: 25.0–29.9; and obesity: ≥30.0.

(continued on next page)

(continued)

Table 3. Characteristics of Analytic Sample of US Adults (N = 3,782), by Support for Changes to Make It Easier to Walk or Bike Even if They Lead to a Higher Cost of Living, SummerStyles Survey, 2018[a]

Characteristics	Supporters, % (95% CI)	Nonsupporters, % (95% CI)	Neither, % (95% CI)	χ^2 *P* value
Retired or not working	43.3 (40.4–46.4)[d]	16.4 (14.3–18.8)	40.2 (37.2–43.3)[d]	
Housing type				
One family house	52.3 (50.3–54.3)[c]	15.7 (14.3–17.2)	32.0 (30.2–34.0)[c]	.005
Apartment or other	44.3 (39.9–48.7)[d]	17.6 (14.4–21.4)	38.1 (33.9–42.5)[d]	
Region[f]				
Northeast	48.9 (44.7–53.1)	16.0 (13.3–19.2)	35.1 (31.0–39.3)	.14
Midwest	48.6 (44.7–52.5)	14.8 (12.4–17.6)	36.6 (32.9–40.5)	
South	53.3 (50.3–56.3)	15.2 (13.2–17.5)	31.5 (28.7–34.4)	
West	49.8 (45.9–53.7)	18.6 (15.7–21.9)	31.6 (28.1–35.3)	
Metropolitan statistical area (MSA)[g]				
Nonmetropolitan	51.9 (49.9–53.9)[c]	15.3 (14.0–16.8)[c]	32.7 (30.9–34.7)	.002
Metropolitan	43.2 (38.7–47.8)[d]	20.7 (17.1–24.7)[d]	36.1 (31.8–40.7)	
Aerobic physical activity level[h]				
Inactive	35.9 (31.5–40.6)[b]	18.4 (15.1–22.3)	45.6 (41.0–50.4)[b]	<.001
Insufficiently active	45.8 (42.4–49.2)	17.6 (15.2–20.2)	36.6 (33.4–40.0)	
Active	57.3 (54.9–59.7)	14.7 (13.0–16.5)	28.0 (25.9–30.3)	
Smoking status[i]				
Current smoker	44.5 (39.2–49.9)[c]	19.7 (15.8–24.3)	35.8 (30.9–41.1)	.02
Former smoker	49.5 (46.3–52.7)[c,d]	17.9 (15.7–20.5)	32.6 (29.7–35.6)	
Never smoker	52.4 (50.0–54.9)[d]	14.6 (13.0–16.4)	33.0 (30.7–35.4)	
Body mass index[j]				
Underweight/normal	51.6 (48.3–54.9)[b]	14.7 (12.5–17.1)	33.7 (30.6–37.0)	.007
Overweight	53.6 (50.5–56.7)	16.9 (14.8–19.3)	29.4 (26.7–32.3)	
Obesity	46.9 (43.8–50.0)	16.7 (14.5–19.1)	36.4 (33.4–39.5)	
How often respondent avoids busy roads to reduce exposure to air pollution when walking, biking, or exercising outdoors				

Abbreviation: NA, not applicable.

[a] Weighted to the total US population as estimated by the annual Current Population Survey by sex, age, annual household income, race and ethnicity, household size, education, census region, and MSA status.

[b] Signifiant linear trend, using orthogonal polynomial contrasts for trends test.

[c, d, e] Values within a column and in the same category that do not share a common superscripted letter are significantly different (Bonferroni corrected *P* < .05), whereas values that do share a common superscripted letter are not significantly different, using pairwise *t* tests.

[f] Regions are defined as the following: Northeast: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, New Jersey, New York, Pennsylvania, and Vermont; Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; South: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, Maryland, North Carolina, Oklahoma, South Carolina, Virginia, Tennessee, Texas, West Virginia, and District of Columbia; West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

[g] An MSA was categorized as metropolitian if it was associated with at least 1 urbanized area that has a population of at least 50,000.

[h] Respondents were classified into 3 activity levels by using the current adult aerobic guideline (1): 1) active, reporting at least 150 min/week of moderate-intensity equivalent physical activity; 2) insufficiently active, reporting some moderate-intensity equivalent physical activity but not enough to meet active definition; and 3) inactive, reporting no moderate-intensity equivalent physical activity that lasted at least 10 min.

[i] Current smoker: respondents who self-reported having smoked at least 100 cigarettes in their lifetime and currently smoked some days or every day; former smoker: respondents who reported having smoked at least 100 cigarettes in their lifetime, and currently smoked not at all; and never smoker: respondents who reported having smoked fewer than 100 cigarettes in their lifetime.

[j] Calculated as weight in kilograms divided by the square of height in meters. Underweight/normal: <25.0; overweight: 25.0–29.9; and obesity: ≥30.0.

(continued on next page)

(continued)

Table 3. Characteristics of Analytic Sample of US Adults (N = 3,782), by Support for Changes to Make It Easier to Walk or Bike Even if They Lead to a Higher Cost of Living, SummerStyles Survey, 2018[a]

Characteristics	Supporters, % (95% CI)	Nonsupporters, % (95% CI)	Neither, % (95% CI)	χ^2 *P* value
Always, usually, sometimes	58.6 (56. 0–61.1)[c]	13.8 (12.1–15.6)[c]	27.6 (25.3–30.1)[c]	<.001
Rarely, never	45.9 (43.0–48.8)[d]	19.1 (17.0–21.5)[d]	35.0 (32.2–37.9)[d]	
Don't know	32.7 (27.5–38.3)[e]	15.2 (11.7–19.6)[c,d]	52.1 (46.4–57.8)[e]	
Agree that neighborhood is experiencing development or revitalization that has caused concerns about higher cost of living				
Agree	65.3 (61.1–69.3)[c]	16.6 (13.7–19.9)	18.2 (15.1–21.8)[c]	<.001
Do not agree	47.3 (45.3–49.3)[d]	16.0 (14.6–17.5)	36.8 (34.8–38.8)[d]	

Abbreviation: NA, not applicable.

[a] Weighted to the total US population as estimated by the annual Current Population Survey by sex, age, annual household income, race and ethnicity, household size, education, census region, and MSA status.

[b] Signifiant linear trend, using orthogonal polynomial contrasts for trends test.

[c, d, e] Values within a column and in the same category that do not share a common superscripted letter are significantly different (Bonferroni corrected $P < .05$), whereas values that do share a common superscripted letter are not significantly different, using pairwise *t* tests.

[f] Regions are defined as the following: Northeast: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, New Jersey, New York, Pennsylvania, and Vermont; Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; South: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, Maryland, North Carolina, Oklahoma, South Carolina, Virginia, Tennessee, Texas, West Virginia, and District of Columbia; West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

[g] An MSA was categorized as metropolitan if it was associated with at least 1 urbanized area that has a population of at least 50,000.

[h] Respondents were classified into 3 activity levels by using the current adult aerobic guideline (1): 1) active, reporting at least 150 min/week of moderate-intensity equivalent physical activity; 2) insufficiently active, reporting some moderate-intensity equivalent physical activity but not enough to meet active definition; and 3) inactive, reporting no moderate-intensity equivalent physical activity that lasted at least 10 min.

[i] Current smoker: respondents who self-reported having smoked at least 100 cigarettes in their lifetime and currently smoked some days or every day; former smoker: respondents who reported having smoked at least 100 cigarettes in their lifetime, and currently smoked not at all; and never smoker: respondents who reported having smoked fewer than 100 cigarettes in their lifetime.

[j] Calculated as weight in kilograms divided by the square of height in meters. Underweight/normal: <25.0; overweight: 25.0–29.9; and obesity: ≥30.0.

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
Volume 19, E75 NOVEMBER 2022

ORIGINAL RESEARCH

Creation and Evaluation of the Illinois Cancer Risk Index as a Predictor of Four Common Cancers

Lei Guo, MA, MS[1]; Margaret E. Wright, PhD[2]; Meredith C. Osias, BA[3]; Mahdi Vaezi, PhD[4]; M. Courtney Hughes, PhD, MS[3]

Accessible Version: www.cdc.gov/pcd/issues/2022/22_0104.htm

Suggested citation for this article: Guo L, Wright ME, Osias MC, Vaezi M, Hughes MC. Creation and Evaluation of the Illinois Cancer Risk Index as a Predictor of Four Common Cancers. Prev Chronic Dis 2022;19:220104. DOI: https://doi.org/10.5888/pcd19.220104.

PEER REVIEWED

Summary

What is already known on this topic?

Information is needed on how social and environmental determinants of health affect outcomes of common cancers to create a measure and identify geographic locations that are most in need of public health intervention.

What is added by this report?

A risk index representing 25 predictors of death from the 4 most common cancers was created by using population-based, county-level data in Illinois. We correlated the index with mortality rates from the 4 most common cancers, and both exhibited similar geospatial distribution across the state.

What are the implications for public health practice?

Professionals in many health fields can adapt our framework to construct indexes and inform public health resource allocations.

Abstract

Introduction

Nearly half of all cancer deaths in the US are attributed to 4 common cancers: lung, colorectal, breast, and prostate. Illinois residents experience higher rates of cancer death from all 4 cancers compared with the US overall. We developed the Illinois Cancer Risk Index (ICRI), which incorporates many predictors of these cancers into a single summary measure, to identify Illinois counties that would benefit most from public health intervention.

Methods

We identified 90 county-level predictors of 4 common cancers, used multicollinearity testing to reduce this number to 61, and applied factor analysis to extract and analyze 4 factors representing 25 variables. Next, we created the ICRI by regressing the 4 factors on our outcome of interest — an age-adjusted common cancers mortality rate (CCMR), incorporating the direction of the β-coefficients from regression models to sum factor scores. Finally, we mapped and assessed the geographic distributions of both ICRI and CCMR by county across the state.

Results

The ICRI was positively associated with the CCMR ($r = 0.59$, $P < .001$) and explained 32.2% of the variance in the CCMR across Illinois. The ICRI showed distinct geospatial patterns across the state, with the highest risk counties located in the east–central, far northern, and southern regions. The CCMR showed similar geospatial patterns.

Conclusion

Our study identifies counties in Illinois that may benefit most from interventions that target multiple cancer risk factors simultaneously. The ICRI may be adapted for use in other geographic locations where data are available.

Introduction

Cancer is the second leading cause of death in the US, with lung cancer accounting for almost one-quarter (23%) of cancer deaths. Other common sites of cancer death are colorectal (9%), female breast (7%), and prostate (5%) (1). Illinois residents experience higher rates of death from all 4 of these cancers compared with the US overall (1). In Illinois, approximately 14,140 deaths from these 4 common cancers are expected in 2022, with wide variation across the state (2).

Determinants of cancer incidence include those related to demographics, social and economic factors, health behaviors, the physical environment, and clinical care (3). For example, Singh et al (4) showed that cancer outcomes in the US correlate with socioeconomic status and race and ethnicity at both individual and population levels. Other researchers have demonstrated that health behaviors and cumulative environmental quality influence cancer risk (5–7). Krieger (8) proposed an ecosocial theory and discussed how social and biologic reasoning and dynamic and ecologic perspectives could affect distributions and result in social inequalities in cancer outcomes. Most cancer risk factors have been examined in isolation with respect to outcomes. Although some indexes have been constructed and used, their data were older (9,10) than ours, focused on just one cancer type (11), or did not include specific risk factors such as individual health behaviors, air pollution, English proficiency, and poor physical and mental health (9–11). Our study included these additional elements of risk and evaluated the index's relation to an outcome that reflects the 4 most common causes of cancer death in Illinois.

This article introduces the Illinois Cancer Risk Index (ICRI), which incorporates predictors of the most common cancer deaths into a single summary measure. The index can be used to identify counties in Illinois where public health intervention is needed most.

Methods

We identified 90 variables related to the following domains of risk: demographics, social and economic factors, health behaviors, the physical environment, and clinical care. We selected these on the basis of published literature (3), ecosocial theory (8), and available Illinois county data. For each variable, the most recent county-level data (2014–2018) were extracted from publicly available sources, including the Area Health Resources Files (12), County Health Rankings and Roadmaps (CHR) (13), the US Department of Agriculture Food Environmental Atlas (14), and the Illinois Environmental Protection Agency Bureau of Air (15).

We normalized each of the 90 variables and expressed them as either per-capita values or percentages by using the following formula:

$$\text{Variable}_{\text{normalized}} = (\text{Variable} - \text{Variable}_{\text{minimal}}) / (\text{Variable}_{\text{maximal}} - \text{Variable}_{\text{minimal}})$$

Data were available for more than 94% of our study variables for all 102 Illinois counties. We replaced missing data for the risk factors (1.1% of data points missing) and site-specific cancer mortality rates (30.4% of data points missing) by using the hot deck imputation method (16). Subsequently, we standardized data for each variable by applying a *z* score standardization to transform the different variables into comparable scales (16). We then conducted redundancy and multicollinearity tests by using variance inflation factors (VIFs) to remove highly correlated variables (VIF >5) and improve the efficiency of factor analysis. Sixty-one variables were retained for subsequent exploratory factor analysis (16).

Exploratory factor analysis was used to extract 13 grouped variables called "factors" from these 61 variables. Analysis was initiated by estimating the variance component with principal component analysis. We used the Kaiser Measure of Sampling Adequacy (Kaiser MSA) in combination with variable communalities (proportion of each variable's variance that the factors explain) to extract these factors. We used the Bartlett χ^2 test to validate the estimated factors (all *P* values < .001). Factors containing 2 or fewer variables or those whose ascribed variables all had factor loadings less than 0.5 were excluded. Four factors representing 25 variables were retained for analysis (Table 1).

Our outcome of interest was the average age-adjusted mortality rate from lung, colorectal, breast, and prostate cancers, which we refer to as the common cancer mortality rate (CCMR). Cancer mortality data were obtained from the Illinois State Cancer Registry and included the years 2014–2018 (15). We replaced missing data with hot-deck imputation for counties that had low counts because of suppressed data.

We used multiple linear regression to assess the bivariate and multivariate relationships between each of the 4 retained factors. We then used this regression analysis information to construct the ICRI. In brief, the direction of each association (the sign of the β-coefficient) was incorporated into the calculation of the ICRI. For each Illinois county, the ICRI was calculated by multiplying the standardized value of each variable by its respective factor loading and then summing all 4 factors together while accounting for the sign of the β-coefficient from regression models. We used the following equation:

$$\text{ICRI} = \text{Factor Score 1} + \text{Factor Score 3} + \text{Factor Score 5} + \text{Factor Score 6}$$

For example, we calculated the ICRI for Cumberland County as

$$\text{ICRI}_{\text{Cumberland County}} = \text{Factor Score 1}_{\text{Cumberland County}} + \text{Factor Score 3}_{\text{Cumberland County}} + \text{Factor Score 5}_{\text{Cumberland County}} + \text{Factor Score 6}_{\text{Cumberland County}} = (-0.49) + (-1.91) + (-3.92) + (-3.05) = -9.37$$

Sensitivity analyses were performed in which the ICRI was regressed individually against breast, colorectal, lung, and prostate cancer mortality rates.

We generated and comparted maps of the ICRI and CCMR by county. We divided county-level CCMRs into quintiles for mapping. For the ICRI map, counties were classified for risk as very low, low, average, high, and very high. Low- and high-risk counties had index values that were more than 0.5 SD but less than 1.5 standard deviations from the mean. Very low-risk and very high-risk counties had an index more than 1.5 standard deviations from the mean. Average risk counties had an ICRI value within 0.5 standard deviations from the mean.

We used SPSS version 27.0 (IBM Corp) to conduct all analyses and constructed maps in ArcGIS Pro 2.8 (Esri Corp).

Results

The communalities of variables (the proportion of common variance found in a particular variable) were all higher than 0.5. All variables correlated with at least 1 other variable (correlation coefficient of at least 0.3), indicating that variables all shared some common variance. We retained 4 factors representing 25 variables because, together, they explained 70.1% of the total variance of the 61 originally identified, noncollinear variables across Illinois. These factors were calculated, along with the variables that each one comprises and their factor loadings (Table 1). Factor 1 included Black race and health behaviors (the variable with the highest factor loading was Black race at 0.931) and explained 27.3% of the total variance. Factor 3 included Hispanic ethnicity, rurality, measures of air pollution, and language barriers (the variable with the highest factor loading was the percentage of non-Hispanic White race at −0.787) and explained 17.8% of the total variance. Factor 5 included aspects of financial security (the variable with the highest factor loading was percentage of people living in poverty at 0.820) and explained 16.7% of the total variance. Factor 6 included density of primary care providers, education, and income ratio (the variable with the highest factor loading was the primary care provider [PCP] rate at 0.553). PCP rate is the ratio of population to primary care physicians and includes practicing physicians specializing in general practice medicine, family medicine, internal medicine, and pediatrics, and explained 8.3% of the total variance.

Regression analysis and index construction

In bivariate regression, Factor 5 was the only significant predictor of the CCMR; Factors 1, 3, and 6 and the ICRI were not. However, in multivariable regression, all 4 retained factors significantly predicted the CCMR. Together, they explained 33.5% of the variance in the county-level CCMR across the state of Illinois (Table 2).

The ICRI had a moderate positive association with the CCMR ($r = 0.59$, $P < .001$) and explained 32.2% of the variance in this outcome (Table 2). The ICRI had a mean of 0 and an SD of 4.6. Winnebago County in the northern portion of the state had the highest index value at 10.76, implying a high risk of mortality from lung, breast, prostate, and colorectal cancers in this area. Cumberland County in the east had the lowest value at −9.37, implying a low risk of mortality from these 4 cancers in that area.

We performed sensitivity analyses by regressing the ICRI against individual cancer mortality rates rather than the CCMR (Table 3). The ICRI explained the largest proportion of variation in prostate cancer mortality (64%), followed by deaths from breast (57%), colorectal (56%), and lung (23%) cancer.

Maps

Thirty-three Illinois counties were classified as at high or very high risk for the 4 common cancers (Figure) based on the ICRI, and 32 counties were classified as low or very low risk. Higher risk counties were primarily located in the east–central portion of the state, with several counties in the far northern and southern portions of the state also classified as such. Counties in the 2 highest CCMR quintiles were located predominantly in the northeast and southern parts of Illinois, with some counties in the central portion.

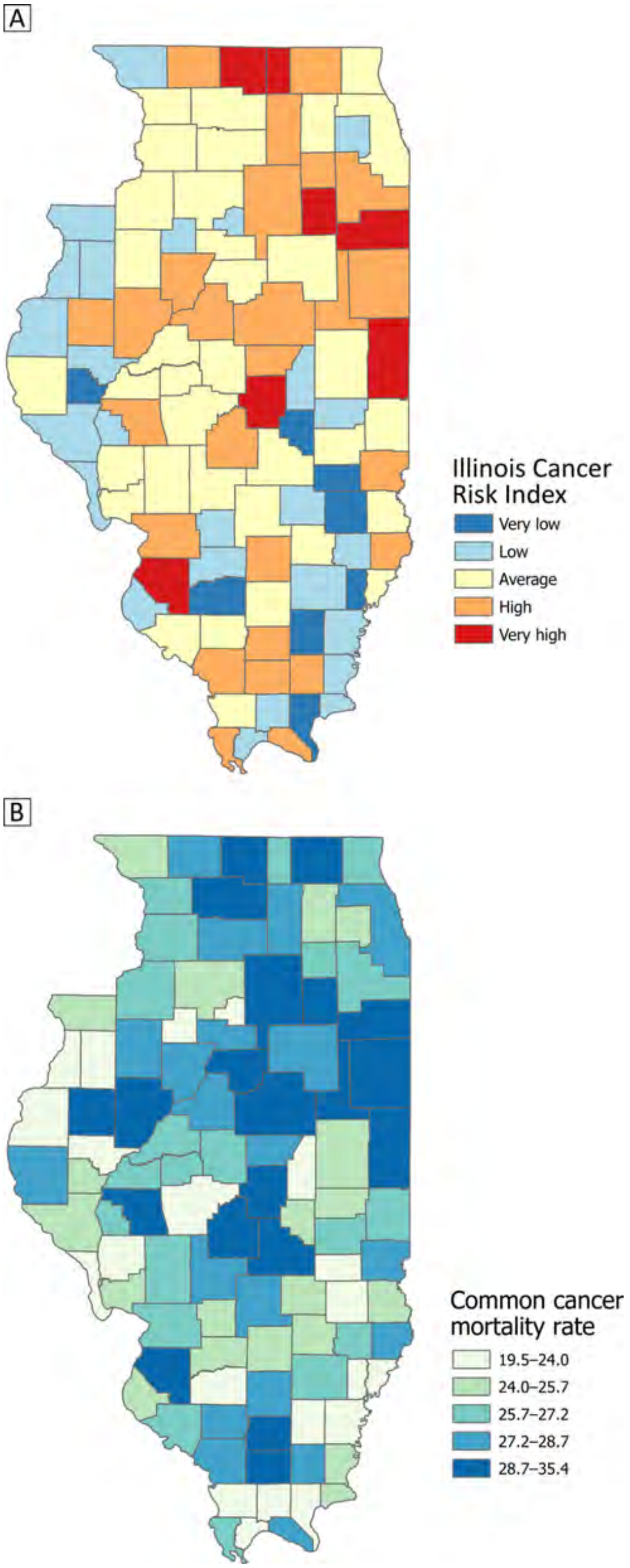


Figure 1. Distribution by county of risk of the 4 most common cancers — lung, colorectal, breast, and prostate — in Illinois. Map A displays risk by the Illinois Cancer Risk Index (ICRI) for each county. Higher risk counties were located in the east-central portion of the state, with some also located in the far northern and southern portions of the state. Map B plots the common cancer mortality rate (CCMR), 2014–2018, for each county in Illinois. Counties in the 2 highest CCMR quintiles were located in the northeast and southern parts of Illinois, with some also located in the central portion of the state.

Discussion

We constructed a novel cancer risk index — ICRI — by using population-based, county-level data from the state of Illinois. The ICRI represented a broad range of determinants of the 4 most common cancers in both Illinois and the US. To the best of our knowledge, our study incorporates one of the largest numbers of cancer risk factors (based on ecosocial theory [8]) to date and has important implications for screening, intervention resource allocation, and access to cancer care.

Our study differs from other reports that also describe cancer risk indexes. Although Scott et al (9) included factors from several domains, our study also examined air pollution, English proficiency, and poor physical and mental health. Wang et al (10) used data from 1998–2000 to focus exclusively on late-stage cancers and did not include rurality, air pollution, English proficiency, poor physical and mental health, or health behaviors such as alcohol consumption, smoking, and diet. In a separate publication, Wang et al (11) used breast cancer mortality throughout Illinois as the outcome of interest and did not include rurality, air pollution, English proficiency, poor physical and mental health, health behaviors, or ratio of population to primary care physicians. Overall, our study evaluated more domains than previous studies and created a unique outcome by averaging the 4 most common causes of cancer death in Illinois.

Demographic variables had high loadings in 2 factors included in our index. For the first factor, the percentage of the population that identified as non-Hispanic Black had the highest factor loading. For the second factor, the percentage of the population that identified as non-Hispanic White had the highest factor loading. Together, these 2 factors explained 45.1% of the total variance of the 61 variables across Illinois. These variables exhibited the connections between social disparities and cancer risks across Illinois counties, showing patterns similar to larger-scale studies. For instance, previously published research found substantially elevated cancer mortality rates at multiple sites among non-Hispanic Black populations compared with non-Hispanic White populations (8). On the other hand, Hispanic populations had lower cancer mortality rates than non-Hispanic White or non-Hispanic Black populations, although Hispanic populations tend to have later-stage diagnoses and poorer quality of life at the national level and across Illinois (1).

In our index, Hispanic ethnicity is loaded passively with air pollution variables. This association may indicate the disparities between racial and ethnic minorities and environmental pollution (7). A recent study showed that Hispanic, Latino, and other minority populations were being exposed to higher levels of dangerous fine particulate matter 2.5 ($PM_{2.5}$) from air pollution than other groups; previous studies also found that the Hispanic population is at higher risk of premature death from exposure to $PM_{2.5}$ air pollution (17), which echo our results. Next, because about 14% of the Illinois population are non–US-born (18), it is important to consider the additional hurdles encountered by this group. Limited English proficiency can burden non–US-born people when they attempt to access health care in the US, although they often have lower cancer mortality rates than people born in the US. Over time, their cancer rates tend to equal or even exceed those of people born in the US as they acculturate, and this tendency is why English proficiency was loaded in our index.

Factor 1 included health behavior–related variables and poor physical and mental health days. Multiple studies have shown that health behaviors are strongly associated with cancer outcomes. For example, smoking and alcohol consumption are known risk factors for multiple cancers (19,20), with smoking the most predominant risk factor for lung cancer (3). Additionally, previous research demonstrated that people who develop mental disorders after a cancer diagnosis may be at higher risk of cancer death (21). Furthermore, mental health treatment offered to cancer patients after diagnosis can improve lung cancer survival, and reductions in the severity of mental illness may manifest in greater self-efficacy for managing chronic conditions and improvements in positive health behaviors, such as physical activity and stress management (22). However, past studies using Illinois data (10,11) and US data (9) did not consider these variables to create their single or multiple cancer sites index.

Living in a rural area — which had a high factor loading in our index — is also associated with higher cancer mortality rates because of limited health care access (23). Another variable, severe housing problems, was loaded in the same factor with rural area. Although housing is not frequently examined with cancer outcomes, studies have found that it is associated with increased cancer mortality disparities between Black and White populations (9,23). Rural residency and the health behavior variables loaded in our factors might represent latent variables related to protective behaviors associated with cancer outcomes.

Household income, median home value, and median gross rent were loaded negatively in Factor 5, echoing the findings from other research that low- and middle-income counties have higher cancer mortality rates than high-income counties in the US (24). On the other hand, the variables “some college,” “PCP rate,” and “income ratio” were positively associated in Factor 6. These results are similar to some studies that found that education, primary care access, and income might have essential roles in cancer treatment and prevention (25,26). For example, use of breast and colorectal cancer screening is 20% to 30% lower among those with only a high school education than among college graduates (25). Also, a greater primary care physician supply was associated with lower cancer mortality (26), and low income is a barrier to health care access (25).

Although the ICRI explained only 32.2% of the variance in the age-adjusted CCMR, the combinations of the multiple cancer risk variables demonstrate the nature of cancer as a heterogeneous disease with many risk factors that may have a long-term impact on health (8). Although resource allocation can affect all the variables analyzed in this article, some variables, such as smoking and alcohol consumption, are also considered modifiable cancer risk factors. In addition, lifestyle modification could significantly reduce the burden of cancer (27); thus, public health professionals in Illinois may use the index to direct risk reduction and health promotion programs and policies at those counties most in need.

Our index explained the largest proportion of variance in prostate cancer mortality rates across Illinois and the smallest proportion for lung cancer. Many variables used in our study may relate to prostate cancer mortality in Illinois compared with other cancer types because early screening often can reduce prostate cancer risk for highly educated people with cancer screening resources (28). The low prediction of lung cancer mortality might be due to the noise introduced by smoking as the strongest risk factor for lung cancer (3). It also indicates that underlying latent sociodemographic variables other than smoking and air pollution could have an impact on lung cancer mortality in Illinois. Future studies could explore more variables specifically related to lung cancer mortality risk. Nevertheless, our index map showed a similar geospatial pattern that matches the county-level CCMR. Both maps demonstrate the compelling need for cancer-related public health resource allocation in east-central and northern Illinois.

Previous Illinois studies examined the associations between several risk factors and late-stage diagnosis for 4 common cancers and created a county-level index for breast cancer (10,11). Our study differs in that we comprehensively focused on ecosocial determinants of cancer risk factors according to ecosocial theory (8). Many variables used to construct our index are also associated with other cancer outcomes and other chronic noncommunicable diseases (27,29). This framework could potentially be used to create a similar scoring system for public health professionals.

Strength and limitations

Our study had several strengths. First, it incorporated numerous cancer risk factors at population levels to create a cancer risk index for Illinois. Our results identified physical and mental health variables and air toxin variables that previous studies did not include. Additionally, the use of a factor analysis framework can explore the underlying trends in the data, increase interpretability, and minimize information loss while reducing dimensionality. Furthermore, we identified and reduced factors from large feature sets associated with common cancer mortality in Illinois, and this information can assist in cancer intervention and prevention program planning at the county level.

Our study had some limitations. First, we only examined the state of Illinois. Further studies should explore larger data sets with counties from multiple US states and abroad. Second, we needed to impute data for some counties because of missing data; this was to ensure that an ICRI value was ascribed to each county. Third, we relied on data that were self-reported and aggregated at the county level, which may obscure nuances in individual behavior and be susceptible to social desirability bias. Fourth, caution must be applied in interpreting the index because we evaluated the ICRI by using its correlation with CCMR, and correlation does not signify causation. Despite these limitations, the calculated index provides informative data to advise public health professionals. To address these limitations in future studies, researchers can use larger and validated data sets and machine learning frameworks, which are becoming increasingly prevalent in cancer research, to model risk factors (30).

Conclusion

This study identified, reduced, and analyzed a substantial number of cancer predictors and incorporated them into a single novel county-level index for the state of Illinois. Our analysis found that the ICRI was moderately associated with the CCMR, which is the average mortality from the 4 most common cancers in the state. Public health professionals may use this framework to target resources and interventions to counties in Illinois that score highest on the risk index and are, therefore, most in need. Future research should apply this framework to construct indexes for other diseases and for multiple geographic locations.

Acknowledgments

This study is supported by an Illinois Innovation Network Seed Grant (2021). No copyrighted materials were used in the conduct of this research or this article.

Author Information

Corresponding Author: Lei Guo, MA, MS, School of Interdisciplinary Health Professions, Northern Illinois University, Wirtz Hall 323, DeKalb, IL 60115. Telephone: 734-358-3970. Email: rguo@niu.edu.

Author Affiliations: [1]School of Interdisciplinary Health Professions, Northern Illinois University, DeKalb, Illinois. [2]University of Illinois Cancer Center, Chicago, Illinois. [3]School of Health Studies, Northern Illinois University, DeKalb, Illinois. [4]Department of Engineering Technology, Northern Illinois University, DeKalb, Illinois.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin 2022;72(1):7–33.
2. 2022–2027 Illinois comprehensive cancer control plan. Springfield (IL): Illinois Department of Public Health. Accessed August 30, 2022. https://dph.illinois.gov/content/dam/soi/en/web/idph/publications/idph/topics-and-services/diseases-and-conditions/cancer/2022-2027_IL-Comp-Cancer-Plan_03092022.pdf
3. Thun MJ, Linet MS, Cerhan JR, Haiman CA, Schottenfeld D, editors. Cancer epidemiology and prevention. Fourth edition. New York (NY): Oxford University Press; 2018.
4. Singh GK, Jemal A. Socioeconomic and racial/ethnic disparities in cancer mortality, incidence, and survival in the United States, 1950–2014: over six decades of changing patterns and widening inequalities. J Environ Public Health 2017;2017:2819372.
5. Hastert TA, Ruterbusch JJ, Beresford SA, Sheppard L, White E. Contribution of health behaviors to the association between area-level socioeconomic status and cancer mortality. Soc Sci Med 2016;148:52–8.
6. Jagai JS, Messer LC, Rappazzo KM, Gray CL, Grabich SC, Lobdell DT. County-level cumulative environmental quality associated with cancer incidence. Cancer 2017;123(15):2901–8. Erratum in Cancer 2019;125(10):1756.
7. James W, Jia C, Kedia S. Uneven magnitude of disparities in cancer risks from air toxics. Int J Environ Res Public Health 2012;9(12):4365–85.
8. Krieger N. Defining and investigating social disparities in cancer: critical issues. Cancer Causes Control 2005;16(1):5–14.
9. Scott LC, Barker LE, Richardson LC. Using population health measures to evaluate the environmental burden of cancer at the county level. Prev Chronic Dis 2019;16:E45.

10. Wang F, Luo L, McLafferty S. Healthcare access, socioeconomic factors and late-stage cancer diagnosis: an exploratory spatial analysis and public policy implication. Int J Public Policy 2010;5(2-3):237–58.
11. Wang F, Guo D, McLafferty S. Constructing geographic areas for cancer data analysis: a case study on late-stage breast cancer risk in Illinois. Appl Geogr 2012;35(1–2):1–11.
12. US Department of Health and Human Services. Area health resource files. Accessed November 14, 2021. https://data.hrsa.gov/topics/health-workforce/ahrf
13. County Health Rankings & Roadmaps. Accessed November 14, 2021. https://www.countyhealthrankings.org/
14. US Department of Agriculture. Food environment atlas. Accessed November 14, 2021. https://www.ers.usda.gov/data-products/food-environment-atlas/
15. Illinois Environmental Protection Agency Bureau of Air. Accessed November 14, 2021. https://www2.illinois.gov/epa/topics/air-quality/Pages/default.aspx
16. Finch WH. Exploratory factor analysis. In: Teo T, editor. Handbook of quantitative methods for educational research. Rotterdam (NL): SensePublishers; 2013.
17. Jbaily A, Zhou X, Liu J, Lee TH, Kamareddine L, Verguet S, et al. Air pollution exposure disparities across US population and income groups. Nature 2022;601(7892):228–33.
18. US Department of Commerce. American community survey; 2018. Accessed November 4, 2021. https://www.census.gov/programs-surveys/acs
19. Siegel RL, Jacobs EJ, Newton CC, Feskanich D, Freedman ND, Prentice RL, et al. Deaths due to cigarette smoking for 12 smoking-related cancers in the United States. JAMA Intern Med 2015;175(9):1574–6.
20. Nelson DE, Jarman DW, Rehm J, Greenfield TK, Rey G, Kerr WC, et al. Alcohol-attributable cancer deaths and years of potential life lost in the United States. Am J Public Health 2013;103(4):641–8.
21. Zhu J, Fang F, Sjölander A, Fall K, Adami HO, Valdimarsdóttir U. First-onset mental disorders after cancer diagnosis and cancer-specific mortality: a nationwide cohort study. Ann Oncol 2017;28(8):1964–9.
22. Berchuck JE, Meyer CS, Zhang N, Berchuck CM, Trivedi NN, Cohen B, et al. Association of mental health treatment with outcomes for US veterans diagnosed with non-small cell lung cancer. JAMA Oncol 2020;6(7):1055–62.
23. Blake KD, Moss JL, Gaysynsky A, Srinivasan S, Croyle RT. Making the case for investment in rural cancer control: an analysis of rural cancer incidence, mortality, and funding trends. Cancer Epidemiol Biomarkers Prev 2017;26(7):992–7.
24. O’Connor JM, Sedghi T, Dhodapkar M, Kane MJ, Gross CP. Factors associated with cancer disparities among low-, medium-, and high-income US counties. JAMA Netw Open 2018;1(6):e183146.
25. Goding Sauer A, Siegel RL, Jemal A, Fedewa SA. Current prevalence of major cancer risk factors and screening test use in the United States: disparities by education and race/ethnicity. Cancer Epidemiol Biomarkers Prev 2019;28(4):629–42.
26. Basu S, Berkowitz SA, Phillips RL, Bitton A, Landon BE, Phillips RS. Association of primary care physician supply with population mortality in the United States, 2005–2015. JAMA Intern Med 2019;179(4):506–14.
27. Stein CJ, Colditz GA. Modifiable risk factors for cancer. Br J Cancer 2004;90(2):299–303.
28. Kilpeläinen TP, Talala K, Taari K, Raitanen J, Kujala P, Pylväläinen J, et al. Patients’ education level and treatment modality for prostate cancer in the Finnish Randomized Study of Screening for Prostate Cancer. Eur J Cancer 2020;130:204–10.
29. Rojas-Rueda D, Morales-Zamora E, Alsufyani WA, Herbst CH, AlBalawi SM, Alsukait R, et al. Environmental risk factors and health: an umbrella review of meta-analyses. Int J Environ Res Public Health 2021;18(2):704.
30. Abdullah Alfayez A, Kunz H, Grace Lai A. Predicting the risk of cancer in adults using supervised machine learning: a scoping review. BMJ Open 2021;11(9):e047755.

Tables

Table 1. Factor Loadings[a] and Sources of Data for Variables, the Illinois Cancer Risk Index, 2014–2018

Variable	Factor name[b]					
	1. Black race and health behaviors	3. Ethnicity, air quality, housing, and rurality	5. Financial needs and unemployment	6. Education, primary care provider, and income ratio	Communalities[c]	Source
Non-Hispanic Black population, 2014–2018, %	0.93[d]	—	—	—	0.96	County Health Rankings (13)
Non-Hispanic White population, 2014–2018, %[b]	—	−0.79	—	—	0.95	
Hispanic population, 2014–2018, %	—	0.79	—	—	0.90	
Rural population, 2010, %	—	0.75	—	—	0.81	
Current smoker[d], %	0.88	—	—	—	0.94	
Excessive drinking[e], %	0.85	—	—	—	0.90	
Poor physical health days[f], mean no.	0.79	—	—	—	0.85	
Poor mental health days[g], mean no.	0.62	—	—	—	0.89	
Limited access to healthy foods, 2014–2018, %	0.53	—	—	—	0.86	
Severe housing problems[h], %	—	0.73	—	—	0.86	
Carbon monoxide[i]	—	0.72	—	—	0.97	Illinois Environmental Protection Agency Bureau of Air (15)
Nitrogen oxides[j]	—	0.72	—	—	0.97	
Average daily $PM_{2.5}$[k]	—	0.72	—	—	0.65	County Health Rankings (13)
Sulfur dioxide[l]	—	0.72	—	—	0.97	Illinois Environmental Protection Agency Bureau of Air (15)
Volatile organic material[m]	—	0.72	—	—	0.97	
Population not proficient in English, %	—	0.69	—	—	0.90	County Health Rankings (13)
Population living in poverty, 2014–2018, %	—	—	0.82	—	0.93	Area Health Resources Files (12)

Abbreviations: — , variable not included in the model.

[a] Factor loadings: correlation coefficients between observed variables and common latent factors.

[b] Factor is a latent variable associated with a set of observed variables that have similar response patterns.

[c] The correlation coeffect *r* and *P* value of bivariate linear regression between factor and common cancers mortality rate.

[d] Percentage of adults who smoked 100 cigarettes in their lifetime and currently smoked, 2014–2018.

[e] Percentage of adults reporting binge or heavy drinking, 2014–2018. Binge drinking = consuming 4 or more drinks on one occasion for a woman or 5 or more drinks on one occasion for a man. Heavy drinking = 8 or more drinks per week for a woman or 15 or more drinks per week for a man.

[f] Number of physically unhealthy days in past 30 days (age-adjusted), 2014–2018.

[g] Number of mentally unhealthy days in past 30 days (age-adjusted), 2014–2018.

[h] Percentage of households with at least 1 of 4 housing problems: overcrowding, high housing costs, lack of kitchen, or lack of plumbing facilities, 2014–2018.

[i] Carbon monoxide stationary point source emission distribution at county level (tons/y), 2014–2018.

[j] Nitrogen oxides stationary point source emission distribution at county level (tons/y), 2014–2018.

[k] Average daily density of fine particulate matter in micrograms per cubic meter ($PM_{2.5}$), 2014–2018.

[l] Sulfur dioxide stationary point source emission distribution (tons/y), 2014–2018.

[m] Volatile organic material point source emission distribution (tons/y), 2014–2018.

[n] Adults aged 25–44 years with some post-secondary education, 2014–2018.

[o] Ratio of household income at the 80th percentile to income at the 20th percentile at county level, 2014–2018.

(continued on next page)

(continued)

Table 1. Factor Loadings[a] and Sources of Data for Variables, the Illinois Cancer Risk Index, 2014–2018

Variable	Factor name[b]					
	1. Black race and health behaviors	3. Ethnicity, air quality, housing, and rurality	5. Financial needs and unemployment	6. Education, primary care provider, and income ratio	Communalities[c]	Source
Median household income, 2014–2018	—	—	−0.81	—	0.95	County Health Rankings (13)
Median home value, 2014–2018	—	—	−0.66	—	0.94	Area Health Resources Files (12)
Median gross rent, 2014–2018	—	—	−0.60	—	0.89	
Unemployment rate, 2014–2018	—	—	0.56	—	0.72	USDA Food Environmental Atlas (14)
Some college[n]	—	—	—	−0.51	0.78	County Health Rankings (13)
Ratio of population to primary care physicians	—	—	—	0.55	0.7	
Income ratio[o]	—	—	—	0.54	0.743	
Correlation coeffect r and P value	$r = 0.37$, $P < .001$	$r = 0.36$, $P < .001$	$r = 0.04$, $P = 0.03$	$r = 0.26$, $P < .001$	—	—

Abbreviations: — , variable not included in the model.

[a] Factor loadings: correlation coefficients between observed variables and common latent factors.

[b] Factor is a latent variable associated with a set of observed variables that have similar response patterns.

[c] The correlation coeffect r and P value of bivariate linear regression between factor and common cancers mortality rate.

[d] Percentage of adults who smoked 100 cigarettes in their lifetime and currently smoked, 2014–2018.

[e] Percentage of adults reporting binge or heavy drinking, 2014–2018. Binge drinking = consuming 4 or more drinks on one occasion for a woman or 5 or more drinks on one occasion for a man. Heavy drinking = 8 or more drinks per week for a woman or 15 or more drinks per week for a man.

[f] Number of physically unhealthy days in past 30 days (age-adjusted), 2014–2018.

[g] Number of mentally unhealthy days in past 30 days (age-adjusted), 2014–2018.

[h] Percentage of households with at least 1 of 4 housing problems: overcrowding, high housing costs, lack of kitchen, or lack of plumbing facilities, 2014–2018.

[i] Carbon monoxide stationary point source emission distribution at county level (tons/y), 2014–2018.

[j] Nitrogen oxides stationary point source emission distribution at county level (tons/y), 2014–2018.

[k] Average daily density of fine particulate matter in micrograms per cubic meter ($PM_{2.5}$), 2014–2018.

[l] Sulfur dioxide stationary point source emission distribution (tons/y), 2014–2018.

[m] Volatile organic material point source emission distribution (tons/y), 2014–2018.

[n] Adults aged 25–44 years with some post-secondary education, 2014–2018.

[o] Ratio of household income at the 80th percentile to income at the 20th percentile at county level, 2014–2018.

Table 2. Associations Between Retained Factors and the Illinois Cancer Risk Index (ICRI) with the Common Cancer Mortality Rate, 2014–2018

	Multivariable linear regression[a]			Bivariate linear regressions				
Factor name	β	SE	*P* value	Intercept	β	SE	*P* value	r^2
Intercept	20.65	0.02	<.001	—	—	—	—	—
Factor 1	0.23	0.08	<.001	20.64	0.76	0.09	<.001	0.14
Factor 3	2.11	0.07	<.001	20.65	1.49	0.08	<.001	0.13
Factor 5	0.04	0.08	.01	20.64	0.19	0.09	.03	0.002
Factor 6	0.38	0.11	.004	20.66	0.65	0.09	<.001	0.07
ICRI	—	—	—	20.63	1.09	0.16	<.001	0.32

Abbreviations: — , variable not included in the model.

[a] r^2 = 0.335.

Table 3. Regression Analysis, Illinois Cancer Risk Index (ICRI) Relative to Breast, Colorectal, Lung, and Prostate Cancer Mortality Rates, 2014–2018

Age-adjusted cancer mortality rates (2014–2018)	Bivariate linear regressions				
	Intercept	β	SE	*P* value	r^2
Breast cancer	12.483	1.37	0.20	<.001	0.57
Colorectal cancer	12.324	0.10	0.13	<.001	0.56
Lung cancer	51.500	0.39	0.25	.02	0.23
Prostate cancer	8.116	1.33	0.16	<.001	0.64

PCD

PREVENTING
CHRONIC DISEASE