frontiers | Research Topics

Geographic inequalities in health and mortality: Factors contributing to trends and differentials

Edited by
Irma Elo and Shannon Monnat

Published in
Frontiers in Public Health

ISSN 1664-8714
ISBN 978-2-8325-2847-1
DOI 10.3389/978-2-8325-2847-1

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Geographic inequalities in health and mortality: Factors contributing to trends and differentials

Topic editors

Irma Elo — University of Pennsylvania, United States

Shannon Monnat — Syracuse University, United States

Citation

Elo, I., Monnat, S., eds. (2023). *Geographic inequalities in health and mortality: Factors contributing to trends and differentials*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2847-1

Table of contents

frontiers | Frontiers in Public Health

TYPE Editorial
PUBLISHED 14 June 2023
DOI 10.3389/fpubh.2023.1217803

OPEN ACCESS

EDITED AND REVIEWED BY
Konstantin G. Arbeev,
Duke University, United States

*CORRESPONDENCE
Irma T. Elo
✉ popelo@pop.upenn.edu

†These authors have contributed equally to this work

RECEIVED 05 May 2023
ACCEPTED 31 May 2023
PUBLISHED 14 June 2023

CITATION
Monnat SM and Elo IT (2023) Editorial: Geographic inequalities in health and mortality: factors contributing to trends and differentials. *Front. Public Health* 11:1217803. doi: 10.3389/fpubh.2023.1217803

Editorial: Geographic inequalities in health and mortality: factors contributing to trends and differentials

Shannon M. Monnat[1,2,3†] and Irma T. Elo[4,5*†]

[1]Department of Sociology, Syracuse University, Syracuse, NY, United States, [2]Center for Policy Research, Syracuse University, Syracuse, NY, United States, [3]Lerner Center for Public Health Promotion and Population Health, Syracuse University, Syracuse, NY, United States, [4]Department of Sociology, University of Pennsylvania, Philadelphia, PA, United States, [5]Population Studies Center, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, United States

KEYWORDS

geographic, health, mortality, United States, inequality

Editorial on the Research Topic
Geographic inequalities in health and mortality: factors contributing to trends and differentials

Introduction

Geographic inequalities in health and mortality in the United States have grown substantially in recent years (1, 2). Mortality rates vary across and within regions (2–5), states (2, 6–8), counties (3, 4, 9–11), and metropolitan status categories (12–14). Mortality trends have been particularly adverse for working-age adults without a 4-year college degree over the past couple of decades (11). This is due largely to increases in drug overdoses, alcohol-related deaths, suicides, and metabolic diseases and to a stagnation in cardiovascular disease mortality rates that had been declining for many years (2). COVID-19 has exacerbated these long-term trends within the U.S. (15–17). At the same time, U.S. life expectancy continues to deteriorate relative to other high-income countries (18). A recent National Academies of Sciences Report on high and rising midlife mortality highlighted the need for investigations of the multi-level and multidimensional drivers of these trends (2).

This special issue aims to improve our understanding of the factors contributing to high and rising geographic inequalities in health and mortality in the U.S. Across the 10 articles comprising this special issue, 29 scholars with diverse disciplinary perspectives representing the fields of demography, sociology, population health, public health, consumer science, political science, and public administration use a variety of theoretical frameworks, data sources, units of analysis (regions, states, counties, and neighborhoods), and modeling approaches to provide a clearer understanding about the places and subpopulations most affected by adverse health and mortality trends and potential explanations for these trends.

Individual studies, key findings, and insights

Starting out the special issue, Montez and Cheng remind us that educational attainment is strongly related to health and mortality in the U.S., but that "not having a college degree is much riskier for health in some U.S. states than others". Their study sought to determine how

variation in economic wellbeing, health behaviors, family factors, and health care availability and affordability among working-age adults helped explain educational disparities in self-rated health in each state. Using data on over 1.7 million adults ages 25–64 from the 2011 to 2018 Behavioral Risk Factor Surveillance System, they found that educational disparities in health differed substantially across states (primarily due to between-state variation in health among those without a college degree) and that educational disparities in self-rated health were the largest in the Midwest and the South. Moreover, in many states in the South and the Midwest, even individuals with college degrees experienced worse health relative to their peers living in other states. They further found that individual-level economic factors (employment and household income) and behavioral factors (smoking and obesity) were key to explaining educational disparities in self-rated health, but the importance of these factors differed across states. In states with larger educational differences in self-rated health, respondents' economic wellbeing was the dominant mechanism linking education to health, whereas in states with smaller educational disparities in health, the contribution of economic mechanisms was smaller, while the role of behavioral mechanisms increased. The takeaway from this paper is that that educational disparities in health are much worse when less-educated adults have limited access to employment and income, and that structural differences across states, such as characteristics of labor markets and labor market policies, may be key to explaining why those with lower education have worse health than their more highly educated peers.

Wolf's paper points to the role of state-level labor market policies in explaining geographic disparities in access to paid sick leave. Wolf considers the combined roles of state paid sick leave (PSL) policies, preemption of PSL, and right-to-work laws on obtaining access to PSL among U.S. workers from 2009 to 2021. Merging data from the U.S. Department of Labor with state policy data, Wolf finds that workers living in states with PSL mandates do indeed have more access to PSL. However, states' adoption of PSL mandates has occurred alongside the adoption of policies preempting lower levels of government from mandating their own PSL provisions, as well as states' adoption of right-to-work laws. In regression models that consider each policy in isolation, a PSL mandate appears to have a larger positive association on access to PSL than the negative associations of both preemption and right-to-work laws. However, when all three policies are considered in the same model, a PSL mandate with no ceiling and a mandate with no ceiling in combination with right-to-work laws appear to be the most important for PSL coverage. This paper illustrates the importance of considering the reality that "people live in more than one policy at a time" (19), and these policies may have exacerbating or countervailing consequences on health outcomes.

Brown et al. test the role of place-based structural racism on state-level Black-White differences in COVID-19 mortality through August of 2022. They operationalize structural racism using seven measures that span educational, economic, political, criminal-legal, and housing sectors. They find substantial variation in both Black-White disparities in COVID-19 mortality rates and structural racism across states. Notably, COVID-19 mortality rates were higher among Black individuals than among White individuals in all states, but the gap was especially pronounced in states with higher structural racism scores. Specifically, whereas Black COVID-19 mortality rates were about 12% higher in states with a structural racism value of two standard deviations below the average, Black COVID-19 mortality rates were over twice as high in states with a structural racism value of two standard deviations above the average. Their findings illustrate that U.S. states are racialized institutional actors that shape geographic disparities in population health.

In a paper focusing on Black-White disparities in infant mortality, Côté-Gendreau and Moran use linked birth and infant death data from the National Center for Health Statistics to compare Black-White and maternal education disparities in infant mortality by region and metropolitan status from 2011 to 2015. They find that infant mortality rates were higher for Black mothers and mothers with lower educational attainment but that these racial and educational infant mortality disparities vary by metropolitan status and region. Whereas educational, regional, and metropolitan status differences in infant mortality are relatively small among White women, there are large differences among Black women. In metropolitan counties, infant mortality rates are significantly lower among Black mothers with at least a 4-year college degree than among less educated Black mothers. However, the educational gradient in infant mortality among Black mothers living in non-metropolitan counties is flat, suggesting that educational attainment is less protective for Black mothers in non-metropolitan counties. They further find that much of this divergence is being driven by the Midwest and the South, with much lower returns to education for non-metropolitan Black mothers in these regions. Similar to Montez and Cheng, this paper's focus on educational attainment draws attention to the considerable geographic variation in education-mortality gradients and how these gradients vary not only by geography but also between Black and White mothers.

The non-metropolitan (or rural) mortality penalty in the U.S. is long-running, large, and growing. However, there are multiple definitions and operationalizations of "rural" that may affect the conclusions we draw about the magnitude of the rural mortality disadvantage. In their brief research report, James et al. determine whether rural mortality disparities from 1968 to 2020 are consistent across three definitions of county-level rural-urban status: the USDA Economic Research Service's Rural-Urban Continuum Code (RUCC) and Urban Influence Code (UIC) and the National Center for Health Statistics' (NCHS) Rural-Urban Classification Scheme for Counties. In addition to comparing mortality trends using a rural-urban dichotomy derived from each classification system, they also consider within-rural variation in mortality rates using disaggregated non-metropolitan classifications (e.g., comparing medium to small non-metro). They find that the rural mortality penalty is remarkably consistent across these different rural-urban classification schemes. For all three operationalizations, the rural mortality penalty emerged in the mid-1980s and has continued to grow over time. They further find that, even when disaggregating across rural subcategories, mortality trends follow similar patterns throughout the time series. Finally, using any of the three operationalizations, they find consistent spatial concentrations of high rural mortality rates throughout the

Southeast and Appalachia. They conclude that "different definitions yielding strongly similar results suggests robustness of" the rural mortality penalty.

The paper by Hendi and Ho further describes widening disparities in life expectancy between metropolitan and non-metropolitan areas between 1990 and 2019 and examines the contribution of smoking (which is the leading cause of premature morbidity and mortality in the U.S.) to the widening non-metropolitan disadvantage. Using death certificate and U.S. Census data, they estimate life expectancy at age 50 and identify causes of death attributable to smoking in 1990–1992 and 2017–2019 across 40 geographic areas cross-classified by region and metropolitan status. They found that the non-metropolitan disadvantage in life expectancy at age 50 increased by 2.17 years for males and 2.77 years for females over this period. They further found that differential changes in smoking-related mortality (larger declines in large cities and coastal areas and smaller declines in non-metropolitan areas in the South and Midwest) were responsible for 19% of the increase in the non-metropolitan life expectancy disadvantage for males and 22% of the increase for females. They conclude that, while differences in education and income contributed to the widening non-metropolitan disadvantage, these factors alone are not enough to explain why smoking-attributable mortality has not declined at the same pace in non-metropolitan areas compared to metropolitan areas. Instead, the characteristics of non-metropolitan places, particularly in the South where there has been a legacy of economic dependence on tobacco, intensive tobacco industry influence, and limited adoption of tobacco control policies, have contributed to the greater burden of smoking-attributable mortality in the non-metropolitan South.

Drug overdoses have been among the largest contributors to increasing mortality rates in the U.S. over the past three decades, with opioids playing a particularly outsized role. The paper by Yang et al. explores county-level variation in rates of opioid use disorders (OUD) among older adult Medicare beneficiaries—a population that is underexplored in the literature on OUD. Using beneficiary-level data from the U.S. Centers for Medicare and Medicaid Services from 2020 and geographically weighted regression models, they find substantial geographic differences in OUD rates among Medicare beneficiaries, with concentrations of high rates in the Pacific region, Four Corners region, mid-Appalachia, Oklahoma, Michigan, and along the Gulf of Mexico coastal region. Rates are lower across much of the Midwest, the Great Plains, and the Northeast. They further find that county-level differences in age and racial/ethnic composition and the share of beneficiaries with various chronic conditions (chronic obstructive pulmonary disease, diabetes, chronic kidney disease, and hypertension) are the primary determinants of county-level variation in OUD rates. Another important finding is that the share of non-Hispanic White beneficiaries, and average number of mental health and chronic physical health conditions play a larger role in predicting OUD rates in some counties than in others. Their findings highlight the importance of considering local area conditions in addressing OUD among older adults.

Debt appears to be an important social determinant of health in the United States (20, 21). The prevalence of high-cost financial services, like payday lenders, has increased substantially in the U.S. since the mid-1990s, leading to increasing debt burden and financial difficulty (22). Yet, the distribution of payday lenders varies substantially across the U.S., with state regulations, such as interest rate caps, preventing loan rollover or repeat borrowing, and assessing borrowers' ability to repay loans playing critical roles in the variation of payday lender placement. Agnew et al. examine whether the presence of payday lenders in a county is associated with premature mortality rates. They merged county-level mortality data with data on the locations of payday lenders in the U.S. from 2000 to 2017, finding that, even after accounting for county-level socioeconomic conditions, the presence of payday lenders is associated with higher rates of all-cause and cause-specific mortality from mental health related causes, homicide, and cardiovascular diseases. However, illustrating the important role of states that is a theme throughout this volume, they find that state regulations partially buffer the relationship between payday lender placement and mortality, especially in counties with high concentrations of payday lenders. The takeaway is that stronger regulations on payday lenders can protect consumers from taking on the types of risky debt that may be harmful for health.

Whereas the papers summarized thus far have focused on states and counties within the United States, the paper by García et al. focuses on neighborhood context (census blocks) in explaining differences in mortality in Puerto Rico. Linking data from the 2000 U.S. Census to the longitudinal Puerto Rican Elderly Health Conditions Project with follow-up mortality through 2021, the team used latent class analysis to identify the effects of neighborhood conditions on all-cause mortality among adults ages 60 and older. They classified neighborhoods into deprivation clusters based on racial/ethnic, age, socioeconomic, and family-structure composition, and housing features. They find that older adults residing in neighborhoods classified as high deprivation or high-moderate deprivation in 2000 had higher risk of death over the study period compared to those in low deprivation neighborhoods. Their finding that neighborhood disadvantage is associated with increased risk of mortality is consistent with similar studies focused on the U.S. and Latin America, but this is the first study examining these relationships for older adults in Puerto Rico—"a segment of the Latino population that is overlooked in U.S.-based neighborhoods research and aging research more broadly".

Mortality rates surged across the globe during the COVID-19 pandemic, but some countries experienced much higher COVID-19 mortality rates than others. The commentary by Zanwar et al. compared reported COVID-19 mortality rates in the U.S. (3,000 per 100,000 population) and India (370 per 100,000 population) as of July 2022 and considered several potential explanations for the observed differences. They identify India's relatively younger age structure and the undercounting of COVID-19 deaths in India as plausible explanations for lower reported COVID-19 mortality rates in India compared to the U.S. They also summarize findings showing large gender, socioeconomic, and rural-urban differences in COVID-19 mortality rates in both the U.S. and India. They warn that the aging of the global population means that future pandemics have the potential to result in even higher mortality rates than during the COVID-19 pandemic, and they encourage developing nations to invest in more resilient health systems to prepare for inevitable future pandemics.

In sum, these papers illustrate large and growing geographic disparities in health and mortality across the United States by highlighting variations across regions, states, counties, and by metropolitan status. They contribute to the growing body of evidence showing that the United States has become increasingly unequal in terms of place-based health disparities. What particularly stands out are the regional disadvantages in the Midwest and the South relative to other regions of the country. This regional variation reflects considerable state-level variation in health outcomes, with the Midwestern and Southern states typically having worse outcomes than states in the Northeast and West. In this regard, state-level policy context is likely to play an important role (7, 8). For example, the paper by Wolf points to the role of labor market policies, Agnew et al. direct attention to state-level regulations of pay-day lenders, and Hendi and Ho to policies related to tobacco control. Brown et al. in turn demonstrate the important role of multidimensional structural racism in the state-level variation of Black-White mortality disparities. Several papers also further our understanding of the consistent non-metropolitan disadvantage that is shown to be robust to the various rural classification schemes employed (James et al.). Although the characteristics of individuals explain some of the documented geographic inequalities, the papers also demonstrate that individual-level characteristics, such as educational attainment, do not confer the same advantages in all states or in non-metropolitan areas, suggesting that local context can have differential consequences for individuals of diverse socioeconomic backgrounds.

Directions for future research

The papers in this collection provide important descriptive information about geographic inequalities in multiple health outcomes. Although these papers do not assess causality, the papers clearly demonstrate the wide variation in health and mortality by various levels of geography, e.g., state, metropolitan status, and census block group. The COVID-19 pandemic has further exacerbated these geographic (as well as racial/ethnic) inequalities in mortality (Brown et al.). They call attention to the need to move away from the sole focus on individual-level determinants of health outcomes to the role of the broader social and structural contexts in which individuals' lives are embedded (2).

Montez and Cheng lay a foundation for research on why education is more important for health in some states than others. The role of states is also implicated in several other papers [e.g., (21); Wolf; Hendi and Ho; Agnew et al.]. State-level factors that merit further investigation include economic environments (e.g., minimum wage laws, earned income tax credit, paid leave policies, and occupational and industrial structure), structural racism, social programs and health care coverage [e.g., program eligibility for Medicaid and temporary assistance for needy families (TANF)], the regulatory environment (e.g., tobacco control policies), and the political environment (7).

Although state-level factors clearly play important roles in U.S. geographic inequalities in health and mortality, they alone are unlikely to fully explain mortality disparities across counties or the rural-urban continuum or within states (10). Local area characteristics, such as educational attainment, demographic characteristics of the population, economic wellbeing, local employment conditions, health care access, and social and political environments also likely play roles at the sub-state level. These local area characteristics may also be more important for explaining health outcomes for some population subgroups than others (e.g., by socioeconomic status, race/ethnicity, and gender) (Montez and Cheng; Côté-Gendreau and Donnelly Moran).

A critical area for future research involves considering the *joint* influence of state and local contexts. While studies in this issue and elsewhere have provided valuable insights on the separate roles of state and local contexts, state and local contexts are likely to have intersecting and synergistic influences on health and mortality. For example, O'Brien et al. (23) found that state policies helped mitigate or exacerbate the effects of county-level deindustrialization on mortality. Specifically, the adverse effect of automation on working-age male all-cause mortality was smaller in states with more generous unemployment insurance (UI) benefits but larger in states with right-to-work laws. Moreover, state UI generosity, Medicaid generosity, and higher minimum wage significantly buffered the adverse effect of automation on suicide mortality, and state Medicaid generosity mitigated the effect of automation on drug overdose mortality. In another example, Wolf et al. (24) found that state laws interacted with county metropolitan status to influence working-age mortality rates. Specifically, state laws that preempted county and city governments from mandating paid sick leave were associated with significantly higher mortality rates in large central metropolitan counties, but not in small metropolitan or non-metropolitan counties. These studies illustrate the need for more research that can identify and explain how state and local contexts concomitantly contribute to geographic disparities in health and mortality. The extent to which internal migration is related to state and local area context and long-term trends in geographic inequalities also merits further investigation, although its role is unlikely to explain the observed patterns (6, 25).

In addition to contextual factors, health-related behaviors, such as smoking and drug use, are implicated in increasing geographic mortality inequalities, suggesting the need for further studies to investigate factors responsible for their spatial patterning (Hendi and Ho; Yang et al.). Future trends in geographic mortality inequalities may be influenced not only by current health behaviors but also by emerging ones (e.g., e-cigarette use), underscoring the need for continued monitoring and coordinated efforts to prevent the uptake of potentially deleterious health behaviors (Hendi and Ho).

Finally, advancing our understanding of the drivers of current geographic inequalities and trends in health and mortality requires longitudinal data at multiple levels, including states, the rural-urban continuum, counties, and neighborhoods. A central repository of longitudinal state-level and county-level characteristics that can be combined into broader geographic units and that can be linked to individual-level survey data and made accessible to the research community at large should be a high priority of funding agencies. Including geocodes in national

survey data would also help further illuminate the role of place on health disparities.

Policy implications

Collectively, the papers in this volume have several policy implications. A consistent theme throughout the papers is the role of place in its association with multiple health outcomes measured at various levels of geography (e.g., regions, states, counties, and metropolitan status). Taken together they point to the importance of policy for not only individual-level social determinants of health but also for the upstream economic, social, and political contexts that affect individuals' access to resources and shape their everyday life experiences [i.e., structural determinants of health or "the causes of the causes of the causes" (26)]. The papers also highlight the fact that individual-level social determinants of health, such as educational attainment, do not confer the same advantages to everyone. Instead, the returns to higher education are conditioned by residential context. Individuals without a 4-year college degree are particularly disadvantaged. State policies that improve employment opportunities, ensure a living wage, and enhance the overall economic wellbeing of less-educated workers should be prioritized, as they could play a role in reducing geographic inequalities in population health.

States can also play an important role in strengthening regulatory policies and public health interventions aimed at preventing detrimental health behaviors, such as drug use, smoking, excessive drinking, and consumption of unhealthy foods. Such policy interventions could include, for example, restrictions on the use of tobacco products in public places, cigarette taxes, smoke free environments, and improved access to drug treatment programs and mental health services. In designing such policies and public health interventions "more attention should be paid to the place-based policies so that the differences in culture, values, attitudes, norms, and socioeconomic conditions across space can be explicitly considered in possible interventions" (Yang et al.). States also play a role in regulating other spheres of influence associated with variation in health. As Agnew et al. write: "Beyond reducing financial difficulties related to paying bills, affording rent, and filing for bankruptcy that have been a focus of existing research, we suggest that regulating higher-cost financial services might advance community public health and protect against premature mortality for some groups."

People live in more than one context at a time, with policies at the federal, state, and local levels trickling down to affect the proximate determinants of health that eventually lead to morbidity and mortality. As such, policymakers at all levels have important roles to play in creating the conditions that enable individuals and families to thrive and achieve healthy longevity.

Author contributions

All authors made equal contributions to conceptualizing this collection of papers, the summary of the findings and research, and policy recommendations.

Funding

The authors acknowledge funding from the National Institute on Drug Abuse (U01DA055972) and support from the National Institute on Aging-funded Center for Aging and Policy Studies at Syracuse University (P30AG066583), two research networks funded by the National Institute on Aging (R24AG065159 and 2R24AG045061), the USDA-supported Rural Population Research Network (W5001), the Syracuse University Lerner Center for Public Health Promotion and Population Health, the National Institute on Aging (R01AG060115); the Eunice Kennedy Shriver National Institute of Child Health and Human Development-Funded Population Studies Center (P2CHD044964), and the NIA-Funded Population Aging Research Center (P30AG012836) at the University of Pennsylvania. The funders played no role in the writing of the manuscript or decision to submit it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Vierboom YC, Preston SH, Hendi AS. Rising geographic inequality in mortality in the United States. *SSM Popul Health.* (2019) 9:100478. doi: 10.1016/j.ssmph.2019.100478

2. National Academies of Sciences Engineering and Medicine. *High and Rising Mortality Rates among Working-Age Adults*. Washington, DC: The National Academies Press (2021). p. 570.

3. Murray CJL, Kulkarni SC, Michaud C, Tomijima N, Bulzacchelli MT, Iandiorio TJ, et al. Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Med.* (2006) 3:e260. doi: 10.1371/journal.pmed.0030260

4. Cossman JS, Cossman RE, James WL, Campbell CR, Blanchard TC, Cosby AG. Persistent clusters of mortality in the United States. *Am J Public Health Res.* (2007) 97:2148–50. doi: 10.2105/AJPH.2006.093112

5. James W, Cossman J, Wolf J. Persistence of death in the United States: the remarkably different mortality patterns between America's Heartland and Dixieland. *Demogr Res.* (2018) 39:897–910. doi: 10.4054/DemRes.2018.39.33

6. Fenelon A. Geographic divergence in mortality in the United States. *Popul Dev Rev.* (2013) 39:611–34. doi: 10.1111/j.1728-4457.2013.00630.x

7. Montez JK, Beckfield J, Cooney JK, Grumbach JM, Hayward MD, Koytak HZ, et al. US state policies, politics, and life expectancy. *Milbank Q.* (2020) 98:668–99. doi: 10.1111/1468-0009.12469

8. Montez JK, Mehri N, Monnat SM, Beckfield J, Chapman D, Grumbach JM, et al. US state policy contexts and mortality of working-age adults. *PLoS ONE.* (2022) 17:e0275466. doi: 10.1371/journal.pone.0275466

9. Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, Morozoff C, Kutz MJ, Huynh C, et al. US county-level trends in mortality rates for major causes of death, 1980–2014. *JAMA.* (2016) 316:2385–401. doi: 10.1001/jama.2016.13645

10. Graetz N, Elo IT. Decomposing county-level working-age mortality trends in the United States between 1999–2001 and 2015–2017. *Spat Demogr.* (2022) 10:33–74. doi: 10.1007/s40980-021-00095-6

11. Woolf SH, Schoomaker H. Life expectancy and mortality rates in the United States, 1959–2017. *JAMA.* (2019) 322:1996–2016. doi: 10.1001/jama.2019.16932

12. Cosby AG, McDoom-Echebiri MM, James W, Khandekar H, Brown W, Hanna HL. Growth and persistence of place-based mortality in the United States: the rural mortality penalty. *Am J Public Health.* (2019) 109:155–62. doi: 10.2105/AJPH.2018.304787

13. Elo IT, Hendi AS, Ho JY, Vierboom YC, Preston SH. Trends in non-hispanic white mortality in the United States by metropolitan-nonmetropolitan status and region, 1990–2016. *Popul Dev Rev.* (2019) 45:549–83. doi: 10.1111/padr.12249

14. Monnat SM. Trends in U.S. working-age non-hispanic white mortality: rural–urban and within-rural differences. *Popul Res Policy Rev.* (2020) 39:805–34. doi: 10.1007/s11113-020-09607-6

15. Zhang CH, Schwartz GG. Spatial disparities in coronavirus incidence and mortality in the united states: an ecological analysis as of May 2020. *J Rural Health.* (2020) 36:433–45. doi: 10.1111/jrh.12476

16. Paglino E, Lundberg DJ, Zhou Z, Wasserman JA, Raquib R, Luck AN, et al. Monthly excess mortality across counties in the United States during the COVID-19 pandemic, March 2020 to February 2022. *medRxiv.* (2022). doi: 10.1101/2022.04.23.22274192. [Epub ahead of print].

17. Sun Y, Cheng KJG, Monnat S. Rural-urban and within-rural differences in COVID-19 mortality rates. *J Rural Soc Sci.* (2022) 37:3. doi: 10.31235/osf.io/jbhvs

18. Ho JY, Hendi AS. Recent trends in life expectancy across high income countries: retrospective observational study. *BMJ.* (2018) 362:k2562. doi: 10.1136/bmj.k2562

19. Beckfield J, Bambra C, Eikemo TA, Huijts T, McNamara C, Wendt C. An institutional theory of welfare state effects on the distribution of population health. *Soc Theory Health.* (2015) 13:227–44. doi: 10.1057/sth.2015.19

20. Batomen B, Sweet E, Nandi A. Social inequalities, debt, and health in the United States. *SSM Popul Health.* (2021) 13:100736. doi: 10.1016/j.ssmph.2021.100736

21. Fitch C, Hamilton S, Bassett P, Davey R. The relationship between personal debt and mental health: a systematic review. *Ment Health Rev.* (2011) 16:153–66. doi: 10.1108/13619321111202313

22. Federal Reserve Bank of New York. *Quarterly Report on Household Debt and Credit: 2019 Q3*. New York, NY: Federal Reserve Bank of New York, Research and Statistics Group (2019). 46 p.

23. O'Brien R, Bair EF, Venkataramani AS. Death by robots? Automation and working-age mortality in the United States. *Demography.* (2022) 59:607–28. doi: 10.1215/00703370-9774819

24. Wolf DA, Montez JK, Monnat SMUS. State preemption laws and working-age mortality. *Am J Prev Med.* (2022) 63:681–8. doi: 10.1016/j.amepre.2022.06.005

25. Couillard BK, Foote CL, Gandhi K, Meara E, Skinner J. Rising geographic disparities in US mortality. *J Econ Perspect.* (2021) 35:123–46. doi: 10.1257/jep.35.4.123

26. Bambra C, Smith KE, Pearce J. Scaling Up: The politics of health and place. *Soc Sci Med.* (2019) 232:36–42. doi: 10.1016/j.socscimed.2019.04.036

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 16 August 2022
DOI 10.3389/fpubh.2022.966434

OPEN ACCESS

EDITED BY
Irma Elo,
University of Pennsylvania,
United States

REVIEWED BY
Lauren Gaydosh,
University of Texas at Austin,
United States
Elizabeth Lawrence,
University of Nevada, Las Vegas,
United States

*CORRESPONDENCE
Jennifer Karas Montez
jmontez@syr.edu

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 10 June 2022
ACCEPTED 01 August 2022
PUBLISHED 16 August 2022

CITATION
Montez JK and Cheng KJ (2022) Educational disparities in adult health across U.S. states: Larger disparities reflect economic factors. *Front. Public Health* 10:966434. doi: 10.3389/fpubh.2022.966434

Educational disparities in adult health across U.S. states: Larger disparities reflect economic factors

Jennifer Karas Montez[1]* and Kent Jason Cheng[2]

[1]Department of Sociology, Syracuse University, Syracuse, NY, United States, [2]Department of Social Science, Syracuse University, Syracuse, NY, United States

Introduction: Education level is positively associated with adult health in the United States. However, new research shows that the association is stronger in some U.S. states than others, and that states with stronger associations also tend to have poorer overall levels of health. Understanding why educational disparities in health are larger in some states than others can advance knowledge of the major drivers of these disparities, between individuals and states. To that end, this study examined how key mechanisms (economic conditions, health behaviors, family, healthcare) help explain the education-health association in each state and whether they do so systematically.

Methods: Using data on over 1.7 million adults ages 25–64 in the 2011–2018 Behavioral Risk Factor Surveillance System, we estimated the association between education level and self-rated health in each state, net of age, sex, race/ethnicity, and calendar year. We then estimated the contribution of economic, behavioral, family, and healthcare mechanisms to the association in each state.

Results: The strength of the education-health association differed markedly across states and was strongest in the Midwest and South. Collectively, the mechanisms accounted for most of the association in all states, from 55% of it in North Dakota to 73% in Oklahoma. Economic (employment, income) and behavioral (smoking, obesity) mechanisms were key, but their contribution to the association differed systematically across states. In states with stronger education-health associations, economic conditions were the dominant mechanism linking education to health, but in states with weaker associations, the contribution of economic mechanisms waned and that of behavioral mechanisms rose.

Discussion: Meaningful reductions in educational disparities in health, and overall improvements in health, may come from prioritizing access to employment and livable income among adults without a 4-year college degree, particularly in Southern and Midwestern states.

KEYWORDS

education, health, disparities, fundamental cause, U.S. states

Introduction

An adult's education level is a robust predictor of their health in the United States. Compared to their less-educated peers, adults with more education have better overall health, are less likely to develop morbidities and disability, and tend to live longer and spend more of those years in good health (1, 2). The magnitude of these disparities is striking. Among U.S. adults in their mid-40s, <15% of those without a high school diploma reported being in excellent health, compared to 24% of those with a high school diploma, over 40% of those with a 4-year college degree, and over 50% of those with a doctorate or professional degree (3). In recent decades, disparities in health between adults with and without a 4-year college degree have become especially pronounced (4, 5).

New research finds that the importance of one's education level for health differs markedly across U.S. states (6–8). As an example, Figure 1 shows the association between education and self-rated health by state among adults ages 25–64 (the associations are adjusted for age, sex, and race-ethnicity differences across states' populations). The association is strongest in West Virginia, where just 69% of adults without a 4-year college degree report being in favorable health compared to 90% of their more-educated peers, a gap of 21 percentage points. The association is weakest in Utah, where 85% of adults without a 4-year degree and 94% of their more-educated peers are in favorable health, a gap of just 9 percentage points. Also intriguing, this new area of research finds that states with the largest disparities in health across education levels tend to have the worst overall health (6–8). In other words, these states are especially disadvantaged. Taken together, these findings imply that understanding why education is a stronger predictor of health in some states than others could advance knowledge of the major drivers of health levels and disparities, between individuals and between states.

A key framework for understanding the education-health association is Fundamental Cause Theory (FCT). It asserts that education is important in contexts with the resources to avoid disease and premature death, yet more-educated persons have greater access to those resources (9–11). Indeed, compared to their less-educated peers, more-educated U.S. adults have greater access to four types of salubrious resources: economic well-being, social ties, healthy behaviors, and quality health care (2, 12). Those four types of resources, or "mechanisms," help explain a large share of the education-health association in the country today (2, 12). Central to the current study, the relevance of each mechanism for explaining the education-health association may vary across states. For instance, unemployment rates, the share of jobs requiring a college degree, and median income vary across states (13). Access to affordable health care for disadvantaged adults also varies across states. Having low education may pose substantial barriers to health care in states that offer minimal levels of Medicaid benefits. As another example, the relevance of smoking for the education-health association may partly depend on states' tobacco control policies (6). If the salience of such mechanisms differs across states, this information may point to reasons why the education-health association is stronger in some states than others (i.e., certain mechanisms may be key) or it may have no bearing on the strength of the association (i.e., a high salience of one mechanism in a state may simply be offset by the low salience of another).

This study examines how the importance of four key mechanisms (economic conditions, health behaviors, social factors, healthcare) linking educational attainment to self-rated health differs across states. Using data spanning 2011–2018 from over 1.7 million U.S. adults ages 25–64 years, it assesses how much of the education-health association in each state is accounted for by these mechanisms, and whether their importance differs across states in a systematic way. In other words, does the importance of these mechanisms vary across states? Such patterns can provide insights into why the education-health association is stronger in some states than others and point to strategies to reduce educational disparities in health and improve overall health.

Materials and methods

Data and sample

We used data from the Behavioral Risk Factor Surveillance System (BRFSS), an annual cross-sectional survey of the noninstitutionalized U.S. adult population aged 18 and older. The BRFSS is the best available data to examine the education-health association within states because the dataset is large, representative of noninstitutionalized adults at the state level, and contains information on educational attainment, health, and the four mechanisms examined in this study.

We used BRFSS data from 2011 through 2018. We start in 2011 when BRFSS expanded the sample to also include households with only cell phones and revised the weighting methodology (14). We restricted the sample to adults ages 25–64 years. The lower age limit was set at 25 because our main exposure is completed education through a Bachelors' degree. The upper limit was set at 64 because some of the mechanisms, such as employment, are most relevant for working ages. The 2011–2018 BRFSS contains 2,172,540 adults ages 25–64 years.

Self-rated health and educational attainment

We examine self-rated health, a valid indicator of overall health (15). BRFSS asks adults, "Would you say that in general your health is excellent, very good, good, fair, or poor?"

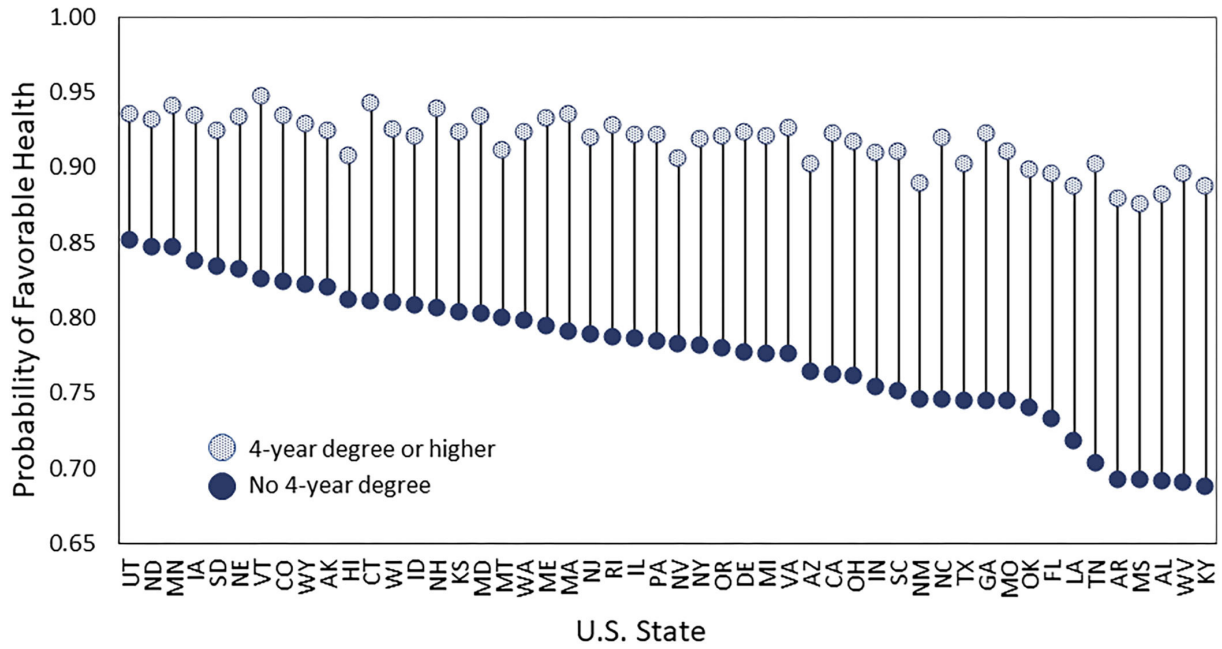


FIGURE 1
Probability of reporting favorable health among adults ages 25–64 by U.S. state. Data are from the 2011–2018 BRFSS, include adults ages 25–64, and are adjusted for age, sex, and race-ethnicity differences between states. Adults who reported that their health was excellent, very good, or good are considered in "favorable" health, unlike those who reported that their health was fair or poor.

We dichotomized the responses, as typically done, so that 1 = excellent, very good, or good (which we refer to as "favorable health") and 0 = fair or poor. The dichotomization is advantageous for the present study because it avoids a complication that would arise from using all responses in ordinal logit models, as the proportional odds assumption is likely to be violated in some U.S. states but not others.

To capture educational attainment, BRFSS asks respondents, "What is the highest grade or year of school you completed?" It has six response categories: never attended school or kindergarten, grades 1–8, grades 9–11, grade 12 or GED, 1–3 years of college, and four or more years of college. Our preferred specification is a dichotomous indicator, where 1 = four or more years of college (we call this group college graduates). It reflects studies showing that the health of U.S. adults has bifurcated, with college graduates doing well and others doing poorly (5, 16, 17) and that health disparities between states are largest for those without a 4-year college degree (6, 7).

Hypothesized mechanisms

We examined four types of mechanisms: economic conditions, health behaviors, social factors, and healthcare. We refer to them as mechanisms because they are hypothesized to be key pathways linking education to health (2, 12) and because this term is prominently used in FCT (10). Although the term "mechanism" often has a causal connotation, we make no causal claims in this analysis.

All mechanisms were measured as continuous, ordinal, or binary variables to facilitate the mediation analysis described below. The two economic factors were employment status (1 = currently employed) and annual household income. To obtain information on employment status, BRFSS provides eight possible employment categories (e.g., employed for wages, self-employed, out of work for 1 year or more) and asks respondents to select the category that best describes them currently. The BRFSS asks respondents about their annual income from all sources. It provides the responses in categories of varying widths ($<$ \$10,000; \$10,000 to $<$ \$15,000;…;\$50,000 to $<$ \$75,000; and $\geq$ \$75,000). We converted this measure into a continuous one based on the recommendation of a validation study, which found that using the upper limit of each category provided the best overall match to the actual income distribution (18).

Three behavior-related mechanisms included smoking, heavy drinking, and obesity. BRFSS assigns a smoking status to respondents based on their answers to questions about past and current cigarette smoking. We included a binary indicator of smoking status, with never smoker = 1 and current and former smoker = 0. The survey provides a binary indicator of heavy drinking (defined as more than 14 drinks per week among men and more than 7 drinks per week among women) based on respondents' answers to questions about the frequency and quantity of alcoholic beverage consumption during the past 30

days. As a proxy for health-related behaviors, we also included a measure of obesity, defined as a body mass index of 30 or higher. The BRFSS calculates BMI based on respondents' reports of their height and weight without shoes.

For the social mechanisms, we included two measures of family, given that family composition differs across education levels and is considered one of the key social mechanisms liking education to health (12). The BRFSS does not contain measures of other social factors such as friendships or loneliness. Specifically, we included the self-reported number of children under 18 years of age in the household (top coded at 10) and self-reported marital status (1 = married). Lastly, the analysis incorporated two healthcare mechanisms related to healthcare availability and affordability. The availability question asked adults if they currently had "any kind of health care coverage, including health insurance, prepaid plans such as HMOs, or government plans such as Medicare, or Indian Health Service." The affordability question asked, "was there a time in the past 12 months when you needed to see a doctor but could not because of cost?"

Covariates

We included calendar year and three self-reported covariates, age, sex, and race/ethnicity. We accounted for age, sex, and race/ethnicity because they are related to both educational attainment and health and because their relevance for the education-health association may differ across states. For instance, a recent study showed that higher education does not provide the same degree of cardiometabolic health benefits for Black adults as it does for White or Hispanic adults (19). We measured age in 5-year groups, from 25 to 29 through 60–64 years. The BRFSS provides sex as female or male. The BRFSS combines respondents' answers to a question about Hispanic/Latinx/Spanish origin and a question about which group (White; Black or African American; American Indian or Alaskan Native; Asian; Pacific Islander) best represents their race into a single variable identifying respondents as non-Hispanic Black, non-Hispanic Other, non-Hispanic White, and Hispanic.

Methods

We estimated logistic regression models using the form below, where b_1 is the coefficient of interest. The vectors b_2, b_3, b_4, and b_5 represent the coefficients for the economic, behavioral, family, and healthcare mechanisms, respectively. The b_6 vector contains coefficients for the covariates, age, sex, race-ethnicity, and calendar year.

$$\ln\left(\frac{p}{(1-p)}\right) = b_0 + b_1 college + \mathbf{b_2} economic + \mathbf{b_3} behaviors + \mathbf{b_4} family + \mathbf{b_5} healthcare + \mathbf{b_6} covariates$$

We estimated a model for each state, which achieves the aims of the study because it allows the importance of the mechanisms in accounting for the education-health association to differ across states. Alternatively, achieving these aims with one model containing all 50 states would require interactions between each state and education and the nine mechanisms (i.e., nearly 500 interaction terms). The notional simplicity of a one-model approach is outweighed by the complexity of hundreds of interaction terms in the mediation analysis.

To examine the contribution of the hypothesized mechanisms to the education-health association within states, we used the method developed by Karlson, Holm, and Breen (KHB) to assess mediation in non-linear probability models (20). It decomposes the difference in the logit coefficient of a variable X (in our case, *college*'s coefficient, b_1) between models with and without the mechanisms Z (i.e., economic, behavioral, family, healthcare), into the portion attributable to Z, while accounting for the rescaling of the X coefficient that occurs across nested non-linear probability models.

A few respondents were missing information on some variables. In preliminary analyses, we assessed several approaches for handling the missing information, such as excluding respondents with missing data or using multiple imputation (details are in Supplementary Tables 1, 2). Because the findings were similar for both approaches, we chose the former one for our main analyses, which includes 1,716,757 adults. All models were estimated with Stata MP 17.

Results

We first describe a few key descriptive statistics from Table 1. Among U.S. adults ages 25–64 during 2011–2018, 83% reported being in favorable health. This percentage ranged from 76% in West Virginia to 89% in Minnesota and Vermont. The percentage of college graduates ranged from 21% in West Virginia to 45% in Massachusetts. States differed in several of the mechanisms. For example, the percentage of adults who had never smoked ranged from 45% in Kentucky and West Virginia to 72% in Utah, and percentage of those who were employed ranged from 63% in West Virginia to 82% in North and South Dakota. In contrast, some mechanisms differed little across states, such as the prevalence of heavy alcohol consumption and the number of children in the household.

We then estimated the 50 state-specific logistic regression models predicting favorable health. In all 50 states, having

TABLE 1 Weighted descriptive statistics of U.S. adults ages 25–64 by state.

	Favorable health	College graduate	Employed	Household income	Never smoked	Heavy drinker	Obese	Married	Number of children in the home	Healthcare coverage	Healthcare not affordable
AL	78	25	64	49,321	52	6	39	61	0.85	83	21
AK	86	29	74	58,956	51	8	32	65	0.99	85	15
AZ	82	29	68	51,637	57	7	32	63	1.04	83	18
AR	77	22	65	46,983	48	6	39	63	0.94	82	21
CA	82	32	70	52,866	64	7	28	64	0.97	84	16
CO	87	40	76	58,150	57	7	24	67	0.92	86	15
CT	88	40	77	59,708	57	7	29	65	0.83	91	12
DE	85	31	75	55,882	54	7	34	61	0.89	89	14
FL	82	29	69	50,674	55	8	31	60	0.83	79	22
GA	82	30	69	51,158	58	6	34	61	0.92	79	21
HI	86	32	78	56,737	58	9	27	62	0.91	92	9
ID	86	27	73	53,558	60	7	31	71	1.17	81	18
IL	84	34	73	55,217	57	7	33	63	0.91	86	14
IN	82	26	72	53,240	50	6	36	65	0.97	85	17
IA	88	30	80	57,719	54	8	36	70	1.00	91	10
KS	85	34	76	55,967	55	6	36	69	1.00	85	15
KY	78	24	66	51,360	45	7	38	63	0.87	87	18
LA	79	24	67	49,896	52	7	39	57	0.93	81	21
ME	85	29	74	54,107	47	9	32	67	0.76	87	13
MD	87	40	77	59,822	61	6	33	62	0.87	90	12
MA	88	45	77	60,140	57	8	26	64	0.80	94	10
MI	83	29	69	53,815	50	8	35	63	0.89	88	16
MN	89	37	81	59,970	55	8	30	68	0.95	91	11
MS	77	22	65	45,213	53	6	41	56	0.91	78	24
MO	83	30	72	53,953	50	8	35	65	0.91	85	16
MT	85	31	74	52,752	53	9	28	67	0.90	83	15
NE	87	32	80	56,780	56	8	35	69	1.03	86	14
NV	82	24	70	52,169	56	7	29	61	0.97	79	19
NH	88	37	77	60,749	52	8	30	69	0.81	89	12
NJ	85	40	75	59,052	59	5	29	65	0.88	87	15
NM	80	26	67	46,997	56	6	32	60	0.99	83	19
NY	85	37	72	53,866	59	6	28	60	0.86	88	14
NC	82	31	71	52,257	54	6	35	63	0.82	82	19
ND	88	31	82	59,903	53	8	36	69	0.93	90	9
OH	83	28	72	53,784	50	7	35	63	0.90	89	14
OK	80	26	69	51,190	51	5	37	65	0.96	82	19
OR	84	33	69	53,885	55	9	31	66	0.84	86	17
PA	84	32	73	56,104	52	7	33	63	0.85	89	14
RI	85	34	73	55,822	54	7	30	62	0.80	89	14
SC	82	27	70	50,348	52	7	37	61	0.87	81	20
SD	88	30	82	56,686	52	7	33	69	0.99	89	12
TN	79	26	68	49,686	51	5	37	62	0.85	83	19
TX	82	29	72	52,011	61	8	36	65	1.04	74	21
UT	88	33	76	59,690	72	5	28	74	1.43	86	15

(Continued)

TABLE 1 Continued

	Favorable health	College graduate	Employed	Household income	Never smoked	Heavy drinker	Obese	Married	Number of children in the home	Healthcare coverage	Healthcare not affordable
VT	89	37	79	56,878	52	9	28	66	0.75	91	10
VA	85	39	76	57,941	57	6	32	65	0.88	87	15
WA	86	34	72	57,661	57	8	30	67	0.87	87	14
WV	76	21	63	48,091	45	4	40	64	0.80	85	20
WI	86	30	77	56,440	53	9	33	66	0.90	90	13
WY	87	27	76	57,528	53	7	31	68	0.96	82	16
Min	76	21	63	45,213	45	4	24	56	0.75	74	9
U.S.	83	31	72	53,897	56	7	32	63	0.92	84	17
Max	89	45	82	60,749	72	9	41	74	1.43	94	24

All numbers are percentages, except for household income ($) and number of children in the home (N=1,716,757).

a college degree is associated with a significantly higher probability of reporting favorable health but the magnitude of the association differs across states, consistent with prior research and Figure 1. Supplementary Figure 1 displays the college coefficient estimated from each of the 50 models along with 95% confidence intervals (it shows that all 50 coefficients are significantly different from zero). Supplementary Figure 2 adjusts the confidence intervals so that comparisons between states can be made (it shows significant differences between many states). We used the 50 state-specific regression models to answer our two research questions, as described below.

How do the mechanisms contribute to the education-health association across states?

Figure 2 shows how much of the education-health association in each state is explained by the nine mechanisms in total (detailed model results are in Supplementary Table 3). The total contribution of the mechanisms, shown as the dashed gray line, differs considerably across states. They explain as little as 55% of the education-health association in North Dakota and as much as 73% in Oklahoma, an 18 percentage-point difference. To explore systematic patterns, Figure 2 shows states sorted from left to right in ascending order of the strength of the education-health association. As a group, the mechanisms are not much better at explaining stronger or weaker associations. This is evidenced by the relatively horizontal dashed gray line and weak correlation ($r = 0.23$, $p = 0.11$) between the total contribution of the mechanisms and strength of the education-health association across the states.

How does the importance of each mechanism vary across states? Do mechanisms vary?

Using the same models from above, we examine the contribution of each of the nine mechanisms to the education-health association within states (i.e., the sum of the contribution of the nine mechanisms = the total contribution). The contributions of income, employment, smoking, and obesity were large and unequal across states, as shown in Figure 2 (the contributions of the other five mechanisms were small and differed little across states, as shown in Figure 3). Figure 2 reveals several intriguing patterns. First, income is the dominant contributor across all 50 states. Nevertheless, income's contribution varies considerably from just 23% in North Dakota to 37% in Oklahoma. Second, the contributions of both income and employment rise across states as the association becomes stronger (i.e., their contribution is larger in states on the right side of Figure 2 than for those on the left side). We can quantify this pattern: the correlation between the strength of the education-health association and the contribution of income is 0.39 ($p < 0.01$) and the correlation between the strength of the education-health association and the contribution of employment is 0.71 ($p < 0.001$). Third, in contrast to economic mechanisms, the contribution of behavior-related mechanisms decreases across states as the education-health association becomes stronger. Consequently, in states like West Virginia and Tennessee with strong education-health associations, the contribution of employment (17.7%) is more than double that of smoking (8.3%), while in states like Hawaii and South Dakota with weak associations, smoking (12.5%) contributes more than employment (9.9%). Taken together, these three patterns are consistent with the notion of mechanisms changing from context to context in the FCT.

FIGURE 2
Contribution of four key mechanisms to the education-health association in U.S. states. Data are from the 2011–2018 BRFSS and include adults ages 25–64 years. States are ordered from left to right in ascending order of the strength of their education-health association. *r* = correlation between the strength of the education-health association in each state and the percentage contribution of each mechanism.

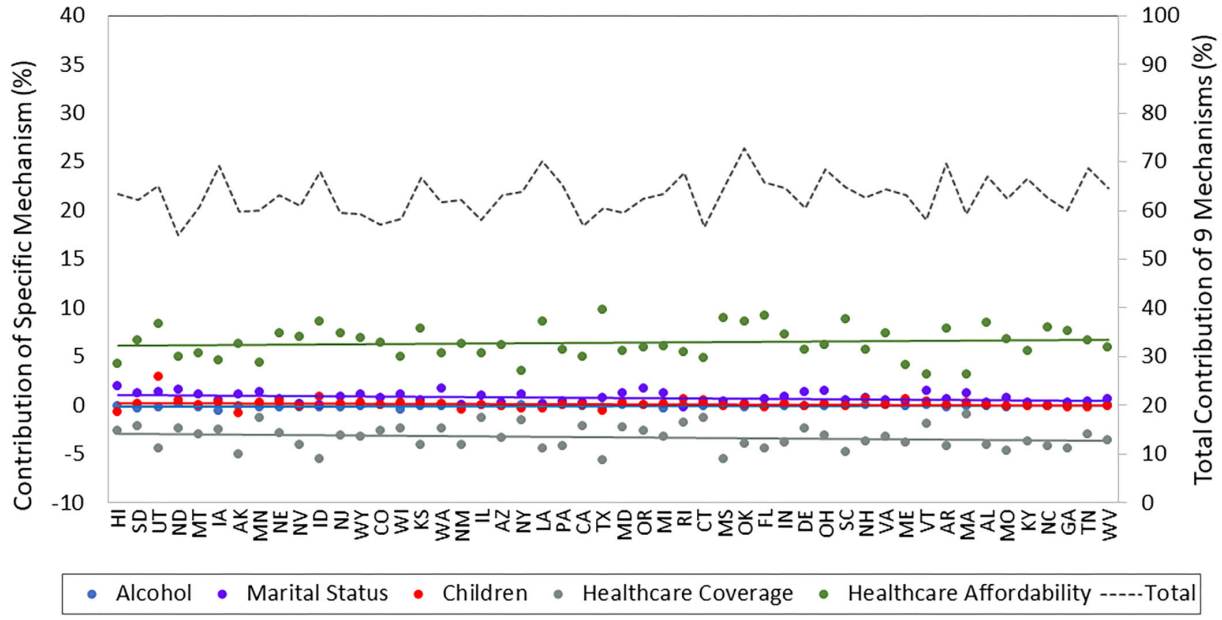
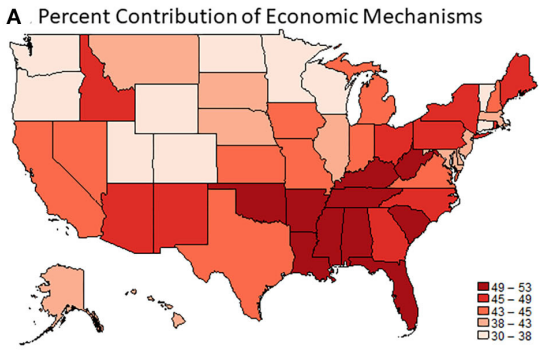


FIGURE 3
Contribution of other mechanisms to the education-health association in U.S. states. Data are from the 2011–2018 BRFSS and include adults ages 25–64 years. States are ordered from left to right in ascending order of the strength of their education-health association.

Another view of how mechanisms vary across states is provided in Figure 4. Panel A shades states according to the contribution of the economic mechanisms (income, employment) to the association, where darker shades of red indicate a larger contribution. Panel B shades states according to the contribution of the behavioral mechanisms (smoking,



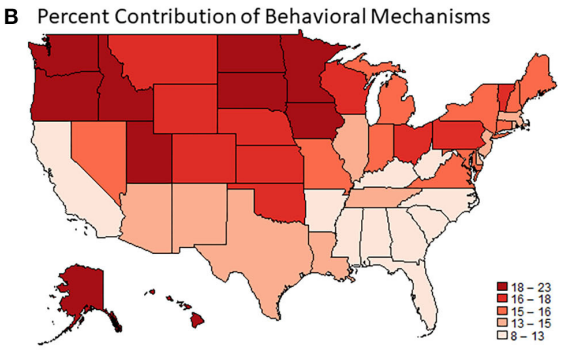


FIGURE 4
The contribution of economic **(A)** and behavioral **(B)** mechanisms to the education-health association in U.S. states. Data are from the 2011–2018 Behavioral Risk Factor Surveillance System and include adults ages 25–64. The magnitude of the education-health association is the difference in the probably of favorable health among adults with at least a 4-year college degree minus the probability of favorable health among adults without a 4-year college degree. The probabilities are estimated from state-specific logistic regression models which include an indicator for college degree, age, sex, race-ethnicity, and calendar year.

obesity) to the associations. The two panels are near mirror images of each other: states where economic mechanisms are particularly important in explaining the education-health association, such as states in the South and Appalachia, tend to be the same states where behavior-related mechanisms are least important.

Robustness checks

We replicated the analyses using different regression models and specifications of education and self-rated health. We first assessed whether our findings were consistent when using a linear probability model instead of a logistic regression model. The findings from these analyses (Supplementary Figure 3) corroborate our main findings. That is, the contribution of economic and behavioral mechanisms to the education-health association were sizable but differed systematically across states such that the importance of economic mechanisms was higher, where the relevance behavioral mechanisms were lower, as the education-health association increased. For example, in the state with the strongest education-health association, income and employment accounted for 48% of it, while smoking and obesity accounted for 11%; in the state with the weakest association, income and employment accounted for 28% and smoking and obesity accounted for 20%. In the linear probability models, the importance of employment rivaled that of income in many Midwestern and Southern states where the education-health association was strongest (e.g., WV, TN, KY, AR, AL, MS, GA).

Next, we assessed whether our findings were consistent when using a different specification of education level. Recall that the BRFSS provides a 6-category measure of education: never attended school or kindergarten, grades 1–8, grades 9–11, grade 12 or GED, 1–3 years of college, and four or more years of college. We created a pseudo-continuous measure by imputing approximate years of schooling (0, 4.5, 10, 12, 14, 18 years) to each category. Although this measure uses all information in the BRFSS, it has only six possible values, thereby posing challenges to estimating a linear relationship with self-rated health. Nevertheless, we replicated the analyses using this measure with both the logistic and linear probability models. The overall findings were similar to those when using the binary measure. That is, the contribution of economic and behavioral mechanisms was large but differed across states such that when the importance of economic mechanisms was higher, that of behavioral mechanisms was lower, as the education-health association became stronger (Supplementary Figures 4, 5). Among three of the four analyses (two regression models x two measures of education), income generally contributed more than employment to the association. The exception was the logistic model with the pseudo-continuous measure where employment generally contributed more than income. Unsurprisingly, given the drawback of the pseudo-continuous measure mentioned previously, the mechanisms did not explain as much of the education-health association with the pseudo-continuous measure as they did with the binary measure. Using logistic regression, the mechanisms collectively explained 55 to 73% of the education-health association in each state when using the binary measure and 34 to 55% when using the pseudo-continuous measure. Lastly, our results were robust to using an OLS with all five values of self-rated health (Supplementary Figure 6).

Discussion

New research has shown that the association between educational level and health is stronger in some U.S. states than others, and that states with stronger associations also tend to have poorer overall levels of health. Understanding why

educational disparities in health are larger in some states than others can advance knowledge of the major drivers of these disparities, between individuals and states. To that end, this study examined how key mechanisms—economic conditions, health behaviors, family, and healthcare—help explain the education-health association in each state and if they do so systematically. Below, we summarize four key findings.

The first set of findings is descriptive. Specifically, the strength of the association between education level and self-reported health differed markedly across U.S. states. It differed across states mainly because the health of nongraduates differed, consistent with prior work (6, 7). For instance, as shown in Figure 1, 90% of adults in West Virginia with at least a 4-year degree were in favorable health, as were 94% of their peers in Utah, while 69% of adults in West Virginia without a 4-year degree were in favorable health, as were 85% of their peers in Utah. This pattern comports with the notion that higher education acts as a "personal firewall" to protect health across contexts (7). Nevertheless, there appears to be limits on how much protection a college degree affords in contexts that are highly problematic for health. For instance, in many states in the South and Midwest (e.g., KY, WV, AL, MI, AR) where nongraduates had strikingly worse health than the rest of the country, the health of graduates also suffered considerably. Such states are especially disadvantaged: they have large disparities in health across education levels and relatively low overall levels of health.

Second, in states where the education-health association was especially strong, economic conditions were the dominant mechanism linking education to health. This suggests that educational disparities in health are exacerbated when less-educated adults have especially limited access to employment and income needed for health-sustaining resources such as nutritious food and safe housing. To the extent that these states improved opportunities for desirable employment and livable wages among college nongraduates, the largest educational disparities in health in the country may be substantially reduced. In other words, states that provide opportunities for economic well-being among nongraduates—for example, through higher minimum wage, earned income tax credits, worker protections, and robust labor markets—may be able to disrupt the pathway from education to economic conditions to health. One interpretation is that structural factors, particularly labor markets, are central for explaining the largest educational disparities in health in the country.

A third key finding is that in states where the link between education level and economic conditions was not as strong, health-related behaviors were more relevant in explaining the education-health association. Specifically, looking across states, as the association became weaker, the contribution of economic mechanisms to the association fell while that of behavior-related mechanisms rose. In states with the weakest associations, smoking rivaled employment as the second most important contributor (income was generally the most important). This pattern would have been obscured if we had only examined the total contribution of the mechanisms, as the total did not rise or fall across states according to the strength of the association. In other words, the importance of certain mechanisms varies from context to context, consistent with a core premise of FCT. Even though the importance of the mechanisms varies across contexts, our findings point to improving employment and income among nongraduates as a potentially effective strategy, as the largest disparities are in states where economic conditions are the dominant mechanism linking education to health.

Fourth, the mechanisms often hypothesized to explain the education-health association (economic conditions, behaviors, social factors such as family, and healthcare) were better able to explain the association in some states than others. Collectively, the total contribution of the mechanisms accounted for as little as 55% of the association in North Dakota to as much as 73% in Oklahoma, a range of 18 percentage points. This range largely reflects the varying contribution of economic conditions. In general, the more closely that education was tied to economic conditions in a state, the more of association that we explained. Among the 10 states where we were best able to explain the association, economic conditions were the single dominant mechanism in some (LA, AR, TN, AL, RI) and shared a high degree of importance with behaviors in others (ID, IA, KS, OK, OH). Among the 10 states where we were least able to explain the association (ND, CT, CA, CO, IL, VT, WI, WY, MA, MD), factors other than those examined in this study also carry considerable weight.

Our findings generally comport with FCT. The fact that college graduates had better health than nongraduates in all 50 states supports FCT's assertion that educational disparities in health persist because more-educated adults use their resources to secure health advantages across contexts. In addition, our finding that nongraduates' health differed markedly across states aligns with FCT's claim that it is essential to understand what puts lower SES individuals "at risks of risks." In states where the education-health association was strongest, less-educated adults were at particularly high risk of adverse economic conditions. Also consistent with FCT, we find evidence of mechanisms varying across places, such that education-health association exists across all 50 states even though the mechanisms that help explain the association vary in importance across states.

Implications for reducing health disparities between states and individuals

Our findings suggest that strategies to weaken educational disparities in health, and improve overall levels of health, might benefit by incorporating both national and state-level elements. Income may be one of the national elements. Regardless

of states' political orientation, demographic composition, macroeconomic conditions, or any other characteristic, income was a central mechanism linking higher education to better health. This suggests that improving opportunities for higher incomes (e.g., raising the federal minimum wage to keep up with inflation) among college nongraduates may be a first-order strategy for reducing health inequalities. Also relevant is our finding that the largest educational disparities in health are in the Midwest and South, where nongraduates are especially disadvantaged in income and employment. This suggests that the largest reductions in educational disparities in health may come from prioritizing improvements in economic conditions in these parts of the country. Our findings also suggest that certain state-specific strategies may be beneficial. For example, in states like Washington and Minnesota, a focus on health behaviors may be key, while in states like Ohio and Pennsylvania, a two-pronged approach that focuses on both economic conditions and health behaviors may be required (as evident from Figure 4).

Limitations and future research

Despite the strengths of the data and analysis, the study has some limitations. First, our data are cross-sectional and lack retrospective information about respondents' lives, so we cannot assert a temporal order between education, the mechanisms, and health. It is possible that the order is reversed for some respondents. For instance, unfavorable health in childhood can truncate schooling, an effect that may be most severe in states lacking educational supports and other compensatory resources. This may exacerbate the magnitude of educational disparities in health, as states where poor childhood health presents major obstacles to obtaining higher levels of education are likely to be the same states where higher levels of education are immensely important for obtaining health enhancing resources The lack of retrospective information also means that the mechanisms only reflect the time of survey. Having information such as employment and marital histories may have helped account for more of the education-health association. Moreover, our short time series did not allow us to examine temporal mechanism swapping. We were only able to assess how mechanisms varied across place, not across time. Second, our study was not designed to assess causality. We do not claim that education caused the mechanisms which, in turn, caused health. We rely on existing literature using causal methods [e.g., (21)] that identified effects of education on health-related outcomes to judiciously interpret our findings.

It is also important to consider that the BRFSS sampling frame excludes incarcerated persons. Because incarcerated adults tend to have relatively low levels of education and poorer health, the size of the education-health association in states with high incarceration rates, including many states in the South and Midwest (22), may be underestimated. Thus, our findings may be even more pronounced if the BRFSS contained incarcerated persons, because states with large educational disparities in health tend to be those with high incarceration rates. Incarceration may also operate as a mechanism through which low education results in poor health, given the many pernicious downstream consequences of current and former incarceration on employment, income, families, social ties, health, and more.

Another potential shortcoming is that our study lacked information on immigration and interstate migration. The proportion of a state's population who are immigrants could affect our findings, given that the education-health association is weaker for some immigrant groups and their health tends to be more favorable than US-born individuals. Supplementary analyses provide some assurance that our findings are not materially affected. Specifically, there is little correlation between the percentage of immigrants in a state and the strength of the state's education-health associations ($r = -0.15$, $p = 0.30$) and there is a small and non-significant correlation between the percentage of immigrants and the contribution of the two dominant mechanisms, employment and income, to the association. Interstate migration could potentially affect our findings to the extent that education or health influences interstate migration. Although we do not rule out this possibility, findings from other studies suggest that it does not materially alter our findings. For example, one study of the education-disability association across states showed the cross-state pattern persisted after limiting the sample to non-movers (7), another study concluded that interstate migration of less- or more- educated adults does not explain the growing health divides across states (23), and a third study found that the benefits of education for health were mainly shaped by their adulthood contexts, not their childhood contexts (24).

This study laid a foundation for a new line of research on why the salience of educational attainment for health differs across U.S. states. It borrowed an approach from decades of research on why the salience of education for health has grown over time (9), which has examined how the purported mechanisms linking education to health have been changing over time. Rather than examining how mechanisms change over time, we examined how they differed across place. Even though these previous studies and the current study examined mechanisms, the ultimate goal is to uncover clues about the structural level factors that made those mechanisms salient and have the potential to reduce health disparities. Given the prime role of employment and income in accounting for the largest disparities in Midwestern and Southern states that we identified, an important next step is to investigate the structural factors (e.g., states' minimum wage levels, earned income tax credits, paid leave laws) that lie at the root of

the disparities and patterns that we documented. Ultimately, reducing the disparities will require identifying and addressing these structural factors.

It may also be informative to examine additional mechanisms, such as occupation, drug use, non-familial social relationships, exposure to discrimination, and lifetime exposures to the mechanisms. It may also be fruitful to examine how and why the education-health association differs across states for specific demographic subgroups (e.g., for gender and race/ethnic groups), and how and why the association differs across local areas. Our study is a first step toward a better understanding of how geographic contexts shapes the importance of one's education level for their health.

Conclusions

Not having a 4-year college degree is much riskier for health in some U.S. states than in others. It is especially risky in Southern and Midwestern states. In general, these states have the largest educational disparities in health and the lowest overall levels of health: these states are especially disadvantaged. Meaningful reductions in educational disparities in health, and overall improvements in health, may come from prioritizing access to employment and livable income among adults without a 4-year college degree, particularly in Southern and Midwestern states.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: www.cdc.gov/brfss/index.html.

Author contributions

JKM developed the idea for the project and wrote the first draft of the manuscript. KJC conducted the statistical analyses, created the graphics, and edited the final version of the manuscript. Both authors contributed to the article and approved the submitted version.

Funding

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Numbers R01AG055481 and R24AG045061.

Acknowledgments

We appreciate comments on earlier drafts of this manuscript from members of the project Educational Attainment, Geography, and U.S. Adult Mortality Risk, members of the Policy, Place, and Population Health Lab at Syracuse University, and members of the NIA-funded Network on Life Course Health Dynamics and Disparities in 21st Century America.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.966434/full#supplementary-material

References

1. Zajacova A, Lawrence EM. The relationship between education and health: reducing disparities through a contextual approach. *Annu Rev Public Health.* (2018) 39:273–89. doi: 10.1146/annurev-publhealth-031816-044628

2. Hummer RA, Lariscy JT. Educational Attainment and Adult Mortality. In: Rogers RG, Crimmins EM, editors. *International Handbook of Adult Mortality.* Springer (2011). p. 241–61. doi: 10.1007/978-90-481-9996-9_12

3. Zajacova A, Hummer RA, Rogers RG. Education and health among U. S working-age adults: a detailed portrait across the full educational-attainment spectrum. *Biodemography Soc Biol.* (2012) 58:40–61. doi: 10.1080/19485565.2012.666122

4. Case A, Deaton A. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *PNAS.* (2015) 112:15078–83. doi: 10.1073/pnas.1518393112

5. Case A, Deaton A. Life expectancy in adulthood is falling for those without a BA degree, but as educational gaps have widened, racial gaps have narrowed. *PNAS.* (2021) 118:e2024777118. doi: 10.1073/pnas.2024777118

6. Montez JK, Hayward MD, Zajacova A. Educational disparities in U. S adult health: US states as institutional actors on the association. *Socius.* (2019) 5:1–14. doi: 10.1177/2378023119835345

7. Montez JK, Zajacova A, Hayward MD. Disparities in disability by educational attainment across U. S states. *Am J Public Health.* (2017) 107:1101–8. doi: 10.2105/AJPH.2017.303768

8. Montez JK, Zajacova A, Hayward MD, Woolf S, Chapman D, Beckfield J. Educational disparities in adult mortality across U.S. states: how do they differ and have they changed since the mid-1980s? *Demography.* (2019) 56:621–44. doi: 10.1007/s13524-018-0750-z

9. Clouston SAP, Link BG. A retrospective on fundamental cause theory: state of the literature and goals for the future. *Annu Rev Sociol.* (2021) 47:131–56. doi: 10.1146/annurev-soc-090320-094912

10. Link BG, Phelan J. Social Conditions as Fundamental Causes of Disease. *J Health Soc Behav.* (1995) 80—94. doi: 10.2307/2626958

11. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *J Health Soc Behav.* (2010) 51:S28–40. doi: 10.1177/0022146510383498

12. Mirowsky J, Ross CE. *Education, Social Status, and Health*. New York: Aldine de Gruyter (2003).

13. Torpey E, Watson A. *Education Level and Jobs: Opportunities by State: U.S. Bureau of Labor Statistics* (2014). Available online at: https://www.bls.gov/careeroutlook/2014/article/education-level-and-jobs.htm (accessed July 20, 2022).

14. MMWR. Methodologic changes in the Behavioral Risk Factor Surveillance System in 2011 and potential effects on prevalence estimates. *Morb Mortal Wkly Rep.* (2012) 61:410–3. Available online at: https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6122a3.htm (accessed January 15, 2022).

15. Schnittker J, Bacak V. The increasing predictive validity of self-rated health. *PLoS ONE.* (2014) 9:e84933. doi: 10.1371/journal.pone.0084933

16. Case A, Deaton A. *Deaths of Despair And the Future of Capitalism*. Princeton: Princeton University Press (2020). doi: 10.1515/9780691217062

17. Sasson I, Hayward MD. Association between educational attainment and causes of death among white and black US adults, 2010-2017. *JAMA.* (2019) 322:756–63. doi: 10.1001/jama.2019.11330

18. Hest R. Four methods for calculating income as a percent of the federal poverty guideline (FPG) in the Behavioral Risk Factor Surveillance System (BRFSS). In: Center SHADA, editor. Robert Wood Johnson Foundation (2019).

19. Hargrove TW, Gaydosh L, Dennis AC. Contextualizing educational disparities in health: Variations by race/ethnicity, nativity, and county-level characteristics. *Demography.* (2022) 59:267–92. doi: 10.1215/00703370-9664206

20. Karlson KB, Holm A, Breen R. Comparing regression coefficients between same-sample nested models using logit and probit: a new method. *Sociol Methodol.* (2012) 42:286–313. doi: 10.1177/0081175012444861

21. Halpern-Manners A, Helgertz J, Warren JR, Robert E. The Effects of Education on Mortality: Evidence From Linked U.S. Census and Administrative Mortality Data. *Demography.* (2020) 57:1513–41. doi: 10.1007/s13524-020-00892-6

22. Widra E, Herring T. *States of Incarceration: The Global Context 2021*. Prison Policy Initiative (2021). Available online at: https://www.prisonpolicy.org/global/2021.html (accessed July 21, 2022).

23. Couillard BK, Foote CL, Gandhi K, Meara E, Skinner J. Rising geographic disparities in US mortality. *J Econ Perspect.* (2021) 35:1–27. doi: 10.1257/jep.35.4.123

24. Kemp BR, Montez JK. Why does the importance of education for health differ across the United States? *Socius.* (2020) 6:1–16. doi: 10.1177/2378023119899545

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 07 September 2022
DOI 10.3389/fpubh.2022.942842

OPEN ACCESS

EDITED BY
Shannon Monnat,
Syracuse University, United States

REVIEWED BY
Andrew Stokes,
Boston University, United States
Dana Glei,
Georgetown University, United States

*CORRESPONDENCE
Arun S. Hendi
arun.hendi@princeton.edu

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 13 May 2022
ACCEPTED 16 August 2022
PUBLISHED 07 September 2022

CITATION
Hendi AS and Ho JY (2022) Smoking and the widening inequality in life expectancy between metropolitan and nonmetropolitan areas of the United States.
Front. Public Health 10:942842.
doi: 10.3389/fpubh.2022.942842

Smoking and the widening inequality in life expectancy between metropolitan and nonmetropolitan areas of the United States

Arun S. Hendi[1]* and Jessica Y. Ho[2]

[1]Office of Population Research and Department of Sociology, Princeton University, Princeton, NJ, United States, [2]Department of Sociology and Criminology and Population Research Institute, The Pennsylvania State University, University Park, PA, United States

Background: Geographic inequality in US mortality has increased rapidly over the last 25 years, particularly between metropolitan and nonmetropolitan areas. These gaps are sizeable and rival life expectancy differences between the US and other high-income countries. This study determines the contribution of smoking, a key contributor to premature mortality in the US, to geographic inequality in mortality over the past quarter century.

Methods: We used death certificate and census data covering the entire US population aged 50+ between Jan 1, 1990 and Dec 31, 2019. We categorized counties into 40 geographic areas cross-classified by region and metropolitan category. We estimated life expectancy at age 50 and the index of dissimilarity for mortality, a measure of inequality in mortality, with and without smoking for these areas in 1990–1992 and 2017–2019. We estimated the changes in life expectancy levels and percent change in inequality in mortality due to smoking between these periods.

Results: We find that the gap in life expectany between metros and nonmetros increased by 2.17 years for men and 2.77 years for women. Changes in smoking-related deaths are responsible for 19% and 22% of those increases, respectively. Among the 40 geographic areas, increases in life expectancy driven by changes in smoking ranged from 0.91 to 2.34 years for men while, for women, smoking-related changes ranged from a 0.61-year decline to a 0.45-year improvement. The most favorable trends in years of life lost to smoking tended to be concentrated in large central metros in the South and Midwest, while the least favorable trends occurred in nonmetros in these same regions. Smoking contributed to increases in mortality inequality for men aged 70+, with the contribution ranging from 8 to 24%, and for women aged 50–84, ranging from 14 to 44%.

Conclusions: Mortality attributable to smoking is declining fastest in large cities and coastal areas and more slowly in nonmetropolitan areas of the US. Increasing geographic inequalities in mortality are partly due to

these geographic divergences in smoking patterns over the past several decades. Policies addressing smoking in non-metropolitan areas may reduce geographic inequality in mortality and contribute to future gains in life expectancy.

KEYWORDS

mortality, life expectancy, smoking, inequality, urban-rural differences

Introduction

American mortality is undergoing an unprecedented stagnation. Since 2010, life expectancy gains have been among the slowest on record for the US, and life expectancy declined for three consecutive years between 2014 and 2017. Between 2010 and 2018, life expectancy increased by less than a tenth of a year (1–3). Considerable geographic variation underlies these national-level trends, with poor performance concentrated in nonmetropolitan areas and specific regions of the country. Some parts of the country—coastal areas and big cities—continue to post robust increases in life expectancy, while others—rural areas and the South and Appalachia—experience much slower rates of improvement (4, 5).

The divide between metropolitan and nonmetropolitan areas has grown considerably between 1990 and the present. While American cities of the early 1990s faced a number of social and economic dilemmas that limited their capacities for promoting healthy and long lives, the situation today is quite different. While they still face difficulties relating primarily to issues of equality, metropolitan areas today tend to have better outcomes along several dimensions, including educational attainment, public health infrastructure and outreach, and economic activity (4). Nonmetropolitan areas have experienced either slower improvement or deterioration along these same dimensions (6). In short, metropolitan areas have prospered while nonmetropolitan areas have been left behind, and this has been manifested in widening metro/nonmetro gaps in mortality.

Mortality inequalities are among the starkest manifestations of inequity in our society. Prior research suggests geographic inequality in mortality has increased over time, and that these inequalities have reached substantial magnitudes (4, 7–11). Where people live influences what policies are in place, their access to and quality of health care, the social and economic conditions they experience, and what health behaviors they practice. These differences are the most commonly proposed explanations for geographic inequalities and their growth over time (4, 7–12).

Cigarette smoking, the leading cause of premature morbidity and mortality in the United States, is one potential explanation that reflects all of the above dimensions. There is a vast literature that relates smoking to elevated levels of mortality and large and growing inequalities in mortality along a number of dimensions (13–20), and smoking is known to have contributed to past mortality variation among states. However, there has been relatively little research on the role of smoking-attributable mortality in explaining the metro-nonmetro divergence in U.S. mortality since the early 1990s. Figure 1 shows that while ever smoking and current smoking prevalence were similar between metros and nonmetros in the early 1990s, they have diverged significantly since then. Recent findings that cardiovascular disease, respiratory diseases, and lung cancer are among the key causes of death contributing to rising geographic inequality in mortality (10) also support the hypothesis that smoking likely plays a role in the growth in metro-nonmetro inequality in mortality over the last quarter century.

This study investigates the contribution of smoking to geographic inequality in mortality over the past three decades. Because of the rapidly growing gap between metropolitan and nonmetropolitan areas, we examine mortality inequalities across four metropolitan categories: large central metros, large metro suburbs, small and medium metros, and nonmetropolitan areas. We additionally explore whether the contribution of smoking to metro/nonmetro mortality inequalities is reproduced across ten regions of the country, since smoking-attributable mortality is known to have a strong regional component (21). These analyses shed light on how differences in smoking are contributing to divergent patterns of life expectancy gains across the nation, and in particular to the adverse mortality trends concentrated in nonmetropolitan areas.

Materials and methods

Data

We used the 1990–2019 National Center for Health Statistics (NCHS) Multiple Cause of Death data files, consisting of all deaths occurring in the US. The files contain information on decedents' age, sex, cause of death, and county of residence. These data were combined with Census population estimates to produce all-cause and lung cancer (ICD-9 code 162 for 1990–1998 and ICD-10 codes C33-C34 for 1999–2019) death rates by age, sex, geographic area (described below), and year. We considered four analytic periods: 1990–1992, 2000–2002, 2010–2012, and 2017–2019, with most analyses focusing on the first and last period. We focus on these specific

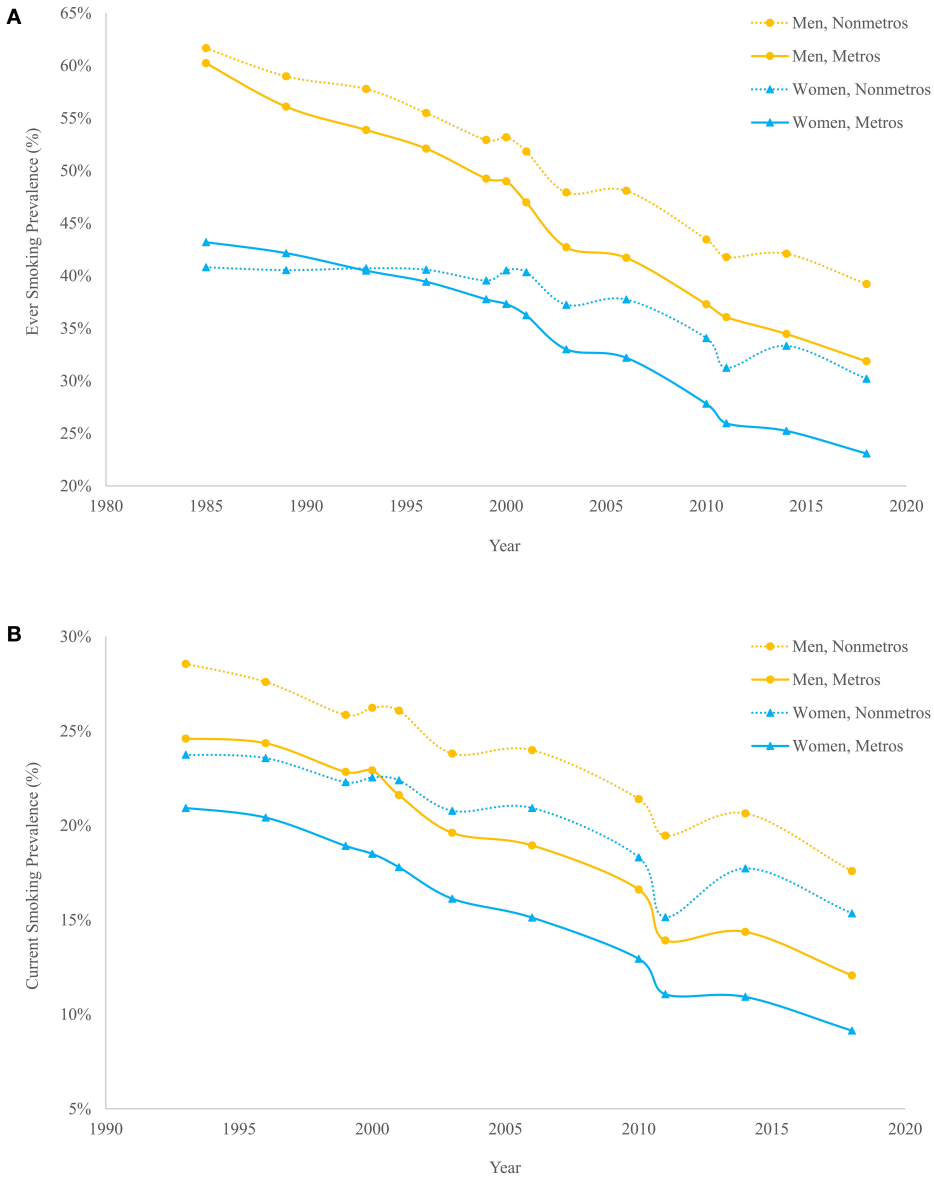


FIGURE 1
Ever smoking **(A)** and current smoking **(B)** prevalence by metropolitan status and sex, 1985–2018. Estimates are authors' calculations based on Current Population Survey Tobacco Use Supplements, 1985–2018. All estimates are based on the population aged 18 and older and are standardized to the 2000 U.S. population age distribution.

periods because the metro-nonmetro mortality divergence commenced in the early 1990s (4, 7). Supplementary analyses (Supplementary Tables S1, S2; Supplementary Figure S1) show that our conclusions hold whether we use either 1990–1992 or 2000–2002 as the baseline period.

Both region and metropolitan/nonmetropolitan residence are key dimensions of geographic inequalities in mortality. To classify counties into metropolitan categories, we used codes developed by the US Department of Agriculture Economic Research Service, which were modified and made available by the NCHS. We used four categories: large central metros, large metro suburbs, small/medium metros, and nonmetropolitan areas. We considered 10 regions: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, Pacific, and Appalachia. The first nine regions were defined using the Census division categorization, while Appalachian counties were defined by the Appalachian Regional Commission classification. Appalachia consists of all of West Virginia and selected counties from 12 other states, which were excluded from

their overlapping Census divisions. Supplementary Table S12 shows the correspondence between Census Region, Census Division, and state. For a subset of analyses, we cross-classified counties by region and metropolitan/nonmetropolitan category to identify 40 distinct geographic units. This 40-category classification has been used in prior studies and captures important features of geographic variation in mortality (4, 10).

Analytic approach

Smoking is causally linked to many chronic diseases, including cardiovascular diseases, respiratory diseases, and cancers (22). In order to capture the total burden of smoking-related mortality, we used an indirect estimation method fitted to US data (21, 23). The method uses excess lung cancer mortality as an indicator of the damage caused by smoking, where excess lung cancer mortality is calculated as the difference between observed lung cancer mortality and the level of lung cancer mortality we would expect to observe among non-smokers. The method then models all-cause mortality as a function of excess lung cancer mortality to produce estimates of the proportion of deaths attributable to smoking by age and sex. The method was developed for ages 50+ because the mortality impacts of smoking manifest primarily at these ages (23). According to the 2018 US life table, 93.9% of Americans can expect to survive to age 50, so this analysis covers the great majority of deaths in the population (2).

The main assumption of this method is that lung cancer mortality accurately proxies for the cumulative burden of smoking, which is likely to be the case since the majority of lung cancer deaths in industrialized societies are attributable to smoking (24). Because lung cancer mortality reflects multiple forms of tobacco smoking, including cigarette and cigar smoking, this assumption means that the indirect estimates capture the broader impact of smoking across multiple product classes. The method's key advantages are that it captures the total burden of smoking-related mortality and relies on vital statistics data rather than self-reported smoking data, which is subject to reporting biases, may not accurately reflect individuals' lifetime smoking histories, and often results in underestimates of smoking-related mortality. We provide additional detail on the methodological approach in the Appendix.

Other studies have used direct approaches to examine smoking-attributable mortality, typically regressing mortality on smoking status to obtain relative risks. While the fine geographic detail used in this study does not allow for replication using direct methods applied to public-use data, prior research covering similar time periods focusing on other subpopulations has found qualitatively similar trends when applying either direct or indirect methodologies (25).

We estimated smoking-attributable mortality by age, sex, period, and geographic area. We computed life expectancy at age 50 with and without smoking using life table techniques and examined the contribution of smoking to life expectancy at age 50 in 1990–1992 and 2017–2019 and the change in its contribution for both the four metro categories and the 40 areas defined above.

We used two measures to quantify inequality and smoking's effect on inequality in mortality. For analyses focusing on the four metro categories, we computed a gradient measure equal to the difference in life expectancy at age 50 between large central metros and nonmetros, both with and without smoking-attributable mortality. We also assessed the contribution of smoking to the change over time in the gradient. In ancillary analyses (not shown here), we computed the difference between nonmetro areas and both large metro suburbs and small/medium metros and found qualitatively similar results.

As a summary measure of mortality inequality across the 40 cross-classified areas, we computed the index of dissimilarity (ID), one of the most commonly used measures of spatial unevenness. We calculated the ID for mortality with and without smoking-attributable deaths in each period to determine how much smoking contributes to geographic inequality in each period and to changes in geographic inequality over time.

The ID is calculated as:

$$ {}_nID_x = \frac{1}{2}\sum_{i=1}^{N}\left|\frac{{}_nD_x^i}{{}_nD_x} - \frac{{}_nP_x^i}{{}_nP_x}\right| $$

where $\frac{{}_nD_x^i}{{}_nD_x}$ is the proportion of national deaths at ages x to $x+n$ occurring in place i, $\frac{{}_nP_x^i}{{}_nP_x}$ is the proportion of the national population aged x to $x+n$ that lives in place i, and N is the total number of places ($N = 40$ geographic units). The ID has previously been used to study residential segregation, occupational and social mobility, and geographic inequality in mortality (10, 26, 27). Its value ranges between 0 and 1, with 0 indicating absolute equality and 1 absolute inequality. Among the ID's useful properties are that it is symmetric, invariant to population size, and easily interpreted as the proportion of national deaths that would need to be reallocated to a different area to achieve geographic equality in mortality. The ID was calculated for each 5-year age group between 50–54 and 80–84, and for an open-ended age group (85+).

Results

Over the past three decades, gains in life expectancy have differed dramatically between metropolitan and nonmetropolitan areas within the US. Metro areas of all types, particularly large central metros, experienced much more rapid gains in life expectancy than nonmetros. Men experienced sizeable life expectancy increases between 1990–1992 and 2017–2019, gaining 4.64 years in large central metros, 3.92 years in large metro suburbs, 3.10 years in small metros, and

2.47 years in nonmetros (Table 1). Women had more modest gains in life expectancy, with values ranging from 0.65 years in nonmetros to 3.42 years in large central metros.

Changes in smoking-attributable mortality are part of the explanation for both the sex and the metro/nonmetro divergences (Table 1). Years of life lost due to smoking were sizeable for men in 1990–1992, ranging from 2.88 to 3.15 years. However, these values were quite similar across metro categories. By 2017–2019, years of life lost due to smoking had declined significantly, but the declines were more rapid in metropolitan areas. In 2017–2019, men lost between 1.30 years to smoking in large central metros and 1.86 years to smoking in nonmetros. Over this period, changes in smoking-attributable mortality were responsible for as much as 1.70 years of the 4.64-year life expectancy gain in large central metros and as little as 1.29 years of the 2.47-year life expectancy gain in nonmetros. All three types of metropolitan areas—large central metros, large metro suburbs, and small metros—experienced more rapid improvements from declines in smoking-related mortality than nonmetros.

The story for women is quite different. Unlike men, women initially experienced a reverse metro/nonmetro gradient in smoking-attributable mortality. In 1990–1992, women in nonmetros lost the fewest years of life to smoking (1.26 years). Women in large central metros lost the most years of life to smoking (1.59 years). By 2017–2019, this situation completely reversed, so that women in nonmetros lost the most years to smoking (1.58 years) while women in large central metros lost the fewest years to smoking (1.30 years). A very clear gradient emerged, wherein nonmetros experienced the greatest increase in years of life lost due to smoking, while the remaining three metropolitan areas experienced either no change or decreases in years of life lost to smoking.

The differential patterns in years of life lost to smoking by metro category contributed to an increase in inequality as measured by the difference in life expectancy between large central metros and nonmetros. The gradient increased by 2.17 years for men, and 19% of that increase was due to smoking. For women, the gradient between large central metros and nonmetros increased by 2.77 years, and 22% of that increase was due to smoking.

Considering finer geographic areas, we see that reductions in smoking-attributable mortality contributed to sizeable gains in life expectancy among men (Figure 2). This trend is evident across all regions and metropolitan/nonmetropolitan categories. The largest life expectancy improvements related to smoking occurred in large central metros in the Southern regions—East South Central (2.34 years), West South Central (1.99 years), and South Atlantic (1.94 years) —and in large metro suburbs in West South Central (2.00 years) and East South Central (1.98 years). The smallest improvements related to smoking were recorded in four nonmetropolitan areas: the Appalachian (1.19 years), East North Central (1.03 years), Mountain (1.00 years), and West North Central regions (0.91 years). Smoking-related improvements in male life expectancy at age 50 were most pronounced in large central metros and least pronounced in nonmetropolitan areas in the majority (six of the ten) of the regions. This pattern of differential life expectancy gains has contributed to a divergence between metropolitan and nonmetropolitan areas.

Life expectancy at age 50 increased for women across all geographic areas; however, most of these increases would have been considerably larger had smoking-attributable mortality not increased. We refer to this negative effect of smoking on longevity as smoking-related declines in life expectancy. Most regions experienced smoking-related declines in female life expectancy. The most notable exception to this pattern is the Pacific region, which experienced smoking-related improvements in life expectancy, with the largest improvement in large central metros (0.39 years) and the smallest improvement in nonmetros (0.12 years). Most nonmetros and small metros experienced smoking-related declines in female life expectancy between 1990–1992 and 2017–2019. These declines were most notable in nonmetropolitan areas of the East South Central (−0.61 years), West North Central (−0.49 years), East North Central (−0.45 years), and Appalachian (−0.43 years) regions. In contrast, nearly all large central metros and some large metro suburbs experienced life expectancy gains due to changes in smoking. These improvements were, however, much smaller than those among men, ranging from 0.01 to 0.45 years.

Changes in the index of dissimilarity

Geographic inequality in mortality increased at all ages above 50 between 1990–1992 and 2017–2019 (Figure 3). Geographic inequality tends to be highest at younger ages but increased more rapidly over time at the older ages. Among men, the ID increased by between 7 and 49% at ages 50–69, while increases in the ID ranged from 63 to 103% for men aged 70 and older. For women, inequality nearly doubled at ages 50–69, and more than doubled at ages 70 and older. The increase in inequality was thus more pronounced for women across all ages. Whereas, geographic inequality was significantly lower for women than for men in the early 1990s, by 2017–2019, inequality was higher for women than for men at each age except 85+.

Table 2 shows values for the ID in 1990–1992 and 2017–2019, the contribution of smoking to the ID, and the contribution of smoking to the change over time in the ID. For men, the smoking contribution was highest in the early 1990s, ranging from 32 to 38% between ages 55–74 (column 5). While the percent contribution of smoking to the ID decreased over time, it still remained sizeable in 2017–2019, ranging from 20 to 24% for ages 55–74 (column 6). For women, there was a

TABLE 1 Life expectancy at age 50 with and without smoking-attributable mortality by sex and metropolitan category, 1990–2019.

	Men								
	1990–1992			2017–2019			Change		
	Obs	NS	YLL	Obs	NS	YLL	Obs	NS	YLL
Large central metro	26.41	29.40	2.99	31.05	32.34	1.30	4.64	2.94	−1.70
Large metro suburb	27.31	30.19	2.88	31.23	32.57	1.34	3.92	2.38	−1.54
Small/Medium metro	26.85	29.91	3.06	29.95	31.51	1.56	3.10	1.60	−1.50
Nonmetro	26.34	29.49	3.15	28.80	30.66	1.86	2.47	1.17	−1.29
Gradient	0.08	−0.08		2.24	1.68		2.17	1.77	
	(0.00, 0.15)	(−0.18, 0.01)		(2.17, 2.32)	(1.59, 1.77)		(2.07, 2.27)	(1.64, 1.89)	
Contribution of smoking to widening of gradient							**19%**		
	Women								
	1990–1992			2017–2019			Change		
	Obs	NS	YLL	Obs	NS	YLL	Obs	NS	YLL
Large central metro	31.51	33.11	1.59	34.93	36.24	1.30	3.42	3.13	−0.29
Large metro suburb	31.99	33.58	1.59	34.53	35.99	1.46	2.53	2.40	−0.13
Small/Medium metro	31.98	33.42	1.44	33.58	35.02	1.44	1.60	1.60	0.00
Nonmetro	31.84	33.11	1.26	32.49	34.07	1.58	0.65	0.97	0.32
Gradient	−0.33	0.00		2.44	2.16		2.77	2.16	
	(−0.41, −0.25)	(−0.09, 0.09)		(2.36, 2.52)	(2.07, 2.25)		(2.66, 2.88)	(2.04, 2.29)	
Contribution of smoking to widening of gradient								**22%**	

Obs = observed life expectancy at age 50; NS = life expectancy at age 50 if smoking-related mortality is eliminated; YLL = years of life expectancy at age 50 lost due to smoking-related deaths; and Gradient = difference in life expectancy at age 50 between the large central metro and nonmetro categories. Values in brackets below Gradient values are 95% confidence intervals. Standard errors for Obs and NS values are given in Supplementary Table S3.

complete reversal in the percent contribution of smoking to the ID. In the 1990s, smoking contributed negatively to geographic inequality at all ages (smoking deaths reduced geographic inequality either because they were more evenly distributed than other causes of death or because they were negatively correlated with other causes of death across areas). By 2017–2019, smoking contributed to inequality in every age group except 85+, ranging between 11 and 18% at ages 50–79 (column 6). The smoking contribution tends to be larger at ages 50–79 and diminishes at the older ages. In both periods, smoking contributes to greater inequality for men than for women.

Next, we assess the contribution of smoking deaths to changes over time in the ID between 1990–1992 and 2017–2019 (column 7). A negative value indicates that smoking tended to decrease inequality, while a positive value indicates that smoking deaths increased inequality. Changes in smoking-attributable mortality contributed to decreased inequality among men aged 50–69. At ages 70+, smoking contributed to increased geographic inequality for men. Its contribution was particularly large for the 85+ age group, where it accounted for 24% of the increase in inequality (column 7). In contrast, smoking contributed to increased geographic inequality for women in all age groups except 85+. Smoking was responsible for between 27 and 44% of the increase in inequality for women aged 50–79 (column 7).

Discussion

Concerns about mortality stagnation and growing social and economic disparities have generated renewed interest in geographic inequality in American mortality. In this study, we find that smoking has become a major contributor to metropolitan-nonmetropolitan inequality, accounting for approximately one-fifth of the widening of the gap in life expectancy between large central metros and nonmetropolitan areas. This pattern holds for both men and women. This finding builds on prior research showing that geographic inequality in mortality increased in recent decades due to a range of causes of death and that smoking plays a role in regional differences in mortality (5, 10, 21, 28). We further show that the uneven geographic distribution of changes in smoking-attributable mortality is driving a substantial portion of the increase in inequality in life expectancy in the US. Large central metros and their suburbs are reaping the benefits of rapid

FIGURE 2
Contribution of the change in smoking-attributable mortality to the change in life expectancy at age 50 for men **(A)** and women **(B)** by region and metropolitan category, 1990–92 to 2017–19. Positive/negative values indicate that changes in smoking-attributable mortality contributed to an increase/decrease in life expectancy. These values are YLLs as defined in Table 1. Standard errors are given in Supplementary Tables S5, S6.

reductions in smoking-attributable deaths, while nonmetros are being left behind.

These differential patterns have important implications for geographic inequality, which has increased substantially over the past three decades. Smoking-attributable deaths are responsible for roughly one-third of the increase in geographic inequality in mortality for women aged 50–79 when considering geographic units defined by regions cross-classified by metro category. For men, 10–24% of the increase among those aged 75+ was due to smoking. While smoking-attributable deaths constitute a larger proportion of overall deaths for men than for women, there was a greater geographic divergence in smoking patterns for women that drove their larger increase in inequality. This sex differential in the effects of smoking is largely a result of the differential patterns of smoking initiation and cessation for men vs. women. Men's smoking peaked in the 1950s, while women's smoking

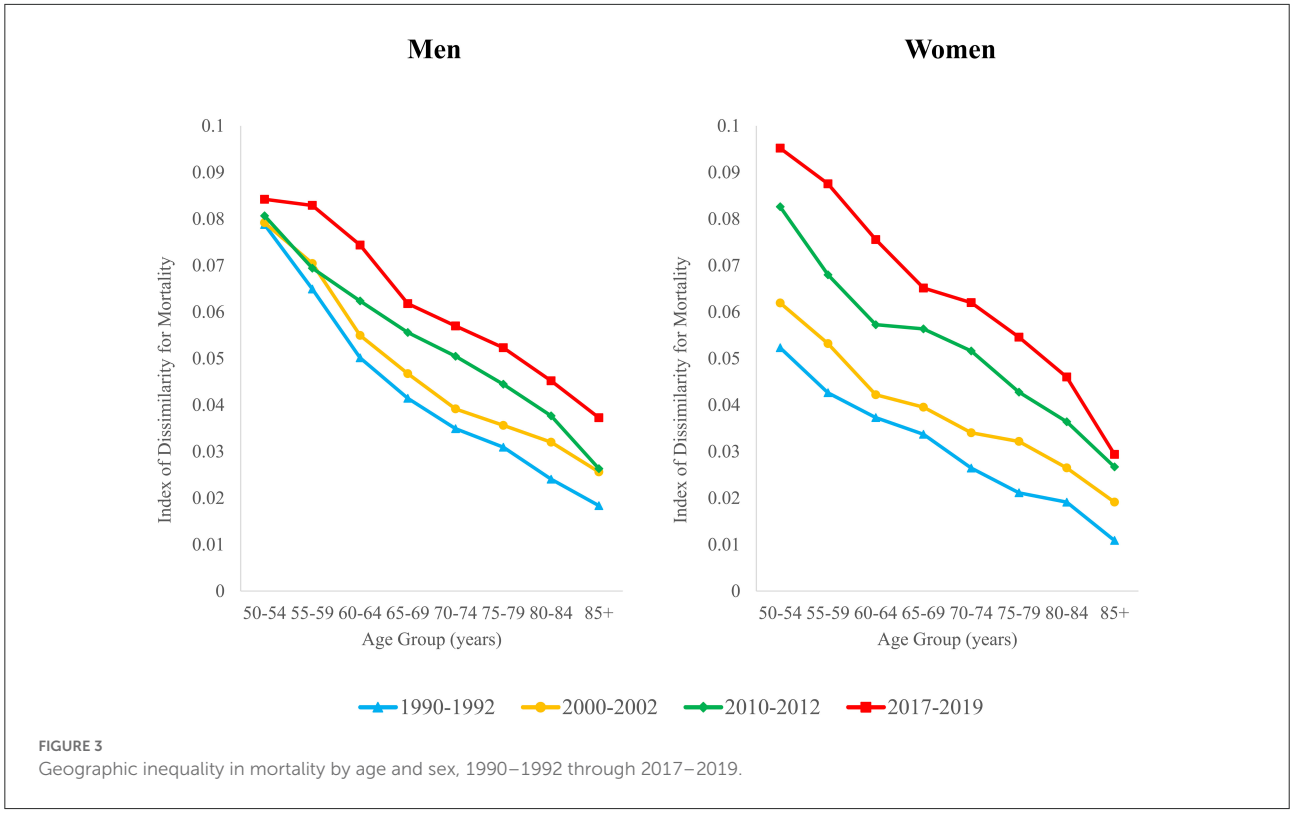


FIGURE 3
Geographic inequality in mortality by age and sex, 1990–1992 through 2017–2019.

TABLE 2 Contribution of smoking-related deaths to geographic inequality in mortality by sex and age, 1990–2019.

		Index of dissimilarity for mortality (ID)		Contribution of smoking to ID		% contribution of smoking to ID		% change in ID due to smoking[a]
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Age group	1990–1992	2017–2019	1990–1992	2017–2019	1990–1992	2017–2019	1990–2019
Men	50–54	0.079	0.084	0.019	0.012	24%	14%	−117% (−141, −94%)
	55–59	0.065	0.083	0.023	0.017	35%	20%	−33% (−37, −29%)
	60–64	0.050	0.074	0.018	0.015	36%	21%	−11% (−12, −9%)
	65–69	0.041	0.062	0.016	0.015	38%	24%	−4% (−5, −2%)
	70–74	0.035	0.057	0.011	0.013	32%	23%	8% (6, 9%)
	75–79	0.031	0.052	0.007	0.009	21%	17%	10% (9, 11%)
	80–84	0.024	0.045	0.003	0.005	10%	12%	13% (12, 14%)
	85+	0.018	0.037	0.000	0.004	−3%	11%	24% (23, 25%)
Women	50–54	0.052	0.095	−0.001	0.011	−2%	12%	28% (25, 31%)
	55–59	0.043	0.088	−0.004	0.016	−9%	18%	44% (42, 46%)
	60–64	0.037	0.076	−0.003	0.010	−7%	13%	33% (32, 34%)
	65–69	0.034	0.065	−0.001	0.007	−3%	11%	27% (25, 28%)
	70–74	0.026	0.062	−0.004	0.008	−14%	13%	32% (32, 33%)
	75–79	0.021	0.055	−0.003	0.006	−14%	11%	27% (26, 28%)
	80–84	0.019	0.046	−0.002	0.002	−9%	5%	14% (13, 15%)
	85+	0.011	0.029	−0.001	−0.001	−5%	−5%	−5% (−5, −4%)

[a]Percent change in ID between 1990–1992 and 2017–2019 due to smoking-attributable deaths, calculated as [(4) – (3)]/[(2) – (1)]. Values in parentheses are 95% confidence intervals.

peaked approximately 15 years later (29). Thus, men's smoking-attributable mortality peaked in the 1990s and women's in the 2000s.

Much of the recent literature on mortality inequality has focused on contemporaneous phenomena as determinants of inequality. Smoking, on the other hand, is an exposure whose mortality effects accumulate over time and manifest decades later, often with a 20- to 40-year lag. The increases in geographic inequality in mortality today are due to changes in smoking behaviors that largely took place in the 1980s through the early 2000s. Even as smoking-attributable mortality declines, the uneven geographic patterning of those declines has contributed to growing geographic inequality.

While the popularity of cigarette smoking for the nation as a whole followed a pattern of rapid uptake followed by decline over the course of the 20th century, this process has occurred unevenly within the country across regions and between metro and nonmetro areas. The earliest data on smoking patterns date from the mid-1950s and suggest that smoking prevalence was higher in metropolitan than nonmetropolitan areas. At this time, the difference in smoking prevalence between urban and rural farm residents was around 11% points for men (52 vs. 41%) and 16% points among women (26 vs. 10%) (30, 31). Regional variation in smoking prevalence was fairly muted among men during this period, although heavy smoking was most common in the Northeast (31, 32). By the mid-1980s, however, smoking had become heavily concentrated in the South among men. The East South Central (35.8%), South Atlantic (34.4%), and West South Central (33.9%) regions had the highest percentage of men who were current smokers, while the Pacific region had the lowest percentage (27.7%) (33). Among women, regional differences were smaller. Smoking prevalence was highest in East North Central (27.5%), New England (25.9%), and the South Atlantic (25.3%), and lowest in the Pacific (22.4%) (33). There was also a reversal in the metro/nonmetro gradient as smoking prevalence declined in metro areas but either increased or stayed the same in nonmetro areas from the mid-1990s through the early 2000s (34). Between the mid-2000s and 2014, smoking declined in nonmetro areas but at much slower rates than in urban areas (30). For the past decade, current smokers in nonmetro areas and regions of the South have been more likely to have begun smoking at earlier ages (i.e., younger than age 16) and to smoke more cigarettes per day (35, 36). Today, nonmetros and parts of the South and Midwest are regarded as lagging far behind the rest of the nation in terms of their progress in reducing smoking and smoking-attributable mortality.

Several factors are thought to contribute to these patterns. These include: fewer and later adoption of tobacco control policies; the countering influence of the tobacco industry, particularly in tobacco-growing areas concentrated in the South; limited access to smoking cessation programs and interventions; and socioeconomic conditions of these areas. Tobacco control policies encompass a spectrum of policies such as excise taxes, media campaigns, and restrictions on smoking in public places and have been found to be effective in reducing smoking prevalence (37). However, nine of the ten states with the lowest excise taxes in 2011 were located in the Midwest (East and West North Central) and the South (South Atlantic, West South Central, and East South Central) (38, 39). Of the 24 states that lacked a comprehensive smoke-free law as of 2015, 17 were located in the South and Midwest (39). Studies have also suggested that tobacco control policies may be much more restricted in scope and less intense in nonmetropolitan areas (30, 40).

Weak tobacco regulations, particularly in tobacco-growing areas in the South, are thought to be related to their history of economic dependence on tobacco coupled with intensive tobacco industry influence. Nonmetros in these areas are viewed as having been particularly dependent on tobacco, and positive attitudes toward tobacco and smoking have persisted (30). Studies have found that in major tobacco growing regions, opposition to smoke-free laws and cigarette taxes was concentrated among tobacco farmers, hospitality associations, and tobacco companies. Tobacco companies sought to promote a pro-tobacco culture and block tobacco-control policies dating from the 1960s and continuing through the 1990s. These efforts included mobilizing farmers growing flue-cured tobacco in the South to block cigarette tax increases, highlighting the benefits that tobacco has brought to these economies, and emphasizing the threat tobacco-control policies pose to farmers and tax revenues (41). One example comes from the Philip Morris publication Smokers Advocate, which included the following as part of an "action alert" to oppose a cigarette tax hike in 1990: "At a time when tobacco is increasingly under attack throughout the rest of the country, North Carolinians need to "circle the wagons" and protect the economic future of as important a crop as tobacco" (42). The RJ Reynolds company created a "Pride in Tobacco" program in the late 1970s that focused on opposing tobacco-control policies in North Carolina, South Carolina, Wisconsin, Ohio, Kentucky, Virginia, and Tennessee. It continued operating through the 1990s (41). As a result, tobacco-growing parts of the South have been much slower to adopt tobacco-control policies, and when they do adopt them, they are more limited (e.g., less comprehensive coverage of workplaces, restaurants, and bars and lower taxes).

Both smoking initiation and cessation influence the risk of dying from a smoking-related cause of death and a population's level of smoking-attributable mortality. Nonmetro areas and regions of the South and Midwest have experienced poor socioeconomic conditions, in part related to deindustrialization. Studies have highlighted that lower education levels and knowledge of the health risks of smoking may be more prevalent in these areas (30). Low absolute and relative levels of education have been tied to high levels of mortality in the U.S. (43, 44). Poor socioeconomic conditions and daily life stressors may lead to smokers continuing to smoke as a form of stress relief (30).

Qualitative studies of rural areas have documented a lack of support for quitting smoking within rural social networks (45). They have also found that the lack of alternative activities in nonmetro areas coupled with few public smoking bans and exposure to other smokers leads to both smoking initiation and continued smoking (45, 46). Economic constraints and limited access to smoking cessation programs and interventions also pose barriers to smoking cessation in these areas. Coverage for smoking cessation treatment services remains low, and some rural smokers have reported perceiving that buying cigarettes is less expensive than purchasing smoking cessation aides (34, 45). Smokers in nonmetropolitan areas may face particular challenges due to lack of smoking cessation programs in their local area, lack of mass media messaging about smoking prevention and treatment, and lack of knowledge of existing resources (34, 45). This is reflected in low use rates of smoking cessation aides such as nicotine lozenges, inhalers, or sprays or smoking cessation counseling in rural areas (45, 47).

Another class of explanations for the diverging life expectancy trends driven by smoking is selection. The populations of metropolitan and nonmetropolitan parts of the country have undergone significant change related to selective migration. More educated, healthier, well-to-do individuals have tended to leave nonmetropolitan areas in favor of large metros and their suburbs, meaning that those left behind in nonmetropolitan areas are likely to be negatively select on these same characteristics. Since education, underlying health, and income and wealth all tend to be negatively associated with both smoking and mortality, this form of selective migration is likely to lead to faster improvements in metropolitan life expectancy and either slower improvements or worsening of life expectancy in nonmetropolitan areas. Cigarette smoking uptake also tends to be concentrated in the teen years, so one's childhood place of residence may matter just as much as where one currently resides. Because of the likelihood that selective migration may be driving some of the trends documented in this study, the results cannot be interpreted as indicative of current place of residence driving 100% of the observed trends. Rather, a host of factors, including migration histories of a place's current population, determines mortality trends.

The main strengths of our study include the use of death certificate data covering the entire US population and the use of an indirect estimation method that captures the full burden of mortality associated with cigarette smoking. There are also several limitations to our study. It is possible that the relationship between lung cancer and all-other-cause mortality has changed over time, which would alter our results. This would be possible if, for example, mortality from causes unrelated to smoking has decreased over time, leading to a tighter, more positive relationship between lung cancer and all-other-cause mortality. We compute ancillary estimates taking into account this change and find that it only minimally influences our findings and does not change our substantive conclusions. Another potential concern is that our study focuses on ages 50+ and thus excludes smoking-related deaths below age 50. While prior research has shown that the smoking-attributable fraction is highly similar for ages 35+ relative to ages 50+ (23), we cannot rule out that smoking may also be important in explaining geographic variation in mortality below age 50. Estimates of the effect of smoking on life expectancy at birth that do not take into account smoking-attributable under-50 mortality are reported for the various geographic units in Supplementary Tables S7, S8. A third limitation is that we use only two measures of inequality: the metro/nonmetro life expectancy gap and the index of dissimilarity. It is possible that other measures of geographic inequality may yield different estimates. Supplementary analyses (Supplementary Table S9) indicate that our conclusions hold whether we use the index of dissimilarity or other measures of inequality, including the Gini coefficient and Theil's index. Finally, this article examines how increasing inequality is tied to smoking and does not examine how the contribution of smoking to geographic inequality might be related to racial and socioeconomic inequalities identified in other studies (25, 48–50). These social inequalities may act as mechanisms linking smoking and geographic inequality in mortality, as more vulnerable groups tend to have higher smoking-related mortality and are more concentrated in high-mortality regions. We find that in some regions of the country, rural areas lag behind in efforts to reduce smoking-attributable mortality. If, in those regions, racial and ethnic minorities are disproportionately concentrated in rural areas, we may expect within-region racial/ethnic disparities to persist or widen. The impacts of widening urban-rural inequality on racial/ethnic disparities and vice versa are nevertheless difficult to predict, since the composition of these areas has also changed over time, likely in a manner that is selective on latent traits predictive of mortality.

In debates surrounding inequality and mortality, researchers have often cast increasing inequality as a natural consequence of improvements in life expectancy. The most advantaged are able to reap the benefits of new knowledge and technologies, which in turn leads to increased inequality (51). What this study adds to the existing literature is identification of metropolitan status as a key dimension along which inequalities in smoking-attributable mortality have emerged over the past three decades. Differences between metropolitan and nonmetropolitan areas are complex and not easily captured by socioeconomic variables alone. Metropolitan status is a distinctly place-based categorization that encompasses differences between areas in their demographic, socioeconomic, environmental, cultural, and health system characteristics (52). For example, the legacy of economic dependence on tobacco and intensive tobacco industry influence has contributed to positive attitudes toward smoking and slower and more limited adoption of tobacco control policies in tobacco-growing nonmetro areas in the South (30, 40, 41). It is not simply that people in nonmetropolitan areas

are poorer or less educated, but rather that the characteristics of these places themselves may lead to a greater burden of smoking-attributable mortality.

The results of this paper suggest that a number of policies can be implemented that would both increase life expectancy and reduce geographic inequality. Cigarette taxes tend to be higher in regions like the Northeast, which are also the areas where smoking-attributable mortality has declined the most. They tend to be lowest in states with large rural populations. In additional analyses (Supplementary Table S10), we find that states with lower cigarette tax rates experienced a greater metro-nonmetro divergence over time in years of life lost to smoking relative to states with higher taxes. Implementing higher cigarette taxes in areas like the South and the Midwest has the potential to reduce geographic inequality and metro/nonmetro inequality in mortality and to contribute to further gains in life expectancy (39, 53). Another potential set of policies encompasses comprehensive smoke-free laws for public areas. States that have not adopted these laws also tend to be concentrated in the South and hold a disproportionate share of the rural population (39). Similarly, tobacco retailer density and tobacco marketing has become more concentrated in rural parts of the country (54, 55). States with large rural populations can implement policies that would restrict retail tobacco growth, which would likely have the effect of decreasing nonmetropolitan smoking rates at the national level. Given the lag between smoking initiation or cessation and the mortality effects of smoking, the impacts of instituting any of these policies on reducing inequality would play out in the decades following the implementation of the policies.

While some policies like cigarette taxation tend to have the effect of reducing inequalities, others tend to do the opposite. This may be because of differential implementation, enforcement, and access to resources that make these programs less effective in nonmetropolitan areas. For example, one Kentucky-based study showed that smoke-free laws had different impacts on air quality due to differential enforcement (56). Nonmetros tend to have fewer smoking cessation programs and interventions, and tobacco control policies tend to be more restricted in scope in these areas (30, 40, 57). This would suggest that the federal and state governments should explore the possibility of targeting policies and smoking cessation resources specifically toward nonmetropolitan areas in order to reduce the disproportionate burden of smoking-attributable mortality in nonmetros.

Though the imprint of cigarette smoking on mortality is diminishing, new substances have emerged with the potential to drive new health inequalities. According to the 2020 National Youth Tobacco Survey, one-fifth of high school students are current users of e-cigarettes, up from roughly one-tenth in 2017 (58, 59). E-cigarette use is more common in rural areas at the national level, though there are important regional variations (60). The long-term health impacts of e-cigarette use are not yet well-established, and it is possible e-cigarettes could become new sources of premature mortality and inequalities in mortality in future decades. Other substances, like marijuana delivered through e-cigarettes, could also have long-term effects on mortality. On the other hand, these products may displace traditional cigarettes and thus have countervailing effects on smoking-attributable mortality (61, 62). Future trends in geographic inequality in mortality may be shaped by these new health behaviors, much as today's trends in inequality were partly shaped by the smoking behavior of cohorts in decades past. The findings of this paper and the emergence of these new technologies underscore the need for continued monitoring and coordinated efforts to prevent the uptake of potentially-deleterious health behaviors.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary materials, further inquiries can be directed to the corresponding author/s.

Author contributions

AH and JH designed the study, analyzed and interpreted the data, and drafted and revised the manuscript. AH was responsible for submitting the manuscript for publication. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by grants from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R00 HD083519 and P2C HD047879) and the National Institute on Aging (R01 AG060115).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those

of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.942842/full#supplementary-material

References

1. Arias E. United States Life Tables, 2010. *Natl Vital Stat Rep.* (2014) 63:1–63.

2. Arias E, Xu J. United States Life Tables, 2018. *Natl Vital Stat Rep.* (2020) 69:1–45. doi: 10.15620/cdc:101128

3. Ho JY, Hendi AS. Recent trends in life expectancy across high income countries: retrospective observational study. *BMJ.* (2018) 362:k2562. doi: 10.1136/bmj.k2562

4. Elo IT, Hendi AS, Ho JY, Vierboom YC, Preston SH. Trends in non-hispanic white mortality in the United States by metropolitan-nonmetropolitan status and region, 1990–2016. *Popul Dev Rev.* (2019) 45:549–83. doi: 10.1111/padr.12249

5. Fenelon A. Geographic divergence in mortality in the United States. *Popul Dev Rev.* (2013) 39:611–34. doi: 10.1111/j.1728-4457.2013.00630.x

6. Probst J, Eberth JM, Crouch E. Structural urbanism contributes to poorer health outcomes for rural America. *Health Aff.* (2019) 38:1976–84. doi: 10.1377/hlthaff.2019.00914

7. Cossman JS, James WL, Cosby AG, Cossman RE. Underlying causes of the emerging nonmetropolitan mortality penalty. *Am J Public Health.* (2010) 100:1417–9. doi: 10.2105/AJPH.2009.174185

8. Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, Morozoff C, Mackenbach JP, van Lenthe FJ, et al. Inequalities in life expectancy among Us Counties, 1980 to 2014: temporal trends and key drivers. *JAMA Intern Med.* (2017) 177:1003–11. doi: 10.1001/jamainternmed.2017.0918

9. Singh GK, Siahpush M. Widening rural-urban disparities in life expectancy, U.S., 1969-2009. *Am J Prev Med.* (2014) 46:e19–29. doi: 10.1016/j.amepre.2013.10.017

10. Vierboom YC, Preston SH, Hendi AS. Rising geographic inequality in mortality in the United States. *SSM Popul Health.* (2019) 9:100478. doi: 10.1016/j.ssmph.2019.100478

11. Wang H, Schumacher AE, Levitz CE, Mokdad AH, Murray CJ. Left behind: widening disparities for males and females in Us county life expectancy, 1985-2010. *Popul Health Metr.* (2013) 11:8. doi: 10.1186/1478-7954-11-8

12. Montez JK, Zajacova A, Hayward MD. Explaining Inequalities in Women's Mortality between US States. *SSM Popul Health.* (2016) 2:561–71. doi: 10.1016/j.ssmph.2016.07.004

13. Hiscock R, Bauld L, Amos A, Fidler JA, Munafò M. Socioeconomic status and smoking: a review. *Ann N Y Acad Sci.* (2012) 1248:107–23. doi: 10.1111/j.1749-6632.2011.06202.x

14. Kanjilal S, Gregg EW, Cheng YJ, Zhang P, Nelson DE, Mensah G, et al. Socioeconomic status and trends in disparities in 4 major risk factors for cardiovascular disease among Us Adults, 1971-2002. *Arch Intern Med.* (2006) 166:2348–55. doi: 10.1001/archinte.166.21.2348

15. Khang YH, Lynch JW, Jung-Choi K, Cho HJ. Explaining age-specific inequalities in mortality from all causes, cardiovascular disease and ischaemic heart disease among South Korean male public servants: relative and absolute perspectives. *Heart.* (2008) 94:75–82. doi: 10.1136/hrt.2007.117747

16. Khang YH, Lynch JW, Yang S, Harper S, Yun SC, Jung-Choi K, et al. The contribution of material, psychosocial, and behavioral factors in explaining educational and occupational mortality inequalities in a nationally representative sample of South Koreans: relative and absolute perspectives. *Soc Sci Med.* (2009) 68:858–66. doi: 10.1016/j.socscimed.2008.12.003

17. Kivimäki M, Shipley MJ, Ferrie JE, Singh-Manoux A, Batty GD, Chandola T, et al. Best-practice interventions to reduce socioeconomic inequalities of coronary heart disease mortality in Uk: a prospective occupational cohort study. *Lancet.* (2008) 372:1648–54. doi: 10.1016/S0140-6736(08)61688-8

18. Thomas S, Fayter D, Misso K, Ogilvie D, Petticrew M, Sowden A, et al. Population tobacco control interventions and their effects on social inequalities in smoking: systematic review. *Tob Control.* (2008) 17:230–7. doi: 10.1136/tc.2007.023911

19. Danaei G, Rimm EB, Oza S, Kulkarni SC, Murray CJ, Ezzati M. The promise of prevention: the effects of four preventable risk factors on national life expectancy and life expectancy disparities by race and county in the United States. *PLoS Med.* (2010) 7:e1000248. doi: 10.1371/journal.pmed.1000248

20. Jha P, Peto R, Zatonski W, Boreham J, Jarvis MJ, Lopez AD. Social inequalities in male mortality, and in male mortality from smoking: indirect estimation from National Death Rates in England and Wales, Poland, and North America. *Lancet.* (2006) 368:367–70. doi: 10.1016/S0140-6736(06)68975-7

21. Fenelon A, Preston SH. Estimating smoking-attributable mortality in the United States. *Demography.* (2012) 49:797–818. doi: 10.1007/s13524-012-0108-x

22. US Department of Health and Human Services. *The Health Consequences of Smoking-—50 Years of Progress: A Report of the Surgeon General.* Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention (2014).

23. Preston SH, Glei DA, Wilmoth JR. A new method for estimating smoking-attributable mortality in high-income countries. *Int J Epidemiol.* (2010) 39:430–8. doi: 10.1093/ije/dyp360

24. Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. *Lancet.* (2003) 362:847–52. doi: 10.1016/S0140-6736(03)14338-3

25. Ho JY, Elo IT. The contribution of smoking to black-white differences in US. *Mortal Demogr.* (2013) 50:545–68. doi: 10.1007/s13524-012-0159-z

26. Duncan OD, Duncan B. Residential distribution and occupational stratification. *Am J Sociol.* (1955) 60:493–503. doi: 10.1086/221609

27. Massey DS, Denton NA. The dimensions of residential segregation. *Social Forces.* (1988) 67:281–315. doi: 10.2307/2579183

28. Vierboom YC, Preston SH. Life beyond 65: changing spatial patterns of survival at older ages in the United States, 2000–2016. *Jo Gerontol Series B.* (2020) 75:1093–103. doi: 10.1093/geronb/gbz160

29. Burns DM, Lee L, Shen LZ, Gilpin E, Tolley HD, Vaughn J, et al. *Cigarette Smoking Behavior in the United States. Changes in Cigarette-Related Disease Risks and Their Implication for Prevention and Control.* Bethesda, MD: U.S. Department of Health and Human Services; National Institutes of Health, National Cancer Institute (1997). p. 13-42.

30. Doogan NJ, Roberts ME, Wewers ME, Stanton CA, Keith DR, Gaalema DE, et al. A growing geographic disparity: rural and urban cigarette smoking trends in the United States. *Prev Med.* (2017) 104:79–85. doi: 10.1016/j.ypmed.2017.03.011

31. Haenszel W, Shimkin MB, Miller HP. Tobacco smoking patterns in the United States. *Public Health Monogr.* (1956) 45:1–105.

32. Devesa SS, Grauman DJ, Blot WJ, Fraumeni JF, Jr. Cancer surveillance series: changing geographic patterns of lung cancer mortality in the United States, 1950 through 1994. *J Natl Cancer Inst.* (1999) 91:1040-50. doi: 10.1093/jnci/91.12.1040

33. Marcus AC, Shopland DR, Crane LA, Lynn WR. Prevalence of cigarette smoking in the United States: estimates from the 1985 current population survey. *J Natl Cancer Inst.* (1989) 81:409–14. doi: 10.1093/jnci/81.6.409

34. Doescher MP, Jackson JE, Jerant A, Gary Hart L. Prevalence and trends in smoking: a national rural study. *J Rural Health.* (2006) 22:112–8. doi: 10.1111/j.1748-0361.2006.00018.x

35. Schoenborn CA, Adams PF, Barnes PM, Vickerie JL, Schiller JS. Health behaviors of adults: United States, 1999-2001. *Vital Health Stat.* (2004) 10:1–79. doi: 10.1037/e378902004-001

36. Schoenborn CA, Adams PF, Peregoy JA. Health behaviors of adults: United States, 2008-2010. *Vital Health Stat.* (2013) 10:1–184.

37. Farrelly MC, Pechacek TF, Thomas KY, Nelson D. The impact of tobacco control programs on adult smoking. *Am J Public Health.* (2008) 98:304–9. doi: 10.2105/AJPH.2006.106377

38. Cokkinides V, Bandi P, McMahon C, Jemal A, Glynn T, Ward E. Tobacco control in the United States–Recent Progress and Opportunities. *CA Cancer J Clin.* (2009) 59:352–65. doi: 10.3322/caac.20037

39. Tynan MA, Holmes CB, Promoff G, Hallett C, Hopkins M, Frick B. State and local comprehensive smoke-free laws for worksites, restaurants, and bars - United States, 2015. *Morb Mortal Wkly Rep.* (2016) 65:623–6. doi: 10.15585/mmwr.mm6524a4

40. York NL, Rayens MK, Zhang M, Jones LG, Casey BR, Hahn EJ. Strength of tobacco control in rural communities. *J Rural Health.* (2010) 26:120–8. doi: 10.1111/j.1748-0361.2010.00273.x

41. Fallin A, Glantz SA. Tobacco-control policies in tobacco-growing states: where tobacco was king. *Milbank Q.* (2015) 93:319–58. doi: 10.1111/1468-0009.12124

42. Philip Morris USA. *Action Alert: Legislators Likely to Call for Tobacco Tax Hikes to Bail out State Economy*. New York, NY: Smokers' Advocate (1990).

43. Hendi AS. Trends in education-specific life expectancy, data quality, and shifting education distributions: a note on recent research. *Demography.* (2017) 54:1203–13. doi: 10.1007/s13524-017-0574-2

44. Hendi AS. Trends in U.S. life expectancy gradients: the role of changing educational composition. *Int J Epidemiol.* (2015) 44:946–55. doi: 10.1093/ije/dyv062

45. Hutcheson TD, Greiner KA, Ellerbeck EF, Jeffries SK, Mussulman LM, Casey GN. Understanding smoking cessation in rural communities. *J Rural Health.* (2008) 24:116–24. doi: 10.1111/j.1748-0361.2008.00147.x

46. Spivak AL, Givel MS, Monnat SM. Self-interest and public opinion in health policy: smoking behavior and support for tobacco control. *Soc Theory Health.* (2018) 16:20–43. doi: 10.1057/s41285-017-0041-6

47. Wewers ME, Stillman FA, Hartman AM, Shopland DR. Distribution of daily smokers by stage of change: current population survey results. *Prev Med.* (2003) 36:710–20. doi: 10.1016/S0091-7435(03)00044-6

48. Hendi AS, Elo IT, Martikainen P. The implications of changing education distributions for life expectancy gradients. *Soc Sci Med.* (2021) 272:113712. doi: 10.1016/j.socscimed.2021.113712

49. Ho JY, Fenelon A. The contribution of smoking to educational gradients in US life expectancy. *J Health Soc Behav.* (2015) 56:307–22. doi: 10.1177/0022146515592731

50. Denney JT, Rogers RG, Hummer RA, Pampel FC. Education inequality in mortality: the age and gender specific mediating effects of cigarette smoking. *Soc Sci Res.* (2010) 39:662–73. doi: 10.1016/j.ssresearch.2010.02.007

51. Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav.* (1995) 80–94. doi: 10.2307/2626958

52. Moy E, Garcia MC, Bastian B, Rossen LM, Ingram DD, Faul M, et al. Leading causes of death in nonmetropolitan and metropolitan areas- United States, 1999-2014. *MMWR Surveill Summ.* (2017) 66:1–8. doi: 10.15585/mmwr.ss6601a1

53. Bowser D, Canning D, Okunogbe A. The impact of tobacco taxes on mortality in the USA, 1970-2005. *Tob Control.* (2016) 25:52–9. doi: 10.1136/tobaccocontrol-2014-051666

54. Hall J, Cho HD, Maldonado-Molina M, George TJ, Shenkman EA, Salloum RG. Rural-urban disparities in tobacco retail access in the Southeastern United States: Cvs Vs. The Dollar Stores. *Prev. Med. Reports.* (2019) 15:100935. doi: 10.1016/j.pmedr.2019.100935

55. Henriksen L, Schleicher NC, Johnson TO, Roeseler A, Zhu SH. Retail tobacco marketing in rural versus nonrural counties: product availability, discounts, and prices. *Health Promot Pract.* (2020) 21(1_suppl.):27s-36s. doi: 10.1177/1524839919888652

56. Lee K, Hahn EJ, Pieper N, Okoli CT, Repace J, Troutman A. Differential impacts of smoke-free laws on indoor air quality. *J Environ Health.* (2008) 70:24–30, 54.

57. American Lung Association. *Cutting Tobacco's Rural Roots: Tobacco Use in Rural Communities*. Washington, DC (2012).

58. Gentzke AS, Wang TW, Jamal A, Park-Lee E, Ren C, Cullen KA, et al. Tobacco product use among middle and high school students - United States, 2020. *Morb Mortal Wkly Rep.* (2020) 69:1881–8. doi: 10.15585/mmwr.mm6950a1

59. Wang TW, Gentzke A, Sharapova S, Cullen KA, Ambrose BK, Jamal A. Tobacco product use among middle and high school students - United States, 2011-2017. *Morb Mortal Wkly Rep.* (2018) 67:629–33. doi: 10.15585/mmwr.mm6722a3

60. Mumford EA, Stillman FA, Tanenbaum E, Doogan NJ, Roberts ME, Wewers ME, et al. Regional rural-urban differences in e-cigarette use and reasons for use in the United States. *J Rural Health.* (2019) 35:395–404. doi: 10.1111/jrh.12333

61. National Academies of Sciences E-cigarettes and Medicine. *Public Health Consequences of E-Cigarettes*. Washington, DC: The National Academies Press (2018).

62. Selya AS, Foxon F. Trends in electronic cigarette use and conventional smoking: quantifying a possible 'diversion' effect among Us adolescents. *Addiction.* (2021) 116:1848–58. doi: 10.1111/add.15385

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fpubh.2022.993507

OPEN ACCESS

EDITED BY
Shannon Monnat,
Syracuse University, United States

REVIEWED BY
Binod Acharya,
Drexel University, United States
Federico Benassi,
Italian National Institute of Statistics (ISTAT), Italy

*CORRESPONDENCE
Tse-Chuan Yang
tyang3@albany.edu

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 13 July 2022
ACCEPTED 07 September 2022
PUBLISHED 26 September 2022

CITATION
Yang T-C, Shoff C, Choi S-wE and Sun F (2022) Multiscale dimensions of county-level disparities in opioid use disorder rates among older Medicare beneficiaries.
Front. Public Health 10:993507.
doi: 10.3389/fpubh.2022.993507

Multiscale dimensions of county-level disparities in opioid use disorder rates among older Medicare beneficiaries

Tse-Chuan Yang[1]*, Carla Shoff[2], Seung-won Emily Choi[3] and Feinuo Sun[4]

[1]Department of Sociology, University at Albany, State University of New York, Albany, NY, United States, [2]Independent Consultant, Baltimore, MD, United States, [3]Department of Sociology, Anthropology, and Social Work, Texas Tech University, Lubbock, TX, United States, [4]Global Aging and Community Initiative, Mount Saint Vincent University, Halifax, NS, Canada

Background: Opioid use disorder (OUD) among older adults (age ≥ 65) is a growing yet underexplored public health concern and previous research has mainly assumed that the spatial process underlying geographic patterns of population health outcomes is constant across space. This study is among the first to apply a local modeling perspective to examine the geographic disparity in county-level OUD rates among older Medicare beneficiaries and the spatial non-stationarity in the relationships between determinants and OUD rates.

Methods: Data are from a variety of national sources including the Centers for Medicare & Medicaid Services beneficiary-level data from 2020 aggregated to the county-level and county-equivalents, and the 2016–2020 American Community Survey (ACS) 5-year estimates for 3,108 contiguous US counties. We use multiscale geographically weighted regression to investigate three dimensions of spatial process, namely "level of influence" (the percentage of older Medicare beneficiaries affected by a certain determinant), "scalability" (the spatial process of a determinant as global, regional, or local), and "specificity" (the determinant that has the strongest association with the OUD rate).

Results: The results indicate great spatial heterogeneity in the distribution of OUD rates. Beneficiaries' characteristics, including the average age, racial/ethnic composition, and the average hierarchical condition categories (HCC) score, play important roles in shaping OUD rates as they are identified as primary influencers (impacting more than 50% of the population) and the most dominant determinants in US counties. Moreover, the percentage of non-Hispanic white beneficiaries, average number of mental health conditions, and the average HCC score demonstrate spatial non-stationarity in their associations with the OUD rates, suggesting that these variables are more important in some counties than others.

Conclusions: Our findings highlight the importance of a local perspective in addressing the geographic disparity in OUD rates among older adults.

Interventions that aim to reduce OUD rates in US counties may adopt a place-based approach, which could consider the local needs and differential scales of spatial process.

KEYWORDS

opioid use disorder, multiscale geographically weighted regression, spatial heterogeneity, geographic disparity, county

Introduction

Life expectancy in the United States (US) has lagged behind other developed countries (1, 2). Since 2014, the US has witnessed a slight decrease in life expectancy, a phenomenon that is not observed in any other countries (3) before the COVID-19 pandemic. It has been suggested that the opioid epidemic contributes to this public health concern (4, 5) and several scholars have investigated the determinants of opioid-related deaths in US counties (6–8). However, little attention has been paid to opioid-related outcomes among older adults (age $\geq$ 65), which increasingly contribute to the ongoing opioid epidemic (9). One study uses Medicare data between 2013 and 2018 and reported that the prevalence of opioid use disorder (OUD) among older adults has increased by more than 3-fold. Specifically, there were approximately 4.6 OUD cases per 1,000 beneficiaries in 2013 and this overall prevalence soared to 15.7 in 2018. The elevating trend is universal across all racial/ethnic, gender, and socioeconomic groups (10).

There are three major reasons why older adults are vulnerable to OUD. First, due to the aging process, older adults are more likely to suffer from physical pain and mental illness than younger populations (11, 12). As such, older adults are frequent recipients of prescription opioids to manage their health conditions. As exposure to prescription opioids is positively associated with the development of OUD (13), older adults are likely to develop a particularly high risk of OUD over time. Second, older adults' vulnerability to OUD, due to declining health conditions, may be further compounded by social and psychological risk factors associated with life course events (e.g., retirement and bereavement), such as social isolation, depression, and helplessness (9, 12, 14). These factors may aggravate the risk of OUD. Third, compared with younger adults, older adults are less likely to realize the negative consequences of opioid use (15) and are more likely to overlook OUD symptoms due to fear of substance-use stigma (11). Importantly, baby boomers (70–80 million people) are generally more tolerant or accepting of recreational substance use (11) than other generations, which likely makes baby boomers have lower perceived risk for not taking opioids as prescribed by their doctors.

Despite the unique vulnerability and challenges faced by older adults, little research has investigated the determinants of OUD until recently. Applying negative binomial regression to a 2017 county-level dataset, a study (16) finds that in a county, the number of older Medicare beneficiaries with OUD is associated with not only the beneficiaries' characteristics (e.g., average age), but also a county's socioeconomic conditions (e.g., social isolation). An individual-level study reports similar associations in that the risk of OUD is higher among socioeconomically marginalized older adults and those who reside in socially isolated and disadvantaged areas (17). Although these findings shed some light on the extant literature, the following gaps remain. First, the spatial distribution of OUD rates among older adults in US counties is unknown and the question of whether there is a geographic disparity in OUD rates across space has not been investigated. Second, previous research has suggested that spatial heterogeneity exists in county-level drug-overdose mortality and some scholars have explored this topic (7, 8). Nonetheless, it is unknown whether spatial heterogeneity is also embedded in the county-level patterns of OUD rates among older adults. Finally, most prior studies adopt a global modeling perspective to understand how OUD rates are shaped by other factors. This global modeling perspective assumes that the spatial process that leads to the observed ecological data is homogeneous but this assumption has been found to be unrealistic in empirical research (18). No prior research has applied a local modeling perspective to the research of OUD rates among older adults.

This study aims to fill these gaps by applying multiscale geographically weighted regression (MGWR) to a dataset of 3,108 contiguous US counties and county-equivalents and investigating three dimensions of spatial process, namely *level of influence*, *scalability*, and *specificity* (details in the next section). MGWR is a recently developed spatial analysis method that allows researchers to explore spatial non-stationarity (19, 20) and the three dimensions are drawn from the strengths of this local analysis perspective (21).

Materials and methods

Data sources and measures

This study assembles the analytical dataset from multiple national sources and focuses on the counties in the contiguous US (N = 3,108). The data from the Centers for Medicare

& Medicaid Services (CMS) include beneficiary-level data from 2020 that is drawn from 3 CMS data files: (i) the Medicare Beneficiary Summary File (MBSF) Base segment, (ii) MBSF Chronic Conditions segment, and (iii) MBSF Other Chronic and Potentially Disabling Conditions Segment. The data have been limited to those beneficiaries who are 65 years of age or older and who are continuously enrolled in Medicare Fee-for-Services Parts A, B, and D for all 12 months of the 2020 calendar year and for all 12 months of 2019. Continuous enrollment for the previous data year is necessary due to the lookback period used to construct the OUD flag (discussed below). The beneficiary-level data are aggregated to the county-level based on county the beneficiary lives. The 2016–2020 American Community Survey (ACS) 5-year estimates (22) serve as the major source for the county-level socioeconomic features.

The dependent variable is the *OUD rate among older Medicare beneficiaries* (per 1,000 beneficiaries), which is defined as the total number of beneficiaries with OUD divided by the total number of beneficiaries in a county. OUD is defined using the overarching opioid use disorder flag that focuses on three opioid-related sub-indicators: (i) diagnosis and procedure code basis for OUD with at least one inpatient claim or two other non-drug claims of any service type with valid International Classification of Diseases, Tenth Revision (ICD-10) diagnosis codes or Current Procedural Terminology (CPT) or Healthcare Common Procedure Coding System (HCPCS) procedure codes, (ii) opioid-related hospitalization or Emergency Department visits, and (iii) use of medication assisted treatment (23). Furthermore, the following Medicare beneficiary characteristics are created at the county-level. *Percentage of female beneficiaries* is calculated by dividing the total number of female beneficiaries by the total number of beneficiaries. The *average age of beneficiaries* (in years) in a county is calculated. *Percentage of non-Hispanic white beneficiaries*, *percentage of non-Hispanic black beneficiaries*, and *percentage of Hispanic beneficiaries* are measured by dividing the number of beneficiaries in each racial/ethnic group by the total number of beneficiaries. Socioeconomically marginalized older adults may be eligible for both Medicaid and Medicare, which is known as dual-eligibility status. This study divides the total number of beneficiaries with dual-eligibility by the total number of beneficiaries to obtain the *percentage of dually eligible beneficiaries*.

The *average number of mental health conditions* is the mean value of beneficiaries' mental health conditions, including anxiety disorders, depressive disorders, bipolar disorder, and schizophrenia and other psychotic disorders. Consistent with defining OUD using the chronic condition and other chronic or potentially disabling condition flags, these mental health conditions are determined using the condition specific flags, which flag the beneficiary as having the condition during the calendar year if they meet the condition specific diagnosis or procedure code basis for that condition.[1] Similarly, the *average number of physical conditions* refers to the mean value of beneficiaries' physical conditions including chronic obstructive pulmonary disease, diabetes, chronic kidney disease, and hypertension, ranging from 0 to 4. These physical conditions are also determined using the condition specific flags. The final beneficiary characteristic is the *average hierarchical condition category (HCC) score*. CMS develops an algorithm to calculate a beneficiary's potential Medicare cost. The HCC score is normalized to 1 and a beneficiary with a score that is <1 is less costly than a beneficiary with a score that is >1 (24).

With respect to socioeconomic features of a county, three composite variables are constructed with the ACS 5-year estimates. Principal component analysis (PCA) is first applied to the following four variables and the PCA score is used to gauge the *social isolation index* among older adults: percentage of older adults with a disability; percentage of older adults who were divorced, separated, or widowed; percentage of older adults having difficulty living independently; and percentage of older adults living in poverty. Each variable has a factor loading higher than 0.65 and more than 60 percent of the total variation can be explained by the first principal component. This social isolation index is designed by the United Health Foundation (25) and has been recently used in opioid-related research (17). Higher values indicate higher levels of social isolation among older adults in a county. In addition, following previous research (26), this study creates the *concentrated disadvantage index* by applying PCA to five variables: logged median family income; unemployment rate; percentage of families headed by women; percentage of the population age 25 and older without a high school degree; and percentage of households receiving public assistance (i.e., cash payments including Temporary Assistance to Needy Families and General Assistance). This measure of concentrated disadvantage focuses on the general population and higher PCA scores reflect stronger concentrated disadvantage. The factor loadings of the five variables are >0.55 and approximately 60 percent of the total variation can be explained by the first principal component. Finally, the average of two standardized variables: percentage of owner-occupied housing units and percentage of households living in the same housing unit for at least 5 years is used to measure *residential stability*. Higher values indicate higher levels of residential stability in a county.

Statistical analysis

The multiscale geographically weighted regression (MGWR) (19) serves as the major analytic technique used in this study.

1 Information on how each of the chronic conditions and other chronic or potentially disabling conditions are defined can be found on the Chronic Conditions Data Warehouse website: https://www2.ccwdata.org/web/guest/condition-categories.

As the MGWR is an extension of GWR (27), it is important to first introduce GWR and then discuss the strengths of MGWR. According to Fotheringham and colleagues (28), a general GWR can be expressed as below (27):

$$y_i = \sum_{j=1}^{k} \beta_{ij} x_{ij} + \varepsilon_i, \quad (1)$$

Where y_i is the dependent variable for location (i.e., county in this study) $i \in \{1, 2, \ldots, n\}$, x_{ij} refers to the jth independent variable ($j \in \{1, 2, \ldots, k\}$) and β_{ij} is the estimated parameter (i.e., coefficient) for x_{ij}. ε_i is the error term. The following matrix form can be used to calibrate the GWR coefficient at each location i:

$$\hat{\beta}_{\mathbf{i}} = (\mathbf{X}^{\mathbf{T}}\mathbf{W}_{\mathbf{i}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{W}_{\mathbf{i}}\mathbf{y}, \; i \in \{1, 2, \ldots, n\}, \quad (2)$$

Where $\mathbf{X}$ is the $n^*(k+1)$ matrix of independent variables (including the intercept), and $\mathbf{y}$ is the n^*1 dependent variable vector, and $\mathbf{W_i}$ is the n^*n spatial weighting matrix for a given location i. In $\mathbf{W_i}$, the spatial weights are obtained with a specific kernel function and a bandwidth. Under the GWR framework, the bandwidth is assumed to be the same across all independent variables. That is, the relationship between an independent variable and the dependent variable operates at the same spatial scale (i.e., bandwidth) (27) and the local estimates are calibrated with this assumption.

The constant spatial scale assumption may not be realistic for two reasons. On the one hand, in empirical research, some relationships do not vary by location but others are space-dependent (29). The former is known as the global relationship, whereas the latter refers to spatial heterogeneity. When both types of relationships exist in the observed data, the constant spatial scale assumption may misestimate the local coefficients. On the other hand, the differences in culture, social structures, norms, and values across space may lead the association between an independent and a dependent variable to operate at different spatial scales. For example, when observing the spatial correlation of a variable in a smaller spatial scale (e.g., counties within a state), researchers tend to identify similarities; however, when the spatial scale becomes larger (e.g., counties across multiple adjacent states), scholars are likely to find differences (18). As such, the constant spatial scale assumption may not fully reflect the spatial data generating process underlying the observed patterns.

MGWR aims to address these methodological issues by relaxing the constant spatial scale (i.e., bandwidth) assumption and allowing variable-specific optimized bandwidth (19, 30). This is the major difference between a MGWR and a GWR model. A MGWR model can be formulated as a generalized additive model (19):

$$\mathbf{y} = \sum_{j=1}^{k} \mathbf{f_j} + \varepsilon \quad (3)$$

Where, $\mathbf{f_j}$ is a smooth function applied to the jth independent variable (31). Under the MGWR framework, each smooth function is a spatial parameter surface calculated with a bandwidth that is specific to the jth independent variable. MGWR calibrates estimates using a back-fitting algorithm (19). That is, compared with GWR, MGWR is more general and each independent variable has its own bandwidth, which forms a data generating process that allows not only global but also localized associations, which may operate at different spatial scales. It should be emphasized that MGWR standardizes all variables in the back-fitting algorithm, which facilitates the comparison of estimated coefficients across the unit of analysis. The adaptive bi-square kernel is used in this study to address the uneven spatial distribution of observations. Other technical details of MGWR can be found elsewhere (20, 28).

Dimensions of geographic disparities

Yang and colleagues have recently exploited the strengths of MGWR to investigate three dimensions of spatial process and geographic disparities (21), namely *level of influence*, *scalability*, and *specificity*. Extending the three dimensions to this study, we define the level of influence as the percentage of older Medicare beneficiaries affected by a certain independent variable across the contiguous US counties. Based on the local estimates of an independent variable, we can first identify the counties where this independent variable is statistically significant and then sum the total number of beneficiaries in these counties. We then divide the sum by the total number of beneficiaries in the entire study area. If a variable is found to influence more than 50 percent of the entire population, this variable will be categorized into the primary influencer group; otherwise (i.e., $\leq$ 50 percent), it is a secondary influencer.

Regarding *scalability*, it can be defined with the calibrated bandwidth of a variable. Scalability has three groups: global, regional, and local. According to Yang et al. (21), when a calibrated bandwidth of a variable is >75 percent of the global bandwidth (i.e., the total number of counties in this study), it can be defined as a "global" factor. If the bandwidth of a variable is between 75 and 25 percent of the global bandwidth, it is regarded as a "regional" factor. When the bandwidth of a variable is smaller than 25 percent of the global bandwidth, this variable falls into the "local" factor group. While this interquartile range approach may be arbitrary, it has been used to detect spatial non-stationarity in GWR analysis (28).

The third dimension, *specificity*, takes advantage of the standardized coefficients yielded by MGWR. In this study, each county will have its own estimates of the independent variables and they can be compared *within* each county. Such a comparison helps researchers to identify the independent variable that has the strongest association (regardless of estimated direction) with the dependent variable. That is, an independent variable may demonstrate the strongest association in some counties but not in others. We will visualize the specificity dimension to show the uniqueness of a certain variable across space. It should be noted that in a conventional ordinary least squares (OLS) regression model, the magnitude of the standardized coefficient of an independent variable increases with the variance of this independent variable. This pattern may make coefficient comparisons problematic (32). Nonetheless, this concern cannot be directly applied to the MGWR framework because each variable has its own bandwidth and the comparison is within a county or the same population (21).

Analytic strategy

This study conducts analysis in three phases: (I) conducting descriptive analysis and visualizing key variables, (II) implementing the OLS regression, which estimates the relationships between the independent variables and the dependent variable with data for all counties, and (III) using MGWR to obtain the local estimates (20). As MGWR generates abundant local parameter estimates, this study uses summary statistics and maps (33) to present the findings. Furthermore, the Monte Carlo method (20) is used to formally test whether spatial non-stationarity exists.

Results

Descriptive findings

Table 1 presents the descriptive statistics of the variables used in this study. On average, the OUD rate is 15.35 per 1,000 older Medicare beneficiaries in a county, which is comparable with the OUD rate reported in recent research (10). Regarding the county-level demographic composition of older Medicare beneficiaries, almost 60 percent of beneficiaries are female and the average age of beneficiaries is 75.83 years old. Slightly more than 88 percent of beneficiaries are non-Hispanic white, and non-Hispanic blacks and Hispanics account for 4.95 and 3.25 percent, respectively. Regarding dual-eligibility status, 15.85 percent of beneficiaries are eligible for both Medicare and Medicaid. The average numbers of mental conditions and physical conditions are 0.38 and 1.31, respectively. With respect to potential financial cost and health, the average HCC score is 1.08, indicating that the average financial burden at the county-level is greater than the population average. As the social isolation index and concentrated disadvantage index are created with PCA, they have a mean value of 0 with a standard deviation close to 1. Residential stability follows a similar pattern.

TABLE 1 Descriptive statistics of the variables used in this study ($N = 3{,}108$).

Variable	Mean	SD	Min	Max
Opioid use disorder (OUD) rate (per 1,000 beneficiaries)	15.35	10.05	0.00	148.22
Percentage of female (%)	58.21	2.37	46.08	73.33
Average age of beneficiaries (%)	75.83	0.73	72.07	79.47
Percentage of non-Hispanic (NH) white (%)	88.41	13.05	3.05	100.00
Percentage of non-Hispanic (NH) black (%)	4.95	9.36	0.00	75.17
Percentage of Hispanic (%)	3.25	8.97	0.00	96.85
Percentage of dual eligibility (%)	15.85	9.52	0.00	85.02
Average number of mental health conditions (count)	0.38	0.08	0.07	0.89
Average number of physical conditions (count)	1.31	0.22	0.50	2.08
Average hierarchical condition category (HCC) score (count)	1.08	0.12	0.61	1.83
Social isolation index	0.00	1.00	−3.24	4.94
Concentrated disadvantage index	−0.01	0.99	−2.42	6.96
Residential stability	0.01	0.88	−5.68	2.36

The spatial distribution of OUD rates among older Medicare beneficiaries (by quintiles) is shown in Figure 1. Some patterns are notable. Counties with high OUD rates are mainly concentrated in the Pacific Coast and the Four Corners region. Counties in Oklahoma, Michigan, Mid-Appalachian Region, and along the Gulf of Mexico Coastal Region also report high OUD rates. By contrast, counties in Mid-West, Great Plains, and Northeastern states have low OUD rates. These patterns suggest that OUD rates are not evenly distributed across space. The spatial process that generates these patterns is likely to be place-dependent and spatial heterogeneity seems to exist in the data, which will be formally examined with the MGWR analysis.

OLS and MGWR results

The OLS and MGWR results are summarized in Table 2. Specifically, columns (a) and (b) are drawn from the OLS analysis and columns (c) to (i) are based on the MGWR modeling. We discuss the main findings below. First, the global (OLS) estimates [i.e., column (a)] suggest that racial/ethnic composition, health conditions, and HCC score are associated with the OUD rate. For example, higher percentages of non-Hispanic black and Hispanic beneficiaries are associated

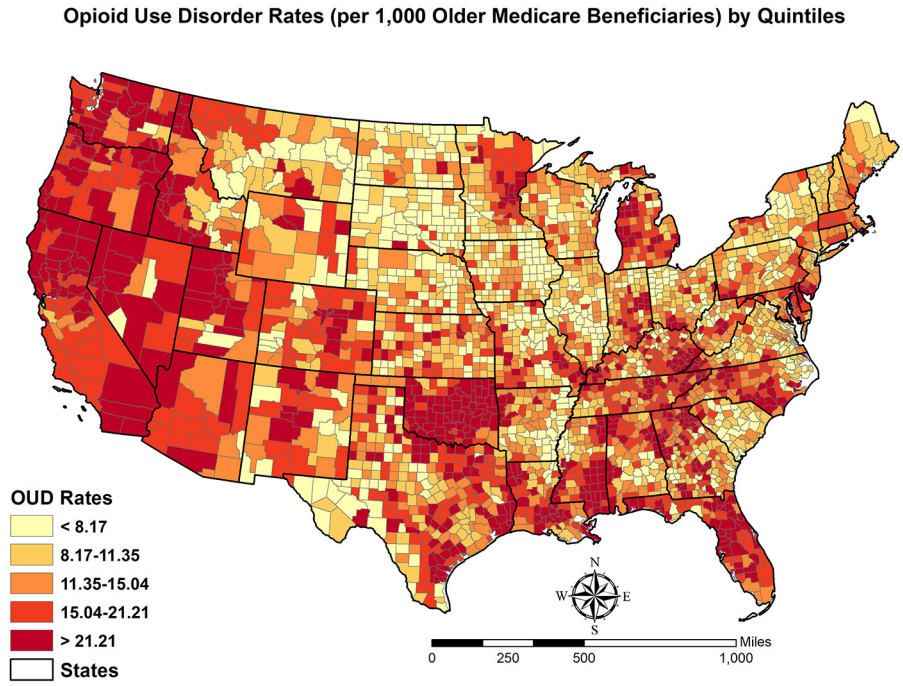


FIGURE 1
Spatial distribution of opioid use disorder rates (per 1,000 older Medicare beneficiaries) by quintiles in contiguous US.

with lower OUD rates at the county-level. Mental health conditions and HCC scores are positively related to OUD rates. It should be noted that the association between mental health conditions and OUD is opposite to that between physical health conditions and OUD. One plausible explanation is that the physical health conditions included in our measures are not strongly associated with pain but they require regular doctor visits, which may increase the awareness of opioid misuse or abuse. Beyond beneficiary characteristics, the OUD rate of a county increases with social isolation and decreases with residential stability. There is no significant relationship between concentrated disadvantages and the rate of OUD. Column (b) includes the variance inflation factors (VIF) among the independent variables. As all VIFs are smaller than 10 (the commonly used criterion), multicollinearity is unlikely to bias the estimates of standard errors of coefficients.

Second, columns (c) to (g) are the summary statistics of the MGWR local estimates. Some variables are estimated to have divergent associations with OUD rates, while others show homogeneous relationships across space. Take the percentage of non-Hispanic white beneficiaries for example, its minimal local estimate [column (e)] is−0.56 but the maximal local estimate [column (g)] is 0.07. By contrast, the local estimates of the percentage of non-Hispanic black beneficiaries range between −0.34 and −0.33, suggesting a highly homogeneous relationship in US counties. The Monte Carlo test for spatial non-stationarity [column (h)] largely echoes the distribution of the local estimates for each variable. Three variables are found to have spatially varying associations with OUD rates, namely the percentage of non-Hispanic white beneficiaries, average number of mental health conditions, and average HCC scores.[2] The three variables also have relatively small bandwidths compared with other covariates. The bandwidth of the percentage of non-Hispanic white beneficiaries is 358, which is comparable with that of the average number of mental health conditions (bandwidth = 384). The average HCC score has the smallest bandwidth of 44.

Third, in terms of model diagnosis, the corrected Akaike Information Criterion (AICc) is much smaller in the MGWR model (6,352.86) than the OLS model (8,262.74), indicating that the MGWR model is preferred and fits the data better.

Spatial non-stationarity in OUD patterns

To further demonstrate spatial non-stationarity, we visualize the MGWR results for the percentage of non-Hispanic white

2 The percentage of Hispanic beneficiaries is marginally significant (p-value = 0.052) so we exclude it from the discussion.

TABLE 2 OLS and MGWR results of opioid use disorder (OUD) rate (per 1,000 older medicare beneficiaries).

	Global estimates (a)	VIF† (b)	Mean (c)	SD (d)	Min (e)	Median (f)	Max (g)	Monte Carlo *p*-value (h)	MGWR Bandwidth (i)‡
Percentage of female	−0.04	1.74	0.03	0.00	0.03	0.03	0.04	0.87	3,106
Average age of beneficiaries	−0.21***	1.50	−0.12	0.02	−0.16	−0.13	−0.09	0.12	2,359
Percentage of NH white	−0.07	9.95	−0.30	0.16	−0.56	−0.34	0.07	0.01	358
Percentage of NH black	−0.17***	5.24	−0.34	0.00	−0.34	−0.34	−0.33	0.94	3,106
Percentage of Hispanic	−0.09*	5.09	−0.26	0.10	−0.53	−0.24	−0.13	0.05	1,474
Percentage of dual eligibility	0.00	2.95	−0.05	0.00	−0.06	−0.05	−0.05	0.98	3,106
Average number of mental health conditions	0.11***	2.79	0.07	0.09	−0.09	0.07	0.31	0.02	384
Average number of physical conditions	−0.11**	4.78	−0.05	0.00	−0.06	−0.05	−0.05	0.57	3,106
Average HCC score	0.34***	4.95	0.41	0.36	−0.27	0.34	2.86	<0.001	44
Social isolation index	0.12***	2.49	−0.01	0.00	−0.01	−0.01	0.00	0.79	3,106
Concentrated disadvantage index	0.01	3.35	0.00	0.00	−0.01	−0.01	0.01	0.66	3,106
Residential stability	−0.05**	1.32	0.01	0.00	0.01	0.01	0.02	0.95	3,106
Intercept	0.00	–	−0.04	0.50	−0.98	−0.13	2.27	<0.001	44
AIC_C	8,262.74		6,352.86						
Adjusted R^2	0.17		0.61						

Significance: *p < 0.05, **p < 0.01, ***p < 0.001.

†The variance inflation factors (VIF) among the independent variables are all smaller than 10, indicating that multicollinearity is not a concern.

‡The bandwidth is determined with the number of nearest neighbors for each location. This is a conventional approach in MGWR.

OLS, ordinary least squares; MGWR, multiscale geographically weighted regression; AICc, corrected Akaike Information Criterion.

beneficiaries, average of mental health conditions, and average HCC scores in Figures 2–4. Before discussing the spatially varying associations with OUD, it should be noted that MGWR estimates are not statistically significant (*p*-value > 0.05) in white areas. Only colored areas refer to statistically significant associations. This visualization method has been commonly used in the geographically weighted regression literature (33).

Map A in Figure 2 shows the spatially varying relationship between the percentage of non-Hispanic white beneficiaries and OUD rates. The significant associations are found mainly to the east of Mississippi River with some exceptions including counties in South Carolina, northern Georgia, and southern Florida. The local associations suggest that higher percentages of non-Hispanic white beneficiaries are associated with lower OUD rates in these areas. We note that in the OLS estimates [Table 2, column (a)], the percentage of non-Hispanic white beneficiaries is not statistically significant. A plausible explanation for this discrepancy is that the positive local estimates offset the negative estimates, which leads to a null global relationship.

Map B in Figure 2 demonstrates how the average number of mental health conditions is related to OUD rates across space. The positive associations between these two variables are clustered in the West of US, particularly in the Pacific Coastal Region and Mountain States. The other two significant clusters are found in Louisiana, Alabama, Georgia, and New Jersey.

Map C in Figure 2 shows the spatial non-stationarity in the relationship between average HCC score and the OUD rate. There are two pockets where the local relationship is positively and strongly related to OUD rate (i.e., red areas). One is around Oklahoma, Eastern Texas and Western Louisiana. And the other is in Mid-Appalachian Regions, especially at the intersection between Virginia and Kentucky. Moreover, most counties in the Mountain States and Pacific Coastal Region are estimated to have a significant relationship.

Multiscale dimensions of geographic disparities in OUD

The three dimensions of spatial process for each independent variable are presented in Table 3. Regarding the first dimension, *level of influence*, 5 variables are identified as primary influencers and 7 variables are secondary influencers. For example, the percentage of non-Hispanic black beneficiaries is a significant factor for OUD rate in every county so that all beneficiaries are affected by this variable. As such, this covariate is identified as a primary influencer. In contrast, approximately

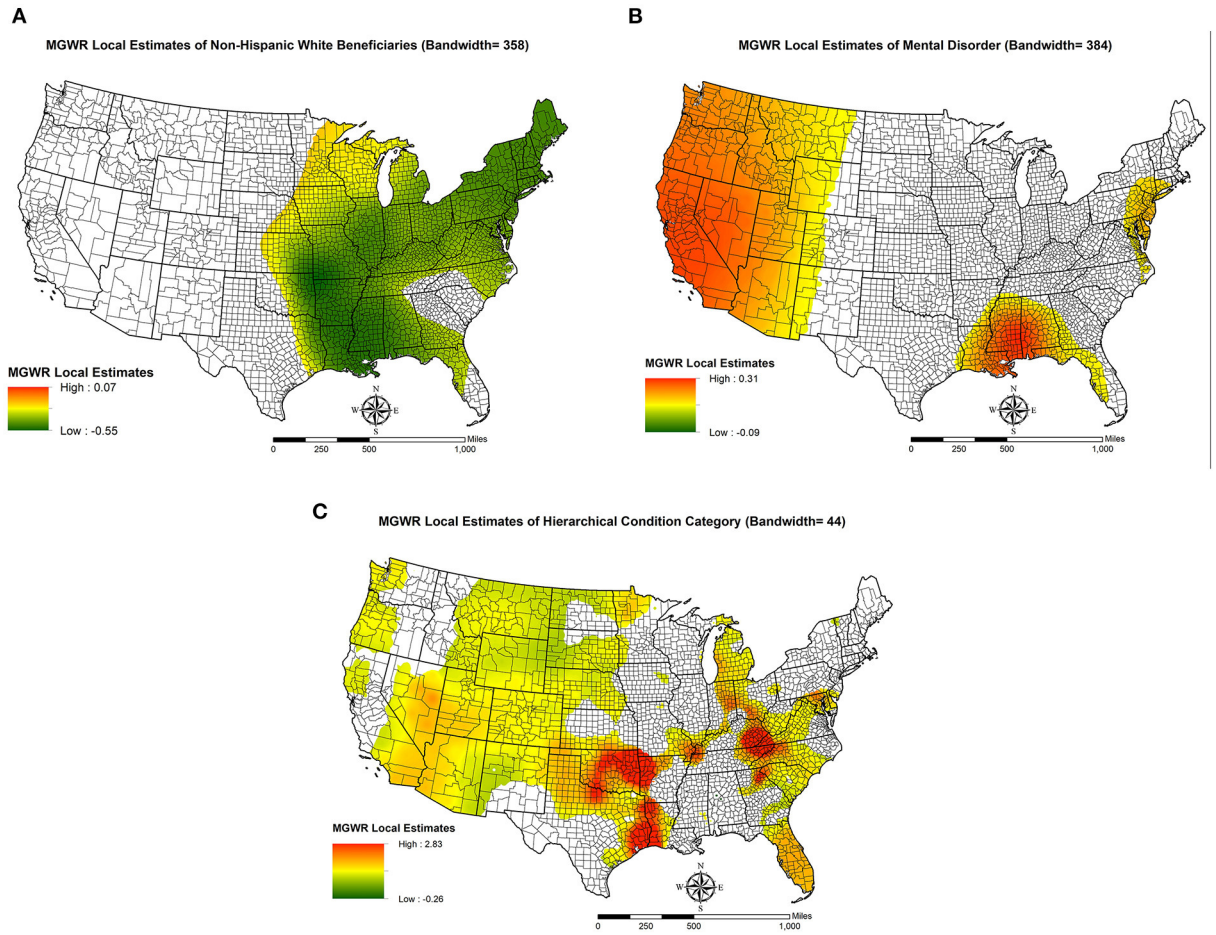


FIGURE 2
Spatial non-stationarity in the relationships between key independent variables and opioid use disorder rates (per 1,000 older Medicare beneficiaries) in US counties. **(A)** MGWR Local Estimates of Non-Hispanic White Beneficiaries (Bandwidth = 358); **(B)** MGWR Local Estimates of Mental Disorder (Bandwidth = 384); **(C)** MGWR Local Estimates of Hierarchical Condition Category (Bandwidth = 44).

41 percent of beneficiaries live in counties where the average number of mental health conditions is a significant factor, which is <50 percent and makes this covariate a secondary influencer.

In terms of the second dimension, *scalability*, 7 independent variables (e.g., percentage of female beneficiaries) have a bandwidth >75 percent of the global bandwidth (3,108*0.75 = 2,331) and they are categorized into the "global scale" group. Three variables have a bandwidth <25 percent of the global bandwidth (3,108*0.25 = 777). For example, the estimated bandwidth for the average HCC score is 44, which indicates that the OUD rate of a focal county is shaped by the nearest 44 counties. As such, this variable is associated with OUD rate at the "local scale". Two independent variables, namely percentage of Hispanic beneficiaries and average age, have a bandwidth between 2,331 and 777 and they are defined as variables that operate to affect OUD rates at the "regional scale".

With respect to *specificity*, 4 variables are found to have the strongest associations with OUD rates in US counties. Among them, average HCC score is estimated to be the most dominant variable in 1,294 of the total 3,108 counties (i.e., 41.6 percent). The percentage of non-Hispanic white beneficiaries is the most dominant factor in 908 counties, which is higher than the percentage of non-Hispanic black beneficiaries (662 counties) and the percentage of Hispanic beneficiaries (244 counties). We visualize the specificity dimension in Figure 3 and observe the following patterns. The average HCC score is the most dominant variable in most counties of Mountain States, such as Utah and Colorado, as well as Oklahoma and Northern Texas. Regarding the percentage of non-Hispanic white beneficiaries, it is mainly clustered in the Northeastern Region, Alabama, Illinois, and Missouri. The percentage of non-Hispanic black beneficiaries is found to have the strongest association with the OUD rate

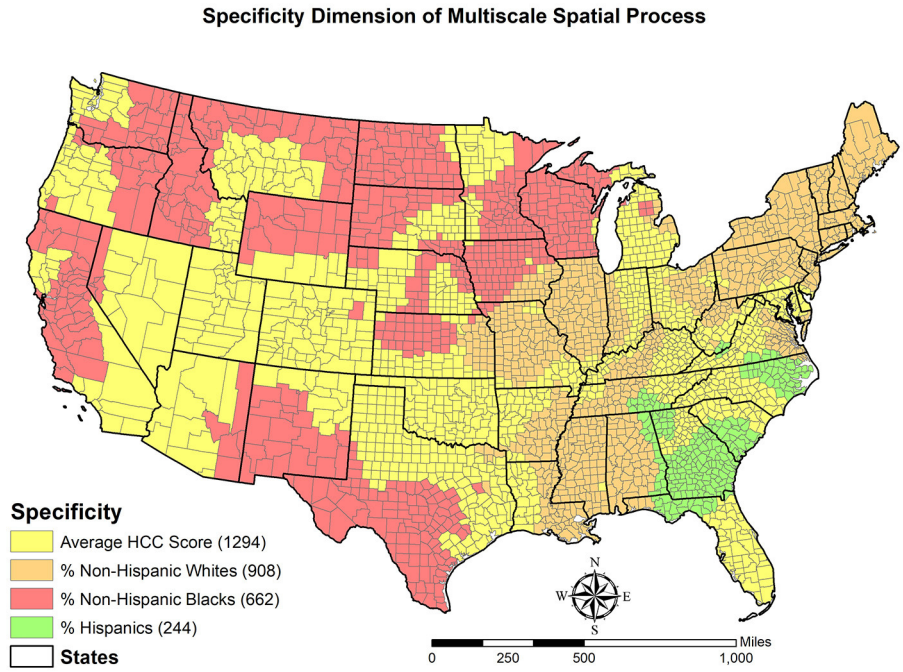


FIGURE 3
Specificity dimension of multiscale spatial process.

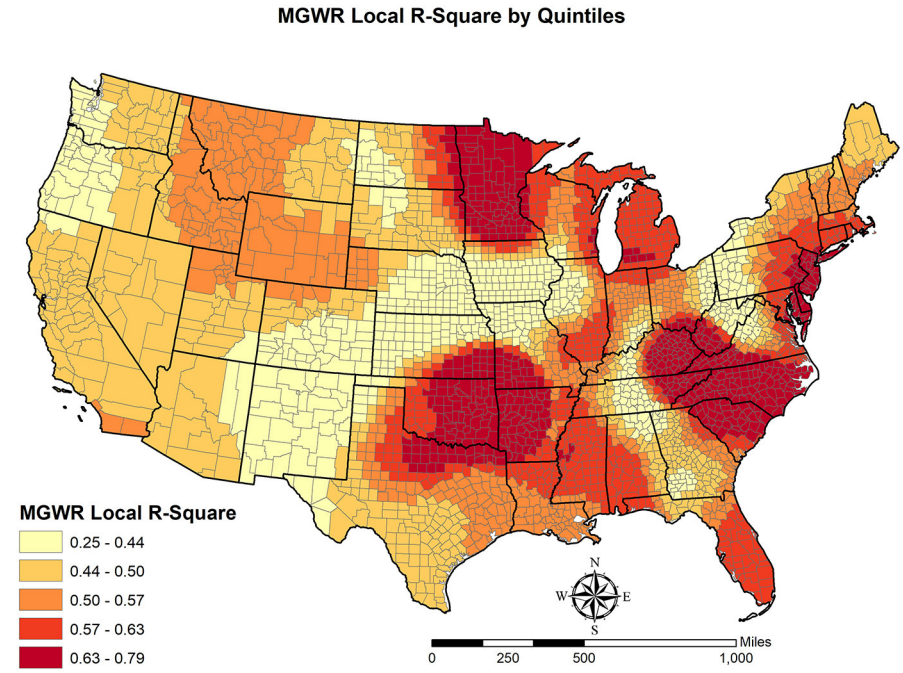


FIGURE 4
Local R-squares based on the multiscale geographically weighted regression.

TABLE 3 Three Dimensions of Multiscale Spatial Process for Each Independent Variable Based on the MGWR Models.

Variable (bandwidth)	Level of influence[a]	Scalability[b]	Specificity[c]
Percentage of female (3,106)	Secondary (0.8%)	Global	0
Average age of beneficiaries (2,359)	Primary (100.0%)	Regional	0
Percentage of NH white (358)	Primary (65.5%)	Local	908 (29.2%)
Percentage of NH black (3,106)	Primary (100.0%)	Global	662 (21.3%)
Percentage of Hispanic (1,474)	Primary (100.0%)	Regional	244 (7.9%)
Percentage of dual eligibility (3,106)	Secondary (0.0%)	Global	0
Average number of mental health conditions (384)	Secondary (40.9%)	Local	0
Average number of physical conditions (3,106)	Secondary (0.0%)	Global	0
Average HCC score (44)	Primary (52.1%)	Local	1,294 (41.6%)
Social isolation index (3,106)	Secondary (0.0%)	Global	0
Concentrated disadvantage index (3,106)	Secondary (0.0%)	Global	0
Residential stability (3,106)	Secondary (0.0%)	Global	0

[a]If the variable affects more than 50% of the total population, it is a primary influencer; otherwise (i.e., $\leq$ 50%), it is a secondary influencer. The percentage of population affected by a factor is included in the parentheses.
[b]If the bandwidth of a variable is larger than 75% of the global bandwidth (i.e., 2,331), it is a global determinant; if the bandwidth is smaller than 25% of the global bandwidth (i.e., 777), it is a local determinant; if the bandwidth is between 75% and 25% of the global bandwidth, it is a regional determinant.
[c]The number and percentage of counties that the focal variable has the strongest significant impact on the dependent variable (i.e., the largest absolute value of the coefficients that are statistically significant).
MGWR, multiscale geographically weighted regression.

in Northern Great Plains, Idaho, Wisconsin, and along the US-Mexico border. Finally, the percentage of Hispanic beneficiaries is concentrated on Southern Georgia and part of the Carolinas.

The specificity dimension further illustrates spatial non-stationarity embedded in the geographic disparities in OUD rates among older adults. Explicitly, the relationship between a certain factor and the OUD rate within a county is not homogeneous across space in that either the direction or the magnitude of this relationship varies by location. In other words, the same change in a covariate may invoke different changes in the OUD rate and the differential responses depends on where a county is located and its surrounding counties.

MGWR also offers the local R-square for each county, which is visualized in Figure 4. The spatial pattern of local R-squares suggest that the model specification of this study fits the observed OUD rates best in the Great Lakes Region, Atlantic Coastal Region, most of the Black Belt region, and Oklahoma and its surrounding states.

Discussion

With the results above, we revisited the three gaps in the extant literature. First, little is known about how OUD rates among older adults are distributed in US counties and the potential geographic disparity in OUD rates has not been explored. This gap can be filled with the exploratory spatial data analysis results of this study. Specifically, based on the 2020 Medicare data, counties with high OUD rates are concentrated in the West of the US with some pockets scattered in Oklahoma, Mid-Appalachian Region, and Florida. By contrast, counties with low OUD rates are in the Mid-West. Such a pattern suggests that the distribution of OUD rates among older adults is uneven and it is likely that spatial heterogeneity exists in US counties. More specifically, OUD rates tend to be place-dependent and different spatial scales reflect different spatial associations. For example, counties in the state of Oklahoma demonstrate a strong spatial dependence (i.e., counties with high OUD rates are nearby) whereas counties in the Northeast Region (e.g., Pennsylvania, New Jersey, and New York) may reflect spatial heterogeneity (i.e., OUD rates vary within thin this region).

Situating this finding into the literature, prior ecological studies either investigate the patterns within a single state or region (34, 35) or explore substance or opioid abuse among the general population (36), rather than older adults. After reviewing 46 published articles, Marks and colleagues (37) conclude that geospatial analysis techniques are commonly used in research of opioid-related outcomes. Nonetheless, no study has adopted a local spatial perspective to investigate the existence of spatial heterogeneity in ecological data. To our knowledge, this study is among the first to present such a spatial pattern of and geographic disparity in OUD rates among older adults in the contiguous US.

Second, several county-level studies have suggested that drug-overdose mortality is spatially heterogenous in that some factors are more important in certain counties than others (7, 8), but whether this argument can be applied to OUD rates among older adults is unknown. Drawing from the MGWR results, we found that at the county-level, only three variables have spatially varying associations with OUD rates and others operate

at the global or regional level. That is, we obtain evidence for spatial heterogeneity underlying the pattern of OUD rates, but such evidence only comes from beneficiaries' characteristics, i.e., percentage of non-Hispanic white beneficiaries, average number of mental health conditions, and average HCC score. These variables demonstrate unique patterns, which are visualized in Figure 2.

How do our findings related to spatial heterogeneity contribute to the literature? For one, without a local modeling perspective, our OLS findings largely echo a recent ecological study (16). For example, social isolation is negatively associated with OUD rates and residential stability decreases OUD rates. With the MGWR results, we can confirm that these two variables (i.e., social isolation and residential stability) have a universal relationship with OUD across space. Furthermore, the relationship between average number of mental health conditions and OUD is only significant in the West of US and part of the Black Belt. The average HCC score also demonstrates a strong spatial heterogeneity pattern. While some scholars have used typology analysis to investigate spatial heterogeneity in opioid-related health outcomes (38), this approach does not explore spatial heterogeneity for each independent variable.

Finally, this study challenges the commonly used global modeling perspective in the literature and identified three dimensions of the spatial process that generates the observed OUD rates. In terms of the level of influence, this study concludes that beneficiaries' characteristics play a larger role in shaping OUD rates than the socioeconomic conditions of a county because racial/ethnic composition of beneficiaries and the average of HCC score are categorized as primary influencers. Regarding scalability, the MGWR results support the argument that different independent variables may operate at different spatial scales to affect OUD rates. All three types of scalability, namely global, regional, and local, are found in this study. This finding is similar to a recent study (21) and indicates that it may not be appropriate to adopt the constant spatial scale assumption. The third dimension, specificity, shows that four factors are estimated to have the strongest association with OUD rates. Among them, the average HCC score is the most dominant in more than 40 percent of the total 3,108 counties, followed by the percentage of non-Hispanic white beneficiaries.

The three dimensions of spatial process take advantage of the strengths of MGWR and serve as an alternative to illustrate and visualize spatial heterogeneity. To our knowledge, the three dimensions have not been applied to ecological OUD studies and this study is the first to describe these dimensions specific to the 2020 OUD rates among older adults. Research on the opioid epidemic has paid attention to middle-aged populations and the opioid use behavior among older adults is often overlooked. As opioid prescription and regulation is less restricted in the US healthcare systems than in other developed countries (39), the spatial heterogeneity and spatial process found in this study may be unique to the study population.

While MGWR has overcome several limitations of GWR, such as multiple testing and developing a local inferential statistics framework (19, 30), it still has some shortcomings and our results should be interpreted with these caveats in mind. First, the multicollinearity among local estimates remains likely to be a concern, even though using different weight matrices in the back-fitting algorithm may minimize multicollinearity. Second, the current MGWR is developed for continuous outcomes that largely follow a normal distribution. When the dependent variable is highly skewed or sparse, the MGWR estimates may not be reliable. Finally, the global estimates cannot be decomposed into MGWR local estimates. As such, the global and local parameters do not have a clear relationship.

This study is subject to several limitations. First, using a different geographic unit (e.g., states) may lead to different findings and conclusions, which is known as the modifiable areal unit problem (40, 41). While the Medicare data can be aggregated to ZIP codes, which is the most granular unit available at CMS, we opted not to use this unit to avoid the small area estimation problem (i.e., few beneficiaries in a ZIP code). Second, starting January 1, 2020, the Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment (SUPPORT) for Patients and Communities Act was enacted. Under the SUPPORT Act, CMS is allowed to pay Opioid Treatment Programs through bundled payments for OUD treatment services including FDA approved medications for OUD and related services (e.g., substance use counseling and periodic assessments). As such, we may observe more beneficiaries with OUD than previous years and analyzing data before 2020 may yield different results. Third, the cross-sectional research design does not allow us to make any causal inference and the findings cannot be generalized to other age populations.

Several policy implications can be drawn from this study. One is that the one-model-fits-all or global approach may not effectively address the increasing OUD rates among older adults. The MGWR findings suggest that policies aiming to lower OUD rates should focus on counties with high average HCC scores and high percentages of non-Hispanic white beneficiaries (the top two variables in the specificity dimension). In addition, higher average age of beneficiaries is also an important factor as this variable is significant in all counties. It may be necessary to prioritize resources to counties with higher concentrations of beneficiaries in the middle-old (ages 75–84) and the old-old (85+) age ranges. Finally, more attention should be paid to the place-based policies so that the differences in culture, values, attitudes, norms, and socioeconomic conditions across space can be explicitly considered in possible interventions.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The beneficiary level data can

be accessed through the Research Data Center. Requests to access these datasets should be directed to Research Data Center, resdac@umn.edu.

Author contributions

T-CY conceptualized this study, led the writing of the manuscript, and conducted the major analysis. CS managed the county-level data and contributed to the writing of the manuscript. S-wC interpreted findings, participated in writing, and critically commented on the manuscript. FS visualized, interpreted findings, and participated in writing. All authors read and approved the final manuscript.

Funding

T-CY received the Interdisciplinary Research Network on Rural Population Health and Aging, which is funded by the National Institute on Aging (R24-AG065159).

Acknowledgments

We acknowledge Stephen A. Matthews for his feedback on the proposed methods and the constructive suggestions from the reviewers.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Arias E, Tejada-Vera B, Ahmad F. *Provisional Life Expectancy Estimates for January through June, 2020 NVSS vital statistics rapid release; report no. 010.* National Center for Health Statistics, U. S. Division of Vital Statistics (2021). Available online at: https://stacks.cdc.gov/view/cdc/100392

2. Ho JY, Hendi AS. Recent trends in life expectancy across high income countries: retrospective observational study. *BMJ.* (2018) 362:k2562. doi: 10.1136/bmj.k2562

3. Woolf SH, Masters RK, Aron LY. Effect of the Covid-19 pandemic in 2020 on life expectancy across populations in the USA and other high income countries: simulations of provisional mortality data. *BMJ.* (2021) 373:n1343. doi: 10.1136/bmj.n1343

4. Case A, Deaton A. *Deaths of Despair and the Future of Capitalism*: Princeton, NJ: Princeton University Press (2020).

5. Stein EM, Gennuso KP, Ugboaja DC, Remington PL. The epidemic of despair among White Americans: trends in the leading causes of premature death, 1999–2015. *Am J Public Health.* (2017) 107:1541–7. doi: 10.2105/AJPH.2017.303941

6. Ruhm CJ. Drivers of the fatal drug epidemic. *J Health Econ.* (2019) 64:25–42. doi: 10.1016/j.jhealeco.2019.01.001

7. Monnat SM, Peters DJ, Berg MT, Hochstetler A. Using census data to understand county-level differences in overall drug mortality and opioid-related mortality by opioid type. *Am J Public Health.* (2019) 109:1084–91. doi: 10.2105/AJPH.2019.305136

8. Rigg KK, Monnat SM, Chavez MN. Opioid-related mortality in rural america: geographic heterogeneity and intervention strategies. *Int J Drug Policy.* (2018) 57:119–29. doi: 10.1016/j.drugpo.2018.04.011

9. Huhn AS, Strain EC, Tompkins DA, Dunn KE. A hidden aspect of the us opioid crisis: rise in first-time treatment admissions for older adults with opioid use disorder. *Drug Alcohol Depend.* (2018) 193:142–7. doi: 10.1016/j.drugalcdep.2018.10.002

10. Shoff C, Yang T-C, Shaw BA. Trends in opioid use disorder among older adults: analyzing medicare data, 2013–2018. *Am J Prev Med.* (2021) 60:850–5. doi: 10.1016/j.amepre.2021.01.010

11. Le Roux C, Tang Y, Drexler K. Alcohol and opioid use disorder in older adults: neglected and treatable illnesses. *Curr Psychiatry Rep.* (2016) 18:87. doi: 10.1007/s11920-016-0718-x

12. Maree RD, Marcum ZA, Saghafi E, Weiner DK, Karp JF. A systematic review of opioid and benzodiazepine misuse in older adults. *Am J Geriatr Psychiatry.* (2016) 24:949–63. doi: 10.1016/j.jagp.2016.06.003

13. Butler MM, Ancona RM, Beauchamp GA, Yamin CK, Winstanley EL, Hart KW, et al. Emergency department prescription opioids as an initial exposure preceding addiction. *Ann Emerg Med.* (2016) 68:202–8. doi: 10.1016/j.annemergmed.2015.11.033

14. Cochran G, Rosen D, McCarthy RM, Engel RJ. Risk factors for symptoms of prescription opioid misuse: do older adults differ from younger adult patients? *J Gerontol Soc Work.* (2017) 60:443–57. doi: 10.1080/01634372.2017.1327469

15. Wang Y-P, Andrade LH. Epidemiology of alcohol and drug use in the elderly. *Curr Opin Psychiatry.* (2013) 26:343–8. doi: 10.1097/YCO.0b013e328360eafd

16. Yang T-C, Shoff C, Kim S. Social isolation, residential stability, and opioid use disorder among older medicare beneficiaries: metropolitan and non-metropolitan county comparison. *Soc Sci Med.* (2022) 292:114605. doi: 10.1016/j.socscimed.2021.114605

17. Yang T-C, Shoff C, Kim S, Shaw BA. County social isolation and opioid use disorder among older adults: a longitudinal analysis of medicare data, 2013–2018. *Soc Sci Med.* (2022) 301:114971. doi: 10.1016/j.socscimed.2022.114971

18. Fotheringham AS. Trends in quantitative methods I: stressing the local. *Prog Hum Geogr.* (1997) 21:88–96. doi: 10.1191/030913297676693207

19. Fotheringham AS, Yang W, Kang W. Multiscale geographically weighted regression (Mgwr). *Ann Am Assoc Geogr.* (2017) 107:1247–65. doi: 10.1080/24694452.2017.1352480

20. Oshan TM Li Z, Kang W, Wolf LJ, Fotheringham AS. Mgwr: a python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS Int J Geo-Inform.* (2019) 8:269. doi: 10.3390/ijgi8060269

21. Yang TC, Matthews SA, Sun F. Multiscale dimensions of spatial process: COVID-19 fully vaccinated rates in U.S. counties. *Am J Prev Med.* (2022) S0749-3797(22)00331-2. doi: 10.1016/j.amepre.2022.06.006

22. U.S. Census Bureau. *2016-2020 American Community Survey 5-Year Public Use Microdata Samples.* (2022). Available online at: https://data.census.gov/cedsci/ (accessed May 1, 2022).

23. Research Data Assistance Center. *Overarching Oud Disorder (Any of the Three Sub-Indicators) - Medicare Only Claims*. (2022). Available online at: https://resdac.org/cms-data/variables/overarching-oud-disorder-any-three-sub-indicators-medicare-only-claims (accessed November 7, 2022).

24. Hoffman AF, Reiter KL, Randolph RK. *Average Beneficiary Cms Hierarchical Condition Category (Hcc) Risk Scores for Rural and Urban Providers.* Chapel Hill, NC: Sheps Center for Health Services Research (2018).

25. United Health Foundation. *America's Health Rankings Senior Report 2018 Saint Paul, Minnesota*. (2018). Available online at: https://assets.americashealthrankings.org/app/uploads/ahrsenior18-finalv1.pdf (accessed August 5, 2019).

26. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science.* (1997) 277:918–24. doi: 10.1126/science.277.5328.918

27. Brunsdon C, Fotheringham S, Charlton M. Geographically weighted regression. *J R Stat Soc Ser D.* (1998) 47:431–43. doi: 10.1111/1467-9884.00145

28. Fotheringham AS, Brunsdon C, Charlton M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships.* New York, NY: John Wiley & Sons (2003).

29. Nakaya T, Fotheringham S, Charlton M, Brunsdon C. Semiparametric geographicallyweighted generalised linear modelling in Gwr 4.0. *Proc. Geocomputation*. (2009). p. 1–5. Available online at: https://mural.maynoothuniversity.ie/4846/1/MC_Semiparametric.pdf

30. Yu H, Fotheringham AS Li Z, Oshan T, Kang W, Wolf LJ. Inference in multiscale geographically weighted regression. *Geogr Anal.* (2020) 52:87–106. doi: 10.1111/gean.12189

31. Wood SN. *Generalized Additive Models: An Introduction with R.* Boca Raton, FL: CRC press (2017).

32. Greenland S, Maclure M, Schlesselman JJ, Poole C, Morgenstern H. Standardized regression coefficients: a further critique and review of some alternatives. *Epidemiology.* (1991) 2:387–92. doi: 10.1097/00001648-199109000-00015

33. Matthews SA, Yang T-C. Mapping the results of local statistics: using geographically weighted regression. *Demogr Res.* (2012) 26:151. doi: 10.4054/DemRes.2012.26.6

34. Albright DL, McDaniel J, Kertesz S, Seal D, Prather K, English T, et al. Small area estimation and hotspot identification of opioid use disorder among military veterans living in the Southern United States. *Substance Abuse.* (2021) 42:116–22. doi: 10.1080/08897077.2019.1703066

35. Brownstein JS, Green TC, Cassidy TA, Butler SF. Geographic information systems and pharmacoepidemiology: using spatial cluster detection to monitor local patterns of prescription opioid abuse. *Pharmacoepidemiol Drug Saf.* (2010) 19:627–37. doi: 10.1002/pds.1939

36. Cerdá M, Gaidus A, Keyes KM, Ponicki W, Martins S, Galea S, et al. Prescription opioid poisoning across urban and rural areas: identifying vulnerable groups and geographic areas. *Addiction.* (2017) 112:103–12. doi: 10.1111/add.13543

37. Marks C, Carrasco-Escobar G, Carrasco-Hernández R, Johnson D, Ciccarone D, Strathdee SA, et al. Methodological approaches for the prediction of opioid use-related epidemics in the United States: a narrative review and cross-disciplinary call to action. *Translational Research.* (2021) 234:88–113. doi: 10.1016/j.trsl.2021.03.018

38. Kolak MA, Chen Y-T, Joyce S, Ellis K, Defever K, McLuckie C, et al. Rural risk environments, opioid-related overdose, and infectious diseases: a multidimensional, spatial perspective. *Int J Drug Policy.* (2020) 85:102727. doi: 10.1016/j.drugpo.2020.102727

39. Fischer B, Keates A, Bühringer G, Reimer J, Rehm J. Non-medical use of prescription opioids and prescription opioid-related harms: why so markedly higher in N Orth a Merica compared to the rest of the world? *Addiction.* (2014) 109:177–81. doi: 10.1111/add.12224

40. Fotheringham AS, Wong DW. The modifiable areal unit problem in multivariate statistical analysis. *Environ Plann A.* (1991) 23:1025–44. doi: 10.1068/a231025

41. Openshaw S. *The Modifiable Areal Unit Problem*. Norwick: Geo Books (1983).

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 18 October 2022
DOI 10.3389/fpubh.2022.993585

OPEN ACCESS

EDITED BY
Shannon Monnat,
Syracuse University, United States

REVIEWED BY
Jason Houle,
Dartmouth College, United States
Adam Lippert,
University of Colorado, United States

*CORRESPONDENCE
Terri Friedline
tfriedli@umich.edu

†These authors have contributed equally to this work and their names are listed alphabetically

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 13 July 2022
ACCEPTED 30 September 2022
PUBLISHED 18 October 2022

CITATION
Agnew M, Bea MD and Friedline T (2022) Payday lenders and premature mortality.
Front. Public Health 10:993585.
doi: 10.3389/fpubh.2022.993585

Payday lenders and premature mortality

Megan Agnew[1†], Megan Doherty Bea[2†] and Terri Friedline[3*†]

[1]Department of Population Health Sciences, University of Wisconsin-Madison, Madison, WI, United States, [2]Department of Consumer Science, University of Wisconsin-Madison, Madison, WI, United States, [3]University of Michigan, Ann Arbor, MI, United States

Relationships between debt and poor health are worrisome as access to expensive credit expands and population health worsens along certain metrics. We focus on payday lenders as one type of expensive credit and investigate the spatial relationships between lender storefronts and premature mortality rates. We combine causes of death data from the Centers for Disease Control and Prevention (CDC) and payday lender locations at the county-level in the United States between 2000 and 2017. After accounting for county socioeconomic and demographic characteristics, the local presence of payday lenders is associated with an increased incidence risk of all-cause and specific-cause premature mortality. State regulations may attenuate these relationships, which provides insights on policy strategies to mitigate health impacts.

KEYWORDS

payday lending, debt, health, premature death, regulation

Introduction

Adverse health conditions such as hypertension, cardiovascular disease, and declines in mental health are associated with debt burdens from ubiquitous access to expensive credit (1–5). Such health concerns are worrisome amidst rising debt (6) and recent deterioration in U.S. life expectancies. While overall premature mortality rates have been declining over the past two decades, since 2014, premature deaths have increased for some groups primarily due to drug overdoses and suicides (7–9). Although debt burdens have been shown to contribute to poor health (2, 5) and even premature death (1, 10, 11), the mechanisms by which communities' increasing access to expensive credit impact mortality are not well understood.

One trend in the availability of expensive credit is the expansion of higher-cost financial services like payday lenders in communities across the United States. The number of these storefront locations has increased nationwide since the mid-1990s (12–16), and the debt that borrowers accumulate from these higher-cost lenders contribute to their financial difficulties such as struggling to pay bills and delaying routine medical care (12, 13). This debt may also have effects in the aggregate, such as by contributing to communities' economic distress and worsened health outcomes regardless of whether any particular resident has borrowed expensive debt. One obstacle to identifying and testing these mechanisms is limited data on the extent of communities' financial services, making it difficult to associate communities' access to expensive credit with residents' health and premature mortality. A broad literature explores the spatial nature of business locational decisions such as fast food restaurants (17, 18), blood and plasma

donation clinics (19, 20), and dollar stores (21, 22) and their associations to community economic distress with implications for public health. However, similar investigations of higher-cost lenders are limited.

In the current study, we investigate whether the presence of payday lenders is associated with premature mortality and hypothesize two mechanisms for explaining these relationships: residents living in areas with a higher number of payday lenders accumulate more higher-cost debt, and a higher density of payday lenders indicates areas' economic distress. We combine novel data including causes of death from the Centers for Disease Control and Prevention (CDC) WONDER database and the locations of payday lenders from InfoGroup USA at the county level between 2000 and 2017. Results indicate that the concentration of payday lenders may matter; though, the associations are conditional on the regulatory environment that informs payday lending practices. Importantly, we find evidence that state regulations can attenuate these relationships, especially for counties with high concentrations of payday lenders. Our findings offer new vantage points regarding the impacts of payday lending regulation. Beyond reducing financial difficulties related to paying bills, affording rent, and filing for bankruptcy that have been a focus of existing research (12–16), we suggest that regulating higher-cost financial services might advance community public health and protect against premature mortality for some groups.

The rise of consumer debt

The use and accumulation of consumer debt are often considered indicators of a person's access to financial services and their integration into the economy (23–26). Consumer debt is increasingly required to subsidize the costs of participating in today's economy (27), which is characterized by reduced collective bargaining power (28–31), low and stagnant wages (32, 33), and widening inequality (34, 35). People rely on debt to cope with these economic trends, as indicated by steadily rising debt burdens from medical expenses, student loans, credit cards, and payday and installment loans (6). Consumer debt increased in the years following the Great Recession and reached $14 trillion in 2019 (6).

Consumer debt is also an area of stratification, where heterogeneity in the quality and cost of certain types of debt may indicate exploitation or exclusion from the economy as opposed to integration (24). Debt from higher-cost, lower quality or "alternative" financial services—such as payday and installment lenders, auto title lenders, and tax refund and anticipation lenders—is expanding, growing by about 6% each year and reaching $141 billion in 2016 (36). The alternative financial services industry has expanded with the advent of online lending, and payday loans in particular comprised $14 billion of all online lending in 2016 (26–38). State regulations that restrict or prohibit certain usurious financial services appear to effectively constrain online lenders from crossing physical geographic boundaries (39–41). As such, increases in online lending may indicate a reinforcement of the industry's spatial ties to economically distressed communities where these lenders' storefronts are disproportionately located, and allude to concerning trends in the rise of consumer debt (16, 23).

The existing literature on debt typically focuses on individual borrowing behavior (see Borck et al. (42), O'Neill et al. (43), and Simone and Walks (44) for exceptions). This includes people's borrowing from the alternative financial services industry (45–48) and the potential consequences to their finances and health (1–5, 12, 13, 49–51). Yet the rising debt burdens of individuals (5) may also accumulate to produce effects that are observable at ecological or community levels, particularly given the extent to which lending and borrowing are spatially arranged (16, 52). For instance, people are more likely to borrow, and to borrow more often, when they live in areas with an increasing concentration of alternative financial services storefronts such as payday lenders (53).

The payday loan is a specific type of higher-cost credit product among the suite of alternative financial services. Payday loans have finance fees and an average annual interest rate of about 400%, which often prevent borrowers from repaying their original loans in full. Fifteen percent of borrowers renew their loans more than 10 times (47, 52). People who borrow payday loans are often younger, between the ages of 25 and 44, and have lower levels of education and income (46, 47). There is evidence of structural racism in borrowing (54), which contributes to Black Americans being more likely to borrow relative to White Americans, all else equal (46, 55). Borrowers of higher-cost debt report using payday loans to afford routine or recurring expenses (47), and experience financial difficulties related to paying bills, affording rent, filing for bankruptcy, and receiving routine medical care (12, 13, 49–51).

The locations of payday lender storefronts

A set of mutually reinforcing policies and practices have created spatial arrangements whereby higher-cost, lower-quality financial services are expanding and disproportionately locating in economically distressed and racially marginalized communities (15, 16, 56–64). Examining the locations of payday lenders in Colorado in 2007, a year when the state passed new legislation regulating payday loans, Gallmeyer and Roberts found that payday lender storefronts were disproportionately concentrated in census block groups with lower median incomes and higher poverty rates (62). Alternative financial services concentrate in White communities that are poor and economically distressed; though, unlike in predominantly White communities, these lenders' presence remains constant in Black

communities regardless of economic indicators like income and poverty (16). Black and Latino communities have nearly twice the number of alternative financial services than do White communities (58, 60)—disparities that are amplified by segregation (16). Notably, the places where payday lenders concentrate could be the same places abandoned by other resources such as grocery stores and hospitals, making residents more susceptible to health-related concerns. In other words, a higher concentration of payday lenders could dissuade the types of development activities that have the potential to improve public health outcomes and enable economic distress, although these potential connections have yet to be evaluated.

The alternative financial services industry's expansion, and growth in payday lender storefronts in particular, has happened more rapidly in some years and in some communities than in others. For example, the number of alternative financial services storefronts increased nearly five-fold nationally between the mid-1980s and -1990s (15), before continuing to grow at an annual rate of 15% (36, 59). The notable growth in storefronts experienced by some communities coincided with the Great Recession in the mid- to late-2000s and the continued rise in consumer debt (58, 63, 64). Check cashers in New York City capitalized on the foreclosure crisis by opening new storefronts in Black and Latino communities between 2006 and 2011 (58). In California, Michigan, Ohio, and Tennessee, the number of new payday lender storefronts peaked between approximately 2006 and 2008, before leveling off in some places (64–68). Michigan's payday lender storefronts initially concentrated their expansion within the state's most populous counties during the early 2000s. Lenders deepened their presence and broadened to other counties across the state after 2005, with notable increases in counties' storefront densities occurring in 2009 and 2013 (67).

State regulation plays a role in where payday lenders are located. Given concerns that payday loans trap borrowers in cycles of debt and worsen their financial difficulties (12, 13, 15), some states have moved to regulate the industry in order to protect their residents. Payday lending densities tend to be lower in states that have strong regulations, including interest rate caps, whereas densities are relatively higher in states with permissive regulations (69). Six states and the District of Columbia currently prohibit payday lending of any kind, while 21 states do not regulate payday lending at all. In the remainder of states, regulation varies between permissive and restrictive with more restrictive regulations capping annual interest rates, preventing rollover or repeat borrowing, and assessing borrowers' ability to repay loans (70–72).

Theoretical mechanisms linking payday lenders to mortality

There are several hypothesized mechanisms through which access to the higher-cost, lower-quality debt made available by the alternative financial services industry may influence premature mortality. One mechanism may operate at the individual level through accumulated debt. Individuals living in communities with higher concentrations of payday lenders tend to use these services at higher rates (12, 53, 73), contributing to their accumulated debt burdens and financial difficulties (13, 48, 50, 51, 74). For example, Friedline and Kepple find that individuals' increased use of alternative financial services is associated with more dense concentrations of higher-cost storefronts in their communities (53). In other words, residents who live in communities with higher concentrations of payday lenders may accumulate more debt. Among people who borrow, the financial burdens of their debt, which is an enduring source of stress that can compound over the life course (4), may place strains on their health and contribute to premature mortality. Higher-cost, lower-quality debt is associated with a range of health effects with implications for premature mortality including weight gain, depression, and suicide (2, 4, 5, 75, 76). Individuals who have accumulated debt such as from payday lenders are more likely to experience negative health consequences, including cardiovascular disease and premature mortality (1–5). In examining debt as a mediator of physical health disparities, Batomen and colleagues find that individuals with the highest amounts of unsecured debt, such as that from payday lenders, experienced an increased risk of premature death due to hypertension and cardiovascular disease, compared to their counterparts with the lowest amounts of unsecured debt (1). Taken together, these findings suggest that the presence of higher-cost, lower-quality alternative financial services like payday lenders in a person's community, as well as the debts that borrowers accumulate from these services, could contribute to rates of premature death.

Another explanatory mechanism may operate as an emergent effect (53, 77–81) and affect all residents in a community regardless of whether or not they borrow payday loans. From one perspective, the presence and or concentration of alternative financial services within a community may be a proxy for economic distress. Residents' longevity may be compromised by the extent to which the presence and or concentration of payday lenders indicate communities' economic marginalization and distress. In Toronto, Canada, a neighborhood's higher density of check cashing storefronts, which served as a proxy for economic distress, was associated with residents' increased risk of premature death (81). In a longitudinal study examining associations between county-level economic distress as indicated by unemployment rates and subprime credit ratings and mortality rates, counties that experienced the greatest distress in 2000 and 2010 had significantly higher baseline mortality rates and rates of increase (79).

Economic distress may also be causally linked to premature mortality. From this perspective, the presence and or concentration of alternative financial services is not simply a

proxy for economic distress. An increase in the concentration of payday lenders may subsequently increase a community's economic distress and therefore drive up premature mortality (82–84). Existing research implies a potential causal relationship between economic distress and premature mortality (79–81) and suggests that the concentration of alternative financial services influences community economic distress (62, 78), even if these relationships are not tested directly.

Prior ecological research finds supportive evidence for effects to emerge at the community level (79, 85–87). Higher concentrations of nuisance establishments like bars and alcohol outlets that often indicate economic marginalization and distress are associated with higher rates of child abuse and neglect, a relationship hypothesized to operate through community-level mechanisms (88–91). Relationships between the spatial arrangements of marijuana dispensaries and communities' crime rates are also hypothesized to operate through community-level mechanisms (92–94). Similar relationships exist between communities' payday lender storefronts and crime rates (80, 95). Along these lines, it is plausible that individuals' increased payday loan debts contribute to premature mortality, and that lenders' presence impacts premature mortality rates vis-à-vis economic marginalization and distress.

The current study

Using national county level data between 2000 and 2017, we examine how changes over time in the concentration of payday lender storefronts are associated with all-cause premature mortality. Among middle-aged Americans, ages 25–64, all-cause mortality rates were declining in 2000, plateaued by 2010, and began to increase after 2010 (8). These trends were especially pronounced from 2010 to 2017 when age-adjusted mortality rates increased by 6% primarily due to a substantial increase in drug overdoses, suicides, and alcoholic liver disease (8). Since most of these premature deaths are highly preventable, it is imperative to identify factors that exacerbate these deaths (81). Our analysis sheds light on two potential mechanisms that may lead to preventable premature deaths. We hypothesize that residents living in counties with higher concentrations of payday lenders have debt burdens that place strains on their health. We also hypothesize that payday lenders themselves may be a proxy of, and potential contributor to, community economic distress, which may worsen community public health outcomes. Our study cannot fully disentangle these mechanisms, but evidence of associations between payday lender presence and premature mortality will offer new pathways for scholarship on debt, access to financial services, and health. Further, a national perspective enables an evaluation of how state-level regulatory environments may impact the relationship between payday lender presence and premature deaths. For example, strong regulations that improve the affordability of payday loan products, such as capped interest rates and fees, limits on loan rollovers, or extensions of time to repayment, may subsequently attenuate any positive relationship.

Data

We combine data from several sources to develop a novel dataset for this study. First, we obtain historical data on payday lender storefront locations in the United States between 2000 and 2017 from InfoGroup. These data include the address, business name, and annual operating status for every payday lending storefront in the United States. We generate an annual file of active payday lenders using Standard Industry Classification business codes and word searches within company names (e.g., "cash advance", "payday"). We then match geocoded business addresses to county boundaries to generate a county-level data file that captures the number of active storefronts in each county and each year. We then bring in data on premature mortality at the county level using data from the Center for Disease Control and Prevention's (CDC) WONDER database (96). We also include county-level socio-economic and demographic information using data from the Census and American Community Survey (97). Finally, we include annual data on state-level payday lending regulations from the National Conference of State Legislatures (72).

Key measures

Our outcome of interest is derived from the count of *premature deaths* in each county in each year. We define premature deaths as deaths from any cause among 20–59 year-olds, following the approach used by Matheson et al. (81). The CDC suppresses mortality counts between 0 and 9 and considers rates that use counts below 20 deaths to be unreliable. As such, we restrict our analytic sample to counties that have 20 or more premature deaths in a given year. Of the 3,134 eligible counties in the United States, 2,626 meet this criteria for at least 1 year between 2000 and 2017. In secondary models, we also evaluate cause-specific premature deaths for cardiac-, mental health- and assault-related deaths[1]. These models use subsets of counties that have non-suppressed counts of these deaths and seek to provide additional insights on possible individual- and community-level mechanisms linking payday lender presence to premature deaths.

Our key variable of interest is a three-level categorical measure of *payday lender presence*. The reference group is 0

1 We use the CDC WONDER database to pull county-level counts of premature deaths due to mental health [IC10 codes - F01-F99 (mental and behavioral health disorders) and X60 - X84 (intentional self-harm)], cardiac [I10 - I51 (e.g., hypertension, heart disease)], and assault (X85-Y09 "assault").

lenders within a county, which we compare to counties that have 1–3 payday lenders, and those that have 4 or more. We base these categories on the average numbers of alternative financial services storefronts found in previous research (16, 60)[2].

A second variable of interest is a constructed measure of *regulatory strength*. This measure refers to the strength of each state's payday lending regulatory environment in a given year. We use a four category measure. The reference group is states with *no regulations*, which is compared to states with *weak regulations, moderate regulations,* and *strong regulations.* A weak regulatory environment is defined as one where the state has a law on the books requiring payday lending licensing and registration. A moderate regulatory environment refers to those that limit rollovers or require lower interest rates. A strong regulatory environment refers to states that have fully prohibited payday lending or have strict interest rate caps set to 36% APR. For our analysis, we include all states, including those that prohibit payday lending. Supplemental models using only states that allow payday lenders produce similar results.

We include several time-varying control variables to better isolate the relationship between local payday lending environments and mortality outcomes. We include continuous measures of the county's *share of poverty*, *share of male residents, share of Black and Latino residents, share of new residents moving into the county in the prior year,* and *share of urban residents* in a given year. Covariates for race, sex assigned at birth, poverty, and urbanicity are standard controls included in analyses on payday lending [see, e.g., Faber (16)]. We additionally include the measure of county residential mobility as a proxy for duration of exposure to the county environment.

Variation in the population at risk of premature death across counties is accounted for using the *population of individuals ages 20–59* as an exposure term in our models, which converts our count of premature deaths to a rate. All measures come from the Census and the American Community Survey (ACS) (97). We create annual measures for years 2008–2017 using the five-year ACS data with the year of interest as the midpoint. For years 2001 - 2007, we use linear interpolation between the 2000 Census and the 2006–2010 ACS (where 2008 is the midpoint) to generate annual estimates.

Table 1 presents descriptive statistics for all variables used in analysis. The premature mortality count for a typical county is 191 but with substantial variation in these counts, which range from 20 in some counties to well over 5,000 annually in large urban counties like Cook County, IL and Los Angeles County, CA. In our data, 38% of county-years have no payday lenders, about 27% have one to three payday lenders, and about 35% have four or more. In this latter category, it is not uncommon for a county to have numerous lenders in a given year; close to 37% of counties in this category (13% of the overall sample) have 10 or more lenders in at least 1 year. Regulatory environments are also mixed; for example, 47% of county observations are in weak state regulatory environments and 24% are in strong state regulatory environments.

TABLE 1 Descriptive statistics.

	Mean (SD) or Proportion
Outcome Variable	
County Premature Mortality (count)	191.27 (483.28)
County Lender Composition	
0 Lenders in County	0.38
1–3 Lenders in County	0.27
4+ Lenders in County	0.35
State Regulatory Strength (share of county-year observations)	
No Regulations	0.21
Weak Regulations	0.47
Moderate Regulations	0.07
Strong Regulations	0.24
Covariates	
Population Ages 20-59	61,413 (177,369)
Share of Male Residents	0.49
Share of Residents in Poverty	0.15
Share of Black/Latinx Residents	0.18
Share of Movers in the Past Year	0.07
Share of Urban Residents	0.49
n	42,230

2,626 counties contributed 42,230 observations over the analytic period. Covariate statistics refer to the underlying continuous measures that were used to create the categorical measures used in models.

Empirical strategy

We first present descriptive associations between the local payday lending environment and premature deaths. Because our outcome of interest is a count variable across space and time, we next fit longitudinal Poisson regressions with random effects (98). Inclusion of the exposure term, population aged 20–59, adjusts the model results to reflect incidence risk of premature deaths at the county level (i.e., converts the count to a rate). We report incidence risk ratios[3], and standard errors are clustered at the county level in all models. All models include state and year fixed effects to account for omitted variables that vary by state and year.

2 Adjusting the categories to evaluate counties with 0 lenders relative to counties with 1–3, 4–10, and 11 or more lenders produces substantively similar results.

3 With our exposure term, incidence risk is defined as the number of premature deaths divided by the population ages 20–59 (i.e., the population at risk of a premature death) for each county in each year.

We proceed in two stages. First, we model the association between payday lender presence and premature death. In this analysis, results from our first Poisson regression (M1) provide estimates from a model of the change in incidence risk of premature deaths as a function of lender presence, net of state and year fixed effects. M2 adds our set of time-varying county-level controls. M3 presents results from a model that interacts payday lender presence with all covariates. We include an interacted model to underscore that associations between payday lender presence and premature death rates may be conditional on other community characteristics, given known demographic and socio-economic disparities in both premature deaths and payday lender locations.

Second, we evaluate how payday lending regulation moderates the relationship between lender density and premature deaths, first with a model interacting lender presence and regulatory strength net of state and year fixed effects (M4), and then adding the set of time varying controls (M5). This set of models provides insights on whether and how regulation of high-cost lending may reduce premature mortality. For ease of interpretation, we do not interact these models with the demographic and socio-economic covariates.

Limitations

Our analysis is not without limitations. Data limitations include suppression of counties with <20 mortality counts. Main models are not spatially weighted to account for geographic clustering of premature death and payday lender counts due to modeling limitations (see Appendix A for discussion). Although we hypothesized individual- and community-level mechanisms to explain higher-cost lenders' effects on premature death, data limitations make direct tests of these mechanisms suggestive. Future research should attempt to elucidate these explanatory mechanisms, particularly how and the extent to which the concentration of payday lenders represents economic distress, encourages economic distress by dissuading other types of development activities, and contributes to poor public health outcomes. Moreover, payday lenders have not been present for very long by our first year of data, meaning that exposure to lenders could be limited and our data do not fully capture cumulative effects. Any associations that are suggestive of cumulative effects may be underestimated. Finally, we look at overall premature death counts and do not examine the effects of payday lender presence on premature deaths by subgroup (e.g., by sex assigned at birth and race/ethnicity), and future research should examine the potential heterogeneity in effects across subgroups.

Results

Descriptive maps of key variables

Figure 1 presents counts of premature deaths and counts of payday lenders for 2 years in our analysis, 2008 and 2011, using all available data from the CDC and InfoGroup. These years correspond with the start and end of the Great Recession, which impacted both premature death rates (99) and use of payday loans (100). The top panel, which features county-level counts of premature deaths from all causes, shows a consistent clustering of premature deaths in the South, with more counties experiencing premature counts in the 75th percentile or higher in 2011. It also shows that the majority of the suppressed CDC data is largely from rural counties in the central and western United States; these counties will be excluded from our analyses.

The bottom panel, which features counts of payday lenders, shows that the count of counties with four or more lenders (top categories) was strongest in 2008, with fewer counties reporting 4 or more payday lenders in 2011. Much of the decline by 2011 is likely due to regulatory interventions that several states enacted, rather than a decline in demand. Use of payday loans increased during the Great Recession (100), but payday lenders also faced increased oversight in several states, which has an impact on where they operate (101). Many of the counties that have suppressed CDC data also have few payday lenders (e.g., central United States). Beyond that, the two panels suggest positive associations between premature deaths and payday lenders; many of the counties with relatively high counts of premature deaths also have four or more payday lenders (see, e.g., parts of Texas and the Florida panhandle). These descriptive correlations also reflect other factors like urbanicity, which will be accounted for in our modeling.

Spatial autocorrelation of variables

Our models do not account for spatial clustering of premature deaths and payday lenders. Given the spatial clustering observed in Figure 1, we assessed the spatial autocorrelation of the dependent and independent variables for the inclusion of spatial weights in our models using Moran's I. There is spatial autocorrelation for our measures of payday lender counts and premature deaths in each year of our analysis ($p < 0.001$). In appendix A, we compare our main results to those from spatially weighted OLS models, which are less ideal for a count outcome but permit an evaluation of potential spatial spillover effects among neighboring counties. The spatially weighted models indicate that spatial autocorrelation does impact results; the models presented below are likely underestimating the effects of payday lender presence on premature mortality.

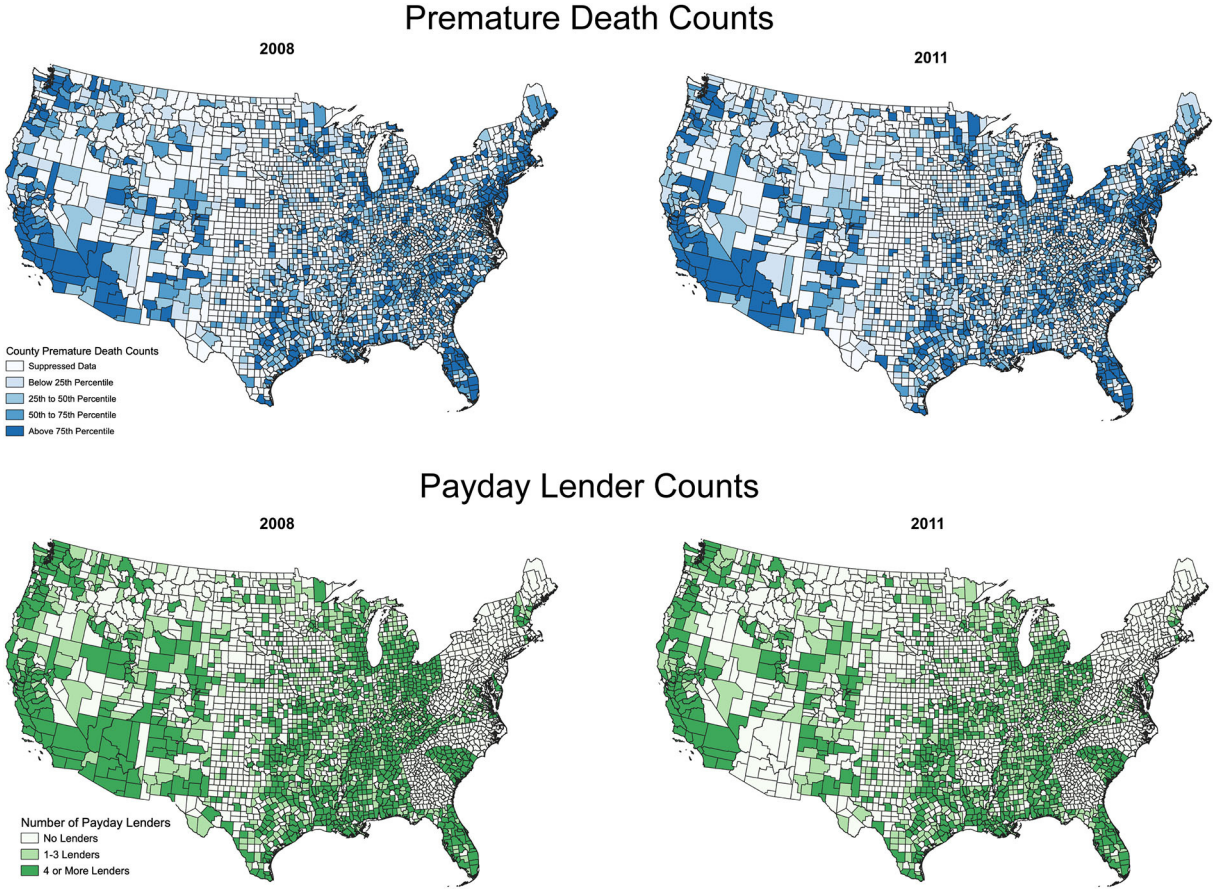


FIGURE 1
Counts of premature deaths and payday lenders over time (by county and year) for counties in the continental United States. Maps created using QGIS software, county boundaries reflect 2010 Census boundaries.

Association between payday lender presence and all-cause premature mortality

Table 2 presents the predicted incidence risk ratios adjusted for state and year fixed effects (M1) and incidence risk ratios that are also adjusted for controls (M2) and interactions with controls (M3). Full model results are available in Appendix B. Model 1, which includes no controls, shows a significant, positive association between counties with four or more lenders and the risk of premature deaths, with a 2% increase in risk of premature death over counties without lenders (RR = 1.020, $p < 0.05$). Counties with 1–3 lenders have little difference in risk of premature death compared to counties without lenders (RR = 1.009, $p > 0.05$). After adjusting for controls, M2 shows that the adjusted risk ratio continues to be significant and positive for counties with 4+ lenders compared to counties without lenders (RR = 1.021, $p < 0.05$). Appendix B also confirms that results for covariates in M2 have directions that are largely in line with expectations. For example, an increase in the share of urban residents is associated with a reduction in risk of premature mortality ($p < 0.05$), in line with prior research that shows that premature deaths are higher in rural areas (89). Further, an increase in the share of mobility in the county is significantly associated with a reduced risk of premature death; other covariates' coefficients have suggestive directions and most are significant (e.g., as the share of male residents increases, there is an association with a slightly elevated risk of premature death, in line with work showing differences in sex assigned at birth in premature deaths ($p < 0.05$).

When the model is interacted (M3), the main effects for counties with 1–3 lenders and 4 or more lenders are significant ($p < 0.05$), as are some interaction effects with county socio-economic and demographic covariates (see Appendix B). As shown in Table 2, counties with 1–3 lenders have a 2.3% increase in risk of premature mortality, and counties with 4 or more lenders have a 1.7% increase in risk, compared to counties without lenders ($p < 0.05$ for both). To put this in context,

10.3389/fpubh.2022.993585

TABLE 2 Unadjusted and adjusted incidence risk ratios for premature deaths, by lender presence.

	M1	M2	M3
No Lenders (reference)	1	1	1
1–3 Lenders	1.009 (0.007)	1.011 (0.006)	1.023* (0.006)
4+ Lenders	1.020* (0.010)	1.021* (0.008)	1.017* (0.007)
Control Variables		Y	Y
Interactions with Controls			Y
State FEs	Y	Y	Y
Year FES	Y	Y	Y
N	42,230 county-years; 2,626 counties		

Incidence risk ratios relative to counties with no lenders; robust standard errors in parentheses; *$p < 0.05$. Exposure term is population ages 20–59. Derived from Models 1–3 in Appendix B using Stata margins and nlcom commands.

Figure 2 presents predicted premature mortality counts by county type. Counties with no lenders have an average predicted premature mortality count of 235.7 deaths when all covariates are at their means. Compared to counties without lenders, the presence of 1–3 lenders was associated with a predicted excess of 4.8 deaths and the presence of 4 or more lenders was associated with a predicted excess of 4.2 deaths[4].

Associations between payday lender presence and cause-specific premature mortality

Though data are more limited, we also evaluate three cause-specific rates of premature mortality, evaluating deaths that stem from mental and behavioral health disorders, cardiac issues, and assaults. Deaths due to mental health (e.g., suicides) in communities with higher-cost lenders may speak to more immediate individual-level mechanisms where stress related to increased debt burdens is linked to premature deaths, given known literature that finds debt can negatively impact mental health [e.g., Fowler et al. (60)]. Cardiac-related deaths may provide a longer-term view of the health consequences of the accumulation of personal debt over time [e.g., Batomen et al. (1), Eisenberg-Guyot et a. (2), Nelson et al. (3), Sweet et al. (4), and Fitch et al. (5)]. Deaths due to assault may speak to broader community-level factors, where the presence of payday lenders as an indicator of economic distress leads to higher risks of these deaths. We know from prior work that the presence of payday lenders has been linked to increased violent crime in local areas [e.g., Kubrin et al. (80)], and assault-related deaths may speak to the more immediate ecological impacts of payday lender presence.

The number of counties with non-suppressed information on these deaths is lower, and thus these analyses are more limited. We evaluate mental health and cardiac related deaths using a shared sample of 1,213 counties with 11,519 observations. We separately evaluate assault related deaths using an even smaller sample of 193 counties and 2,184 observations; this cause of death is rarer and few counties have sufficient counts of deaths to be included in analysis. For the analysis of assault related deaths, we reduce our measure of lender presence to two categories: no lenders, or any lenders: this is because the vast majority of counties with lenders have 4 or more locations (only five counties had one to three lenders). Models for these disaggregated analyses mirror Table 2 using the same covariates and interactions. The smaller sample sizes reduce the precision of estimates, and we caution that these results reflect a select set of counties that may not be nationally representative.

Results in Table 3A show strong evidence of a positive association between payday lender presence and mental-health related deaths. In the model for mental-health related deaths net of covariates (M2), having 1–3 lenders or 4 or more lenders is associated with a substantially higher risk of premature mortality compared to counties with no lenders (RR = 1.178, $p < 0.01$ and RR = 1.167, $p < 0.05$). When interacted with county covariates, the adjusted risks remain elevated, although with large confidence intervals (RR = 1.082, $p > 0.05$, and RR = 1.068, $p > 0.05$, respectively). In these same counties, we see more modest evidence of connections to cardiac-related deaths (M4 - M6). Having 1–3 lenders is positively associated with a higher risk of cardiac-related premature mortality compared to having no lenders; however, the results are only significant for those counties with 1–3 lenders in Model 5 (RR = 1.054, $p < 0.05$). Finally, as shown in Table 3B, in the more limited set of counties for premature deaths due to assaults (M3), there are positive associations between having any lenders in a county compared to no lenders, with magnitudes similar to that for the risk of mental-health related premature deaths (e.g., RR = 1.077, $p > 0.05$ in M3 of Table 3B).

Together, these results show initial support for both individual- and community-level mechanisms, with some indication that the public health impacts of payday lender presence may be more immediate as shown by the large magnitudes of the relative risks for mental health and assault related deaths in counties with lenders. The comparatively more modest associations between cardiac deaths and payday lender presence may be due to the fact that the full cumulative effects of local industry presence on public health have not been realized. In our analytic period, payday lending is relatively new; most storefronts started opening nationwide in the early 2000s, and connections to longer term health issues may not be known for some time. Additional analyses (not shown) of all-cause mortality that include a control

4 Calculated using results from Model 3 in Appendix B, using the margins postestimation commands in Stata. We keep covariates at their means to approximate an average U.S. county during our analytic period.

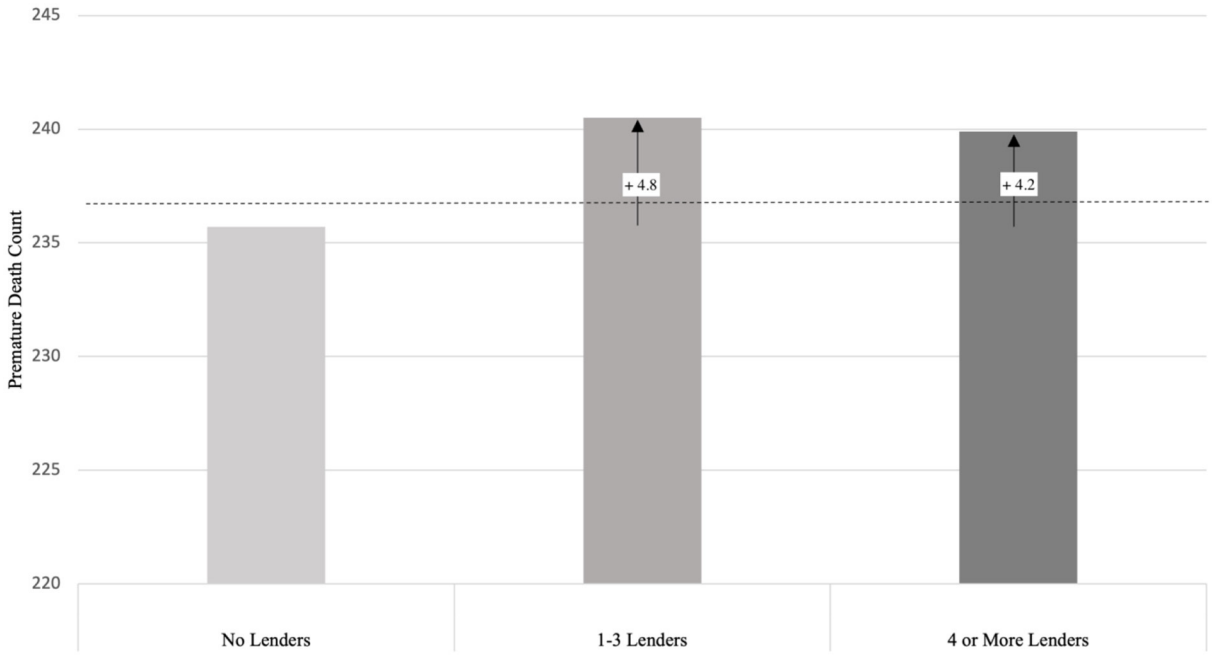


FIGURE 2
Predicted excess premature deaths in counties with lenders. The higher adjusted incidence risk ratios ($p < 0.05$) in M3 of Table 2 for counties with 1 to 3 and 4 or more lenders translate to 4.8 and 4.2 excess deaths, respectively, compared to counties with no lenders. Estimated death counts are derived from Model 3 in Appendix B using Stata *margins* and *lincom* commands. Dashed line represents overall grand mean (238.4 deaths), derived from same model. 2,626 counties contributed 42,230 observations.

TABLE 3A Incidence risk ratios for premature deaths from mental health and cardiac-related causes.

	Mental Health			Cardiac		
	M1	M2	M3	M4	M5	M6
No Lenders (reference)	1	1	1	1	1	1
1–3 Lenders	1.184** (0.076)	1.178** (0.071)	1.082 (0.056)	1.047 (0.024)	1.054* (0.023)	1.043 (0.028)
4+ Lenders	1.139 (0.083)	1.167* (0.079)	1.068 (0.055)	0.984 (0.020)	1.014 (0.018)	0.976 (0.022)
Control Variables		Y	Y		Y	Y
Interactions with Controls			Y			Y
State FEs	Y	Y	Y	Y	Y	Y
Year FES	Y	Y	Y	Y	Y	Y
N	11,519 county-years; 1,213 counties					

Incidence risk ratios relative to counties with no lenders; robust standard errors in parentheses. Exposure term is population ages 20–59. *p < 0.05; **p < 0.01. Full models available upon request.

for state-level credit card debt per capita find that this control has a significant, positive association with premature deaths; though, it does not meaningfully change the payday lender—premature death association. This further suggests that the mechanisms behind the relationship may extend beyond individual-level debt burdens. These analyses remain suggestive; more research is needed to fully understand mechanisms behind the association between communities' payday lender presence and premature deaths.

Impact of regulatory interventions on relationship of interest

We return to our main analysis of all-cause premature mortality and include an interaction between payday lender presence and regulatory strength to understand whether any relationship between lender presence and premature deaths may be dependent upon the regulatory environment. Table 4 presents results of this lender presence by regulation interaction with just

TABLE 3B

	Assault		
	M1	M2	M3
No Lenders (reference)	1	1	1
1 or More Lenders	1.220** (0.091)	1.193* (0.095)	1.077 (0.071)
Control Variables		Y	Y
Interactions with Controls			Y
State FEs	Y	Y	Y
Year FES	Y	Y	Y
N	2,184 county-years; 193 counties		

Incidence risk ratios relative to counties with no lenders; robust standard errors in parentheses. Exposure term is population ages 20–59. *p < 0.05; **p < 0.01. Lender categories collapsed into two groups due to small sample of 1–3 lender counties in this sub-analysis. Full models available upon request.

TABLE 4 Incidence risk ratios for premature deaths by lender presence and regulatory strength.

	M1	M2
No Lenders, No Regulations (references)	1	1
No Lenders, Weak Regulations	1.111*** (0.012)	1.086*** (0.012)
No Lenders, Moderate Regulations	1.160*** (0.023)	1.103*** (0.024)
No Lenders, Strong Regulations	1.071*** (0.016)	1.047*** (0.014)
1–3 Lenders, No Regulations	1.010 (0.008)	1.003 (0.007)
1–3 Lenders, Weak Regulations	1.118*** (0.013)	1.093*** (0.013)
1–3 Lenders, Moderate Regulations	1.143*** (0.024)	1.108*** (0.022)
1–3 Lenders, Strong Regulations	1.068*** (0.016)	1.052*** (0.014)
4+ Lenders, No Regulations	1.078*** (0.016)	1.056*** (0.012)
4+ Lenders, Weak Regulations	1.107*** (0.010)	1.087*** (0.016)
4+ Lenders, Moderate Regulations	1.086*** (0.021)	1.071*** (0.020)
4+ Lenders, Strong Regulations	1.082*** (0.019)	1.054** (0.017)
N	42,230 county-years; 2,626 counties	

Incidence risk ratios relative to counties with no lenders within states with no regulations; robust standard errors in parentheses; ***$p < 0.001$. Exposure term is population ages 20–59. Derived from Models 1 and 2 in Appendix C using Stata *margins* and *nlcom* commands.

state and year fixed effects (M1) and results from a model that also controls for other covariates (M2). These results indicate that the impacts of payday lender presence are conditional on the regulatory context. Full model results are available in Appendix C.

There are significant differences that are large in magnitude for premature death—payday lender associations across regulatory contexts. M2 shows that, relative to counties with no lenders in states without regulations, every other combination of regulatory environment and lender presence has a significantly higher risk of premature mortality. The only exception is among counties with 1–3 lenders in states without regulation, which remains marginally higher (RR = 1.003, $p > 0.05$). In weak and moderate regulatory environments, relative risks range between 1.07 and 1.11, or seven to 11% higher, compared to counties with no lenders and no regulations, while relative risks tend to be lower in strong regulatory environments (around 1.05 for each county type, or 5% higher). These patterns suggest that the variation in the degree of attenuation that regulation can achieve depends on its strength. For example, counties with 1–3 lenders move from a relative risk of 1.09 in a context of weak regulations to 1.11 in a context of moderate regulations; in other words, the risk of premature death actually increases by 2% when moving from weak to moderate regulations. However, the shift from moderate to strong regulations for these counties results in a *reduction in risk* by 6% (moving from 1.11 to 1.05). This suggests that strong regulations do comparatively better in dampening the risk of premature mortality compared to weak and moderate regulations, which actually experience increases in risk compared to even less regulation.

That these patterns are true irrespective of the number of lenders is somewhat puzzling. Similar patterns occur for counties that have no lenders across these regulatory contexts, and we would expect regulation to impact counties with lenders but have little effect on counties without. This might be due in part to compositional changes that occur when regulation takes effect. For example, when a state shifts from moderate to strong regulations, there are some counties where all lenders leave, as shown by prior work (101–103). When this occurs, the county would get reclassified as part of the "no lenders, strong regulations" group in our models. The average adjusted risk of 1.05 for this group might be picking up some of the higher initial risk of counties that had been in the 1–3 or 4+ lender categories under a moderate regulatory context. More research is needed to fully understand why regulation impacts the public health of counties without lenders.

Figure 3 displays estimated relative risks *within* regulatory context, derived from Model 2. This figure does not show increases in levels relative to one common reference group, as shown in Table 4, but rather compares across county lender categories within the same regulatory environment. In the absence of regulation, the relative risk of premature deaths is the highest for counties with four or more lenders (RR = 1.06, $p < 0.001$), and then differences between high concentration counties and those with zero and 1–3 lenders diminish as regulations strengthen. In moderately regulated environments, the risk for counties with four or more lenders is marginally lower relative to that of counties without lenders (RR = 0.97, $p > 0.05$). Relative risks return to parity under strong regulatory environments, all else equal. This within-context comparison

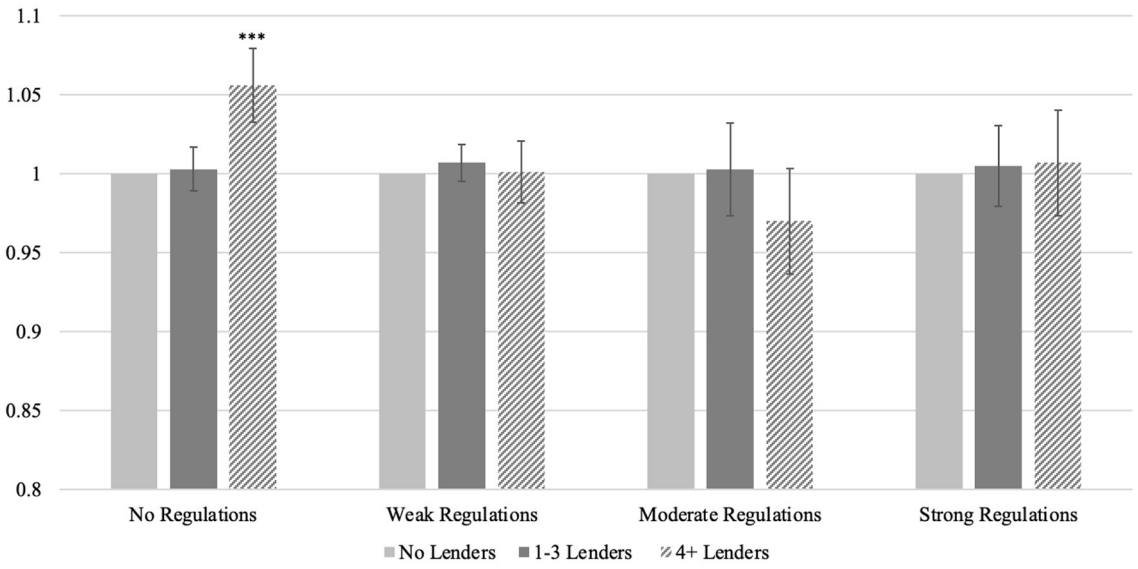


FIGURE 3
Within-regulatory group relative risks for premature deaths. Within-period incidence risk ratios relative to counties with no lenders (left-most column in each group); error bars reflect 95% confidence intervals. ***$p < 0.001$, compared to counties with no lenders. Exposure term is population ages 20–59; 2,626 counties contributed 42,230 observations. Derived from Model 2 in Appendix C using Stata *margins* and *nlcom* commands.

underscores that the lack of regulation impacts public health in counties with high concentrations of lenders the most, and that these counties gain relatively greater public health benefits under contexts with enhanced regulation.

Concluding discussion

The relationship between high debt burdens and poor health raises concerns about the expansion of the payday lending industry, which sells expensive loans that are hard to repay (47, 52) and contribute to borrowers' financial difficulties (12, 13). Robust literatures explore the relationships between payday lenders and financial difficulties [e.g., (12, 77)] and high debt burdens and poor health [e.g., Batomen et al. (1) and Eisenberg-Guyot et al. (2)], laying the groundwork to connect these lines of inquiry. Importantly, as the number of these industry storefronts expanded, some states began trying to protect borrowers by placing new restrictions on payday lenders such as capping usurious interest rates and preventing the renewal or re-borrowing of these loans, which may have public health benefits. We explore these associations spatially at the county level and our findings, described below, offer new pathways for inquiries into payday lending and effective regulation.

Several key findings elucidate any relationships between the distribution of payday lenders and community health. First, we find that after accounting for socioeconomic covariates, the risk of all-cause premature death is significantly higher in counties with four or more payday lenders and one to three lenders, relative to counties without. Secondary analyses that disaggregate causes of premature deaths lend some support for both proposed mechanisms for this association. Higher risks of mental health related deaths and modestly higher risks of cardiac related deaths suggest that residents' longevity in these communities may be compromised by individual-level stress due to increased debt burdens. Higher risks of assault related deaths suggest that exposure to heightened community economic distress as proxied by payday lender presence may also compromise longevity. Because our analysis remains at the county level, it is not possible to fully disentangle community- and individual-level mechanisms; though, taken together, the all-cause and specific cause analyses underscore that payday lender presence is associated with poorer community health, even after accounting for community demographic composition, poverty, and urbanicity.

Second, we find evidence of moderating effects of regulation, whereby the influence of a county's concentration of payday lenders on the risk of all-cause premature death is conditional on regulatory context. These findings provide evidence that better regulations may have beneficial public health impacts in areas with a relatively large number of lenders. This modest attenuation lends some support to the notion that improved regulation may have positive spillover effects on community health. These findings are notable because they allude to the importance of broadening policy conversations on financial regulation to include the effects on social, physical, and mental well-being. Depending on the extent to which

the effectiveness of regulation is evaluated in economic terms, current policy conversations may underestimate the economic benefits of regulation by focusing primarily on financial difficulties and well-being.

Regulation appears to matter even for counties without lenders. This puzzling finding could be explained by the extent to which regulations targeting payday lenders also discourage or supplant other types of storefronts and businesses that contribute to a county's economic marginalization and distress. Similar to regulation's positive spillover effects on community health, perhaps there are also spillover effects onto usurious and other predatory businesses that undermine community health even in absence of payday lenders. Future research should investigate this possibility.

In the United States, geographic inequalities in health and mortality are growing. The substantial spatial variation in mortality rates makes it important to understand the links between local built environments, policy contexts that inform those environments, and public health. For instance, the absence of grocery stores and hospitals from communities—forms of food and healthcare apartheid enabled by policy decisions and that often accompany other indicators of economic marginalization and distress—has implications for public health (104, 105). We provide evidence that the availability of expensive credit also matters, using geographic variation in the presence of payday lenders and connections to mortality. Understanding the contributions of payday and other high-cost lenders to mortality can aid in identifying potential underlying mechanisms and the possibility for regulation to attenuate their effects. We suggest that, in this context, regulation has the potential to protect against premature mortality for some groups. Future research will need to investigate these relationships in the years during and after the COVID-19 pandemic, which notably changed people's life expectancy and experiences with financial difficulties, as well as influenced business turnover and storefront locational decisions.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The Center for Disease Control and Prevention's WONDER database is available for public use (https://wonder.cdc.gov/wonder/help/ucd.html). Annual data on state-level payday lending regulations from the National Conference of State Legislatures is available for public use (https://www.ncsl.org/research/financial-services-and-commerce/payday-lending-state-statutes.aspx). Historical data on payday lender storefront locations in the United States between 2000 and 2017 is available for purchase from InfoGroup or available through some university library subscriptions, such as the University of Michigan. Requests to access these datasets should be directed to https://www.data-axle.com/contact-us/?gclid=CjwKCAjw2rmWBhB4EiwAiJ0mtSffVEq4u6aa2W9P7ScizvyA-RQXNDJt-PpXLJoMQ9EXQHzAMNBPkRoC4kgQAvD_BwE#contact_us_location_3?utm_term=data%20axle%20headquarters&utm_campaign$=$Corporate$+$Brands$+$%7C$+$PR&utm_source$=$google&utm_medium$=$cpc&hsa_acc$=$9152831390&hsa_cam$=$13480869329&hsa_grp$=$123161790413&hsa_ad$=$529715951002&hsa_src$=$g&hsa_tgt$=$kwd-1307381766427&hsa_kw$=$data%20axle%20headquarters&hsa_mt$=$b&hsa_net$=$adwords&hsa_ver$=$3.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

MA's time for this research was supported by the T32 AG00129 grant, awarded to the Center for Demography of Health and Aging at the University of Wisconsin-Madison by the National Institute on Aging.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.993585/full#supplementary-material

References

1. Batomen B, Sweet E, Nandi A. Social inequalities, debt, and health in the United States. SSM Popul Health (2021) 13:100736. doi: 10.1016/j.ssmph.2021.100736

2. Eisenberg-Guyot J, Firth C, Klawitter M, Hajat A. From payday loans to pawnshops: Fringe banking, the unbanked, and health. *Health Aff.* (2018) 37:429–37. doi: 10.1377/hlthaff.2017.1219

3. Nelson MC, Lust K, Story M, Ehlinger E. Credit card debt, stress and key health risk behaviors among college students. *Am J Health Promot.* (2008) 22:400–6. doi: 10.4278/ajhp.22.6.400

4. Sweet E, Kuzawa CW, McDade TW. Short-term lending: Payday loans as risk factors for anxiety, inflammation and poor health. *SSM - Popul Health.* (2018) 5:114. doi: 10.1016/j.ssmph.2018.05.009

5. Fitch C, Hamilton S, Bassett P, Davey R. The relationship between personal debt and mental health: a systematic review. *Mental Health Rev J.* (2011) 16:153–66. doi: 10.1108/13619321111202313

6. Federal Reserve Bank of New York. *Quarterly Report on Household Debt and Credit: 2019 Q3*. New York: Federal Reserve Bank of New York, Research and Statistics Group 2019. Available online at: https://www.newyorkfed.org/microeconomics/hhdc.html

7. Ho JY, Hendi AS. Recent trends in life expectancy across high income countries: retrospective observational study. *BMJ.* (2018) 15:362:k2562. doi: 10.1136/bmj.k2562

8. Woolf SH, Schoomaker H. Life expectancy and mortality rates in the United States, 1959-2017. *JAMA.* (2019) 322:1996–2016. doi: 10.1001/jama.2019.16932

9. Shiels MS, Chernyavskiy P, Anderson WF, Best AF, Haozous EA, Hartge P, et al. Trends in premature mortality in the USA by sex, race, and ethnicity from 1999 to 2014: An analysis of death certificate data. *Lancet.* (2017) 389:1043–54. doi: 10.1016/S0140-6736(17)30187-3

10. Argys LM, Friedson AI, Pitts MM. *Killer debt: The impact of debt on mortality. Federal Reserve Bank of Atlanta Working Paper Series.* (2016). Available online at: https://www.atlantafed.org/-/media/documents/research/publications/wp/2016/14-killer-debt-the-impact-of-debt-on-mortality-2017-04-10.pdf

11. Clayton, M. Liñares-Zegarra J, Wilson JOS. Does debt affect health? Cross country evidence on the debt-health nexus. *Social Sci Med.* (2015) 130:51–8. doi: 10.1016/j.socscimed.2015.02.002

12. Melzer BT. The real costs of credit access: Evidence from the payday lending market. *Q J Econ.* (2011) 126:517–55. doi: 10.1093/qje/qjq009

13. Skiba PM, Tobacman J. Do payday loans cause bankruptcy? *J Law Econ.* (2019) 62:485–518. doi: 10.1086/706201

14. Apgar W, Herbert C. *Subprime Lending and Alternative Financial Service Providers: a Literature Review and Empirical Analysis*. Washington, DC: US Department of Housing and Urban Development. Available online at: http://www.huduser.gov/Publications/pdf/sublending.pdf

15. Caskey JP. *Fringe Banking: Check-Cashing Outlets, Pawnshops, and the Poor*. New York: Russell Sage Foundation (1994). 184 p.

16. Faber JW. Segregation and the cost of money: Race, poverty, and the prevalence of alternative financial institutions. *Social Forces.* (2019) 98:819–48. doi: 10.1093/sf/soy129

17. Shannon J. Food deserts: Governing obesity in the neoliberal city. *Prog Hum Geogr.* (2014) 38:248–66. doi: 10.1177/0309132513484378

18. Yenerall, J, You, W, Hill, J. Investigating the spatial dimension of food access. *Int J Environ Res Public Health*. (2017) 14:866. doi: 10.3390/ijerph14080866

19. Balentine MD. *A Geographical Analysis of US Plasma Centers and Their Economic, Social, and Public Health Impacts [dissertation].* (2020). Greensboro, NC: The University of North Carolina at Greensboro.

20. Robert JC, Mustard CA. Geographic location of commercial plasma donation clinics in the United States, 1980-1995. *Am J Public Health.* (2004) 94:1224–9. doi: 10.2105/AJPH.94.7.1224

21. Shannon J. Dollar stores, retailer redlining, and the metropolitan geographies of precarious consumption. *Ann Am Assoc Geogr.* (2021) 111:1200–18. doi: 10.1080/24694452.2020.1775544

22. Sharkey JR, Horel S, Han D, Huber JC. Association between neighborhood need and spatial access to food stores and fast food restaurants in neighborhoods in Colonias. *Int J Health Geogr.* (2009) 8:1–17. doi: 10.1186/1476-072X-8-9

23. Charron-Chénier R. Predatory inclusion in consumer credit: Explaining black and white disparities in payday loan use. *Sociol Forum.* (2020) 35:370–92. doi: 10.1111/socf.12586

24. Charron-Chénier R, Seamster L. Racialized debts: Racial exclusion from credit tools and information networks. *Crit Sociol.* (2021) 47:977–92. doi: 10.1177/0896920519894635

25. Dwyer RE. Credit, debt, and inequality. *Annu Rev Sociol.* (2018) 44:237–61. doi: 10.1146/annurev-soc-060116-053420

26. Friedline T, Freeman A. The potential for savings accounts to protect young-adult households from unsecured debt in periods of macroeconomic stability and decline. *Soc Serv Review.* (2016) 90:83–129. doi: 10.1086/685791

27. Lin KH, Neely MT. *Divested: Inequality in Financialized America*. Oxford, UK: Oxford University Press. (2020). 240 p.

28. Antonczyk D, Fitzenberger B, Sommerfeld K. Rising wage inequality, the decline of collective bargaining, and the gender wage gap. *Labour Econ.* (2010) 17:835–47. doi: 10.1016/j.labeco.2010.04.008

29. Jacobs D, Myers L. Union strength, neoliberalism, and inequality: Contingent political analyses of U.S. income differences since 1950. *Am Sociol Rev.* (2014) 79:752–74. doi: 10.1177/0003122414536392

30. Wallerstein M, Western B. Unions in decline? What has changed and why. *Annu Rev Polit Sci.* (2000) 3:355–77. doi: 10.1146/annurev.polisci.3.1.355

31. Western B, Rosenfeld J. Unions, norms, and the rise in US wage inequality. *Am Sociol Rev.* (2011) 76:513–37. doi: 10.1177/0003122411414817

32. Autor DH, Katz LF, Kearney MS. Trends in US wage inequality: Revising the revisionists. *Rev Econ Stat.* (2008) 90:300–23. doi: 10.1162/rest.90.2.300

33. Mishel L, Bivens J, Gould E, Shierholz H. *The State of Working America. 12th ed.* New York: ILR Press (2012).

34. Kim C, Sakamoto A. The rise of intra-occupational wage inequality in the United States, 1983 to 2002. *Am Sociol Rev.* (2008) 73:129–57. doi: 10.1177/000312240807300107

35. Piketty T. *Capital in the Twenty-First Century*. Cambridge: Harvard University Press (2014). 816 p.

36. Schmall T, Wolkowitz E. *2016 Financially Underserved Market Size Study. Center for Financial Services Innovation.* (2016) 54 p. Available online at: https://cfsi-innovation-files-2018.s3.amazonaws.com/wp-content/uploads/2017/02/27001759/2016-Financially-Underserved-Market-Size-Study_Center-for-Financial-Services-Innovation.pdf

37. Hecht J. *The Industry, the Customer, the Experience: Online Consumer Lending Industry Trends and Observations*. Miami: Online Lending Alliance Conference (2012).

38. Peterson A. *Predatory Payday Lending*. Washington, DC: Center for American Progress (2013). Available online at: https://www.americanprogress.org/article/predatory-payday-lending/

39. Brown T, Swartz M. Competition policy on consumer financial services: The disparate regulation of online marketplace lenders and bankers. *Compet Policy Int.* (2016) 11:38–45. Available online at: https://www.competitionpolicyinternational.com/wp-content/uploads/2016/03/Competition-Policy-in-Consumer.pdf

40. Nuñez S, Schaberg K, Hendra R, Servon L, Addo M, Mapillero-Colomina A. Online payday and installment loans: Who uses them and why? New York, NY: MDRC (2016). 134 p.

41. Wang H, Overby EM. How does online lending influence bankruptcy filings? Evidence from a natural experiment. *SSRN Journal.* (2017). Available online at: http://www.ssrn.com/abstract=2958916

42. Borck R, Fossen FM, Freier R, Martin T. Race to the debt trap?—Spatial econometric evidence on debt in German municipalities. *Reg Sci Urban Econ.* (2015) 53:20–37. doi: 10.1016/j.regsciurbeco.2015.04.003

43. O'Neill KK, Kennedy I, Harris A. Debtors' blocks: How monetary sanctions make between-neighborhood racial and economic inequalities worse. *Sociol Race Ethnic.* (2022) 8:43–61. doi: 10.1177/23326492211057817

44. Simone D, Walks A. Immigration, race, mortgage lending, and the geography of debt in Canada's global cities. *Geoforum.* (2019) 98:286–99. doi: 10.1016/j.geoforum.2017.10.006

45. Pew Charitable Trusts. *Payday Loan Facts and the CFPB's Impact*. Washington, DC: Pew Charitable Trusts (2016). 2 p.

46. Kindle PA, RB. Do we know what we think we know about payday loan borrowers? Evidence from the Survey of Consumer Finances. *J Soc Soc Welfare.* (2017) 44:19–44. Available online at: https://heinonline.org/HOL/LandingPage?handle=hein.journals/jrlsasw44&div=48&id=&page=

47. Pew Charitable Trusts. *Payday Lending in America: Who Borrows, Where They Borrow, and Why.* Washington, DC: Pew Charitable Trusts (2012). 48 p.

48. Gathergood J, Guttman-Kenney B, Hunt S. How do payday loans affect borrowers? *Evidence from the UK market The Review of Financial Studies.* (2018) 32:496–523. doi: 10.1093/rfs/hhy090

49. Bickham T, Lim Y. In sickness and in debt: Do mounting medical bills predict payday loan debt?. *Soc Work Health Care.* (2015) 54:518–31. doi: 10.1080/00981389.2015.1038410

50. Fitzpatrick K, Coleman-Jensen A. Food on the fringe: Food insecurity and the use of payday loans. *Soc Serv Rev.* (2014) 88:553–93. doi: 10.1086/679388

51. Morgan DP, Strain MR, Seblani I. How payday credit access affects overdrafts and other outcomes. *J Money Credit Bank.* (2012) 44:519–31. doi: 10.1111/j.1538-4616.2011.00499.x

52. Consumer Financial Protection Bureau. *CFPB Data Point: Payday Lending.* Washington, DC: CFPB Office of Research (2014). 33 p.

53. Friedline T, Kepple N. Does community access to alternative financial services relate to individuals' use of these services? Beyond individual explanations. *J Consum Policy.* (2017) 40:51–79. doi: 10.1007/s10603-016-9331-y

54. Baradaran M. *The Color of Money: Black Banks and the Racial Wealth Gap.* Cambridge, MA: Harvard University Press (2017). doi: 10.4159/9780674982284

55. Charron-Chénier R. Payday loans and household spending: How access to payday lending shapes the racial consumption gap. *Soc Sci Res.* (2018) 76:40–54. doi: 10.1016/j.ssresearch.2018.08.004

56. Cover J, Fuhrman Spring A, Garshick Kleit R. Minorities on the margins? The spatial organization of fringe banking services. *J Urban Affairs.* (2011) 33:317–44. doi: 10.1111/j.1467-9906.2011.00553.x

57. Dunham IM, Foster A. Proximate landscapes of economic inclusion in Southeastern Pennsylvania. *Prof Geogr.* (2015) 67:132–44. doi: 10.1080/00330124.2014.907697

58. Faber JW. Cashing in on distress: The expansion of fringe financial institutions during the Great Recession. *Urban Affairs Rev.* (2018) 54:663–96. doi: 10.1177/1078087416684037

59. Federal Deposit Insurance Corporation. Alternative financial services: a primer. *FDIC Q.* (2009) 3:39–47. Available online at: https://www.fdic.gov/analysis/quarterly-banking-profile/fdic-quarterly/2009-vol3-1/fdic140-quarterlyvol3no1-afs-final.pdf

60. Fowler CS, Cover JK, Kleit RG. The geography of fringe banking. *J Reg Sci.* (2014) 54:688–710. doi: 10.1111/jors.12144

61. Friedline T, Dunham IM, O'Brien M. The financial services environment and schools' savings rates in the San Francisco Kindergarten to College Program. *J Consum Affairs.* (2019) 53:1797–824. doi: 10.1111/joca.12264

62. Gallmeyer A, Roberts WT. Payday lenders and economically distressed communities: a spatial analysis of financial predation. *Soc Sci J.* (2009) 46:521–38. doi: 10.1016/j.soscij.2009.02.008

63. Graves SM. Landscapes of predation, landscapes of neglect: a location analysis of payday lenders and banks. *Prof Geogr.* (2003) 55:303–17. doi: 10.1111/0033-0124.5503017

64. Jorgensen M, Akee R. *Access to Capital and Credit in Native Communities: A Data Review, Digital Version.* Tuscon, AZ: Native Nations Institute (2017). 62 p. Available online at: http://nni.arizona.edu/application/files/6514/8642/4513/Accessing_Capital_and_Credit_in_Native_Communities__A_Data_Review.pdf

65. Agarwal S, Gross T, Mazumder B. How did the Great Recession affect payday loans? *Econ Perspect.* (2016) 2:1–12. Available online at: https://fraser.stlouisfed.org/files/docs/historical/frbchi/economicperspectives/frbchi_econper_2016n2.pdf

66. Kasper SD. Payday lending: the case of Tennessee. *J Econ Issues.* (2014) 48:905–25. doi: 10.2753/JEI0021-3624480402

67. Richardson K. *Petty debt: a comprehensive examination of Michigan's Deferred Presentment Industry and its growth between 2000-2018. [dissertation].* (2019). Lund, Sweden: Lunds Universitet.

68. Rothstein D, Dillman JD. *The Continued Growth of Payday Lending in Ohio.* Columbus, OH: Policy Matters Ohio (2008). 18 p.

69. Barth JR, Hilliard J, Jahera JS, Sun Y. Do state regulations affect payday lender concentration? *J Econ Bus.* (2016) 84:14–29. doi: 10.1016/j.jeconbus.2015.08.001

70. Bhutta N. Payday loans and consumer financial health. *J Bank Fin.* (2014) 47:230–42. doi: 10.1016/j.jbankfin.2014.04.024

71. Bhutta N, Goldin J, Homonoff T. Consumer borrowing after payday loan bans. *J Law Econ.* (2016) 59:225–59. doi: 10.1086/686033

72. National Conference of State Legislatures. *Payday Lending State Statutes.* (2021). Available online at: https://www.ncsl.org/research/financial-services-and-commerce/payday-lending-state-statutes.aspx (accessed on February 3, 2022).

73. Carter SP. Payday loan and pawnshop usage: The impact of allowing payday loan rollovers. *J Consum Affairs.* (2015) 49:436–56. doi: 10.1111/joca.12072

74. Dobbie W, Skiba PM. Information asymmetries in consumer credit markets: Evidence from payday lending. *Am Econ J Appl Econ.* (2013) 5:256–82. doi: 10.1257/app.5.4.256

75. Hodson R, Dwyer RE, Neilson LA. Credit card blues: The middle class and the hidden costs of easy credit. *Sociol Q.* (2014) 55:315–40. doi: 10.1111/tsq.12059

76. Sweet E, DuBois LZ, Stanley F. Embodied neoliberalism: Epidemiology and the lived experience of consumer debt. *Int J Health Serv.* (2018) 48:495–511. doi: 10.1177/0020731418776580

77. Allen H, Swanson A, Wang J, Gross T. Early Medicaid expansion associated with reduced payday borrowing in California. *Health Aff.* (2017) 36:1769–76. doi: 10.1377/hlthaff.2017.0369

78. Hundley CL, Wilson RW, Chenault J, Smimble JL. Is fringe banking a component of socially determined health? *Am J Health Stud.* (2017) 32:149–55. doi: 10.47779/ajhs.2017.83

79. Knapp EA, Bilal U, Dean LT, Lazo M, Celentano DD. Economic insecurity and deaths of despair in US counties. *Am J Epidemiol.* (2019) 188:2131–9. doi: 10.1093/aje/kwz103

80. Kubrin CE, Squires GD, Graves SM, Ousey GC. Does fringe banking exacerbate neighborhood crime rates? *Criminol Public Policy.* (2011) 10:437–66. doi: 10.1111/j.1745-9133.2011.00719.x

81. Matheson FI, Creatore MI, Gozdyra P, Park AL, Ray JG, A. population-based study of premature mortality in relation to neighbourhood density of alcohol sales and cheque cashing outlets in Toronto, Canada. *BMJ Open.* (2014) 4:e006032. doi: 10.1136/bmjopen-2014-006032

82. Collin L, Gaglioti A, Beyer K, Zhou Y, Moore M, Nash R, et al. Neighborhood-level redlining and lending bias are associated with breast cancer mortality in a large and diverse metropolitan area. *Cancer Epidemiol Prevent Biomark.* (2021) 30:53–60. doi: 10.1158/1055-9965.EPI-20-1038

83. Denney J, Saint Onge J, Dennis J. Neighborhood concentrated disadvantage and adult mortality: insights for racial and ethnic differences. *Popul Res Policy Rev.* (2018) 37:301–21. doi: 10.1007/s11113-018-9461-9

84. Finkelstein A, Gentzkow M, Williams H. Place-Based Drivers of Mortality: Evidence from Migration. *Am Econ Rev.* (2019) 111:2697–735. doi: 10.3386/w25975

85. Miller S, Soo CK. Do neighborhoods affect credit market decisions of low-income borrowers? Evidence from the moving to opportunity. *Exp Rev Fin Stud.* (2021) 34:827–63. doi: 10.1093/rfs/hhaa060

86. Cylus J, Glymour MM, Avendano M. Health effects of unemployment benefit program generosity. *Am J Public Health.* (2015) 105:317–23. doi: 10.2105/AJPH.2014.302253

87. Cylus J, Glymour MM, Avendano M. Do generous unemployment benefit programs reduce suicide rates? A state fixed-effect analysis covering 1968-2008. *Am J Epidemiol.* (2014) 180:45–52. doi: 10.1093/aje/kwu106

88. Freisthler B. A spatial analysis of social disorganization, alcohol access, and rates of child maltreatment in neighborhoods. *Child Youth Serv Rev.* (2004) 26:803–19. doi: 10.1016/j.childyouth.2004.02.022

89. Freisthler B, Gruenewald PJ. Where the individual meets the ecological: a study of parent drinking patterns, alcohol outlets, and child physical abuse. *Alcohol Clin Exp Res.* (2013) 37:993–1000. doi: 10.1111/acer.12059

90. Freisthler B, Maguire-Jack K. Understanding the interplay between neighborhood structural factors, social processes, and alcohol outlets on child physical abuse. *Child Maltreat.* (2015) 20:268–77. doi: 10.1177/1077559515598000

91. Morton CM, Simmel C, Peterson NA. Neighborhood alcohol outlet density and rates of child abuse and neglect: moderating effects of access to substance abuse services. *Child Abuse Negl.* (2014) 38:952–61. doi: 10.1016/j.chiabu.2014.01.002

92. Contreras C. A block-level analysis of medical marijuana dispensaries and crime in the city of Los Angeles. *Justice Q.* (2017) 34:1069–95. doi: 10.1080/07418825.2016.1270346

93. Freisthler B, Ponicki WR, Gaidus A, Gruenewald PJ, A. micro-temporal geospatial analysis of medical marijuana dispensaries and crime in Long Beach California. *Addiction.* (2016) 111:1027–35. doi: 10.1111/add.13301

94. Kepple NJ, Freisthler B. Exploring the ecological association between crime and medical marijuana dispensaries. *J Stud Alcohol Drugs.* (2012) 73:523–30. doi: 10.15288/jsad.2012.73.523

95. Lee AM, Gainey R, Triplett R. Banking options and neighborhood crime: does fringe banking increase neighborhood crime? *Am J Crim Just.* (2014) 39:549–70. doi: 10.1007/s12103-013-9228-0

96. Centers for Disease Control and Prevention. *Underlying Cause of Death, 1999-2020.* Available online at: https://wonder.cdc.gov/ucd-icd10.html [Accessed June 1, 2022].

97. Manson, Steven, Schroeder, Jonathan, Van Riper, David, Kugler, Tracy, Ruggles, Steven. *National Historical Geographic Information System: Version 16.0.* Minneapolis, MN: IPUMS (2021). Available online at: https://www.nhgis.org/

98. Cameron, AC, Trivedi, PK. *Regression Analysis of Count Data*. New York: Cambridge University Press (2013). 566 p. doi: 10.1017/CBO9781139013567

99. Strumpf EC, Charters TJ, Harper S, Nandi A. Did the Great Recession affect mortality rates in the metropolitan United States? Effects on mortality by age, gender and cause of death. *Soc Sci Med.* (2017) 189:11–6. doi: 10.1016/j.socscimed.2017.07.016

100. Lee J, Kim KT. The increase in payday loans and damaged credit after the Great Recession. *J Fam Econ Iss.* (2018) 39:360–9. doi: 10.1007/s10834-017-9557-0

101. Ramirez SR. Regulation and the payday lending industry. *Contemp Econ Policy.* (2020) 38:675–93. doi: 10.1111/coep.12469

102. Bea M, Amorim M, Friedline T. *Public Cash Assistance and Spatial Predation: How State Cash-Transfer Environments Shape High-Interest Lender Geography*. Madison, WI: University of Wisconsin-Madison (2022).

103. Bea MD, Bley K. (Un)conditional consumer protections in high-cost lending regulation: Impacts on local lending geographies. *J Consum Affairs*. (2022) 1–32. doi: 10.1111/joca.12481

104. Bower KM, Thorpe RJ, Rohde C, Gaskin DJ. The intersection of neighborhood racial segregation, poverty, and urbanicity and its impact on food store availability in the United States. *Prev Med.* (2014) 58:33–9. doi: 10.1016/j.ypmed.2013.10.010

105. Nguyen C, Chernew M. Ostrer I, Beaulieu N. Comparison of healthcare delivery systems in low- and high-income communities. *Am J Account Care* (2019) 7:11–18. Available online at: https://www.ajmc.com/view/comparison-of-healthcare-delivery-systems-in-low-and-highincome-communities

frontiers | Frontiers in Public Health

TYPE General Commentary
PUBLISHED 01 November 2022
DOI 10.3389/fpubh.2022.995751

OPEN ACCESS

EDITED BY
Shannon Monnat,
Syracuse University, United States

REVIEWED BY
Suryakant Yadav,
International Institute for Population Sciences (IIPS), India
Haowei Wang,
Tufts University, United States

*CORRESPONDENCE
Preeti Pushpalata Zanwar
preeti.zanwar@jefferson.edu

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 16 July 2022
ACCEPTED 29 September 2022
PUBLISHED 01 November 2022

CITATION
Zanwar PP, Wallace KL, Soria C and Perianayagam A (2022) Commentary: Examining contextual factors contributing to differentials in COVID-19 mortality in U.S. vs. India. *Front. Public Health* 10:995751. doi: 10.3389/fpubh.2022.995751

Commentary: Examining contextual factors contributing to differentials in COVID-19 mortality in U.S. vs. India

Preeti Pushpalata Zanwar[1,2]*, Katrine L. Wallace[1,3], Christopher Soria[4] and Arokiasamy Perianayagam[5,6]

[1]Jefferson College of Population Health, Thomas Jefferson University, Philadelphia, PA, United States, [2]National Institutes on Aging (NIA) Funded Network on Life Course Health Dynamics and Disparities, University of Southern California, Los Angeles, CA, United States, [3]School of Public Health, University of Illinois Chicago, Chicago, IL, United States, [4]Demography Department, University of California, Berkeley, Berkeley, CA, United States, [5]International Institute for Population Sciences, Mumbai, India, [6]National Council of Applied Economic Research (NCAER), New Delhi, India

KEYWORDS

aging, COVID-19 mortality, excess mortality, disparities, inequalities, COVID-19 pandemic, population demographics

A Commentary on
Examining contextual factors contributing to differentials in COVID-19 mortality in U.S. vs. India

Introduction

As of July 2022, the cumulative, confirmed COVID-19 mortality rates in the United States (U.S.) and India are roughly 3,000 and 370 deaths per 100,000 population, respectively (1). Rates are far lower in India vs. the United States, despite the U.S., dedicating 14% more of their gross domestic product (GDP) on healthcare spending than India (Table 1). However, COVID-19 mortality rates in India have been highly contested. COVID-19 mortality in India has been referred to as the "Indian death paradox" (2) because the death rate is disproportionately low relative to the number of COVID-19 cases the country has experienced. We examine the two most evidence-based and plausible explanations of India's low death rates during the COVID-19 pandemic, which are (1) demographic dynamics and (2) factors contributing to increasing divergence in COVID-19 mortality in the U.S. vs. India, including the systematic undercounting of deaths in India.

TABLE 1 Economic and demographic inequalities in the United States vs. India.

Population characteristics	United States	India
Economic inequalities		
Income classification	High-income	Lower middle income
GNI per capita, USD (2021)	70,430	2,170
GDP spent on healthcare (%)	17	3
Overall development (United Nations)	Developed	Developing
Demographic inequalities		
2022 Population	334.81 M	1.41 B
Men (n, %)	162.76 M, 48.6	730 M, 52
Women (n, %)	171 M, 51.1	675 M, 48.1
Age > 65 years (n, %)	56 M, 16.9	98 M, 7.0
Life expectancy in years		
Overall	79.1	70.4
Female: Male	81.7: 76.6	71.8: 69.2
Age dependency ratio (2021)*	55	48
Fertility rate (births per women) (2022)	1.782	2.159

Notes: M, Millions; B, Billions, GNI, Gross National Income; GDP, Gross Domestic Product.
*Age dependency ratio (% of working-age population).
Source: https://worldpopulationreview.com/country-rankings/high-income-countries.
https://www.worldometers.info/demographics/life-expectancy/.
https://data.worldbank.org/indicator/SP.POP.DPND?locations=IN.
https://www.statista.com/statistics/737923/us-population-by-gender/.
https://www.theglobalstatistics.com/india-population-statistics/#:$\sim$:text=In%202022%2C%20males%20make%20up,total%20population%20at%20675%20million.
https://www.americashealthrankings.org/explore/senior/measure/pct_65plus/state/ALL?edition-year=2022.

Demographic dynamics

Cumulative global COVID-19 data show older adults ages 65 years and above, and older men, to be extremely vulnerable to COVID-19 fatality. Early in the pandemic, 8 out of 10 COVID-19 reported deaths in the United States were in older adults 65 years and above (3). Overall, the age gradients of COVID-19 mortality are steeper in the low mortality settings of high-income countries. The case fatality rate (CFR) of COVID-19 for older adults ages 65 years and above is 15 times higher than for younger adults (3, 4). Low-and middle-income countries display relatively moderate age gradients of COVID-19, where population age structures are comparatively young. In India, data released by the Ministry of Health and Family Welfare (MoHFW) show the age-wise share of COVID-19 deaths at older ages (i.e., 60 years and above) is ∼50%, with 46% of deaths occurring in the prime adult ages of 30–60 years, with a median age at death of 60 years. This is almost 20 years lower than what applies in high-income countries. Although India's death toll is large due to the sheer size of its population, its COVID-19 mortality rate per capita is relatively low due to a relatively young age structure. Currently, only about 7% of India's population is 65 years and older. By 2050, however, that number is expected to approximately double to about 15%, and double again to ∼30% by the year 2100 (5). Since the risk of dying of COVID-19 increases with age globally and is greatest in older adults, one plausible theory is that India, with a younger demographic makeup, has a lower risk of death on a population level.

Additionally, other demographic characteristics, such as gender, contributed to cross-national COVID-19 death rate disparities. In the United States, as of September 2022, 55.1% of COVID-19 deaths were in men, 44.9% of COVID-19 deaths were in women while COVID-19 death counts for other sex and gender minorities are not available (6). Similarly, in India, as of 18 May 2021, 64% of deaths were in men while 36% of deaths were in women, with no data on other sex and gender minorities (7). As of 9 April 2022, the male to female ratio for deaths in the U.S. among confirmed cases was 1.39, with the oldest aged men (80 plus years) at the greatest disadvantage for COVID-19 deaths, followed by women of 80 plus years. However, early in the pandemic in India, men were at a greater disadvantage than women with a case fatality rate (CFR) of 3.3 and 2.9%, respectively, where the CFR was estimated as the ratio of confirmed deaths to the total number of confirmed cases (8).

Causes and factors contributing to differentials in COVID-19 incidence and fatality

Globally, the age pattern of the COVID-19 pandemic cases points to a near normal distribution with three-fifths of the cases concentrated in adults aged 30–65 years. In contrast, the fatality rate is very high for older adults aged 60 years and above compared with younger adults aged <60 years. Second, the age gradients of COVID-19 mortality demonstrate sharp differences between high- and middle-income countries; with stepper age gradients for high-income countries compared with moderating age gradient for lower middle-income countries, such as India. Third, biological sex differences are shown to contribute to excess male COVID-19 deaths at older ages. However, data suggest gender-associated risk of exposure may affect rates of infection and fatality differently for men and women (9). Fourth, during the pandemic, the age-associated chronic diseases has taken millions of lives annually and a larger share of younger lives ages 30–70 years in low- and

middle-income countries (LMIC's) comprise the global burden (10). The premature onset of chronic health conditions among older adults aged 45 years and above is more common in LIMCs with workforces in these ages exposed to heightened risks of contracting COVID-19; this shifts the share of COVID-19 mortality from older adult ages to younger adult ages of 30–60 years. About 90% of COVID-19 deaths of older adults are assigned causes associated with pre-existing chronic health conditions, such as cardiovascular diseases, diabetes, kidney disease, respiratory system disease, and cancer (11, 12). Finally, the huge humanitarian and economic crisis of the COVID-19 pandemic led to a disproportionate burden on the health and mortality impact on the poor and vulnerable, bringing existing socioeconomic disparities in health into sharper focus.

Socioeconomic disparities in COVID-19 incidence and fatality

Across the world, COVID-19 incidence and fatalities are distributed unequally among those with different levels of material and social deprivation (13). People living in congested poor urban settlements, migrant wage earners in low- and middle-income countries, racial minorities in high-income countries, and the frontline healthcare workforce at the bottom of the hierarchical spectrum have borne a disproportionately heavy burden of infection and death. The COVID-19 pandemic has brought such existing socioeconomic health disparities into sharper focus both in the United States and in India (14). India and the United States have different and unique underlying socioeconomic and structural determinants across many domains and levels of influence (9, 15–19). These so called upstream drivers of disparities rooted in historical, social, political, and cultural determinants have intersected during the COVID-19 pandemic and have amplified and exposed the underlying and existing disparities within specific country contexts, thereby creating inequities in COVID-19 mortality. The demographic and socioeconomic disparities in the health gradient provide an important framework to deepen the understanding of, and to mitigate, the health equity effects of the disease (13). According to the OXFAM International report on extreme inequality numbers in India, the lack of universal health coverage in India is one of the drivers of socio-economic inequalities in the health sector, which disproportionately affects health outcomes of marginalized communities (20).

Additionally, the waves of COVID-19 lockdown and related measures across States and Union Territories in India have led to state variations in COVID-19 incidence and deaths (Table 1). There has been a huge humanitarian and economic crisis with heavier burdens of COVID-19-related health consequences for the poor and vulnerable. Excess mortality was greater in rural, less affluent areas in India. With the worldwide healthcare infrastructure and human resources of health set exclusively for COVID-19 priority during the lockdown, people in urgent/emergency need of critical healthcare for pre-existing chronic health conditions, acute conditions, and maternity and childcare, faced either no access to care or extremely limited access. This led to a sizeable number of deaths from other causes; most countries reported disruption to healthcare services for non-communicable diseases (NCDs) (12).

Undercounting of deaths in India

Measuring mortality in India is difficult. About half of India's deaths occur at home (21). Of 10 million deaths every year, over 3 million are not registered and over 8 million have not undergone medical certification (22). A quarter of women's deaths are not counted (23) and registration is especially low in the poorest states, such as Uttar Pradesh and Bihar (24).

In contrast, the completeness and accuracy of the US mortality reporting are generally more reliable, even during the COVID-19 pandemic (25). However, even in the United States, COVID-19 deaths have been somewhat undercounted in certain areas of the country, namely, rural areas and in the southern states (26). The Centers for Disease Control and Prevention (CDC) estimates that overall U.S. excess deaths between March 2020 and March 2022 totaled 1,105,736, which was only 15% more than the 958,864 official death toll reported from COVID-19 over that period (3).

Methodologically, excess mortality is the best way to estimate the total net mortality burden of the COVID-19 pandemic. Excess mortality is the difference between the *observed* deaths and the *expected* number of deaths during the same time period, based on data from years before the pandemic. Thus, the "excess deaths" estimate the extra number of deaths from all causes during the pandemic relative to what would have been expected, had the pandemic not occurred.

In May 2022, the World Health Organization (WHO) released a report estimating the global mortality burden of the COVID-19 pandemic to be 14.9 million excess deaths, in contrast to the 5.4 million COVID-19 deaths that had been officially reported. This report claims, of these deaths, a cumulative 4.7 million deaths due to COVID-19 occurred in India alone (27), whereas India's official estimates place the total at ∼525,000 as of July 2022 (22). Furthermore, two other large studies of deaths in India during the COVID-19 pandemic also estimated India's excess mortality to be 6–8.3 times higher than expected (22, 24), with much variability in the ratios of observed to expected between Indian states (from 0.96 in Goa to 26.7 in Bihar) (24). For a more realistic count of excess mortality, researchers must wait until all-cause mortality data for the COVID-19 period (March 2020 to December 2021) is released from the Sample Registration System (SRS), possibly in 2024.

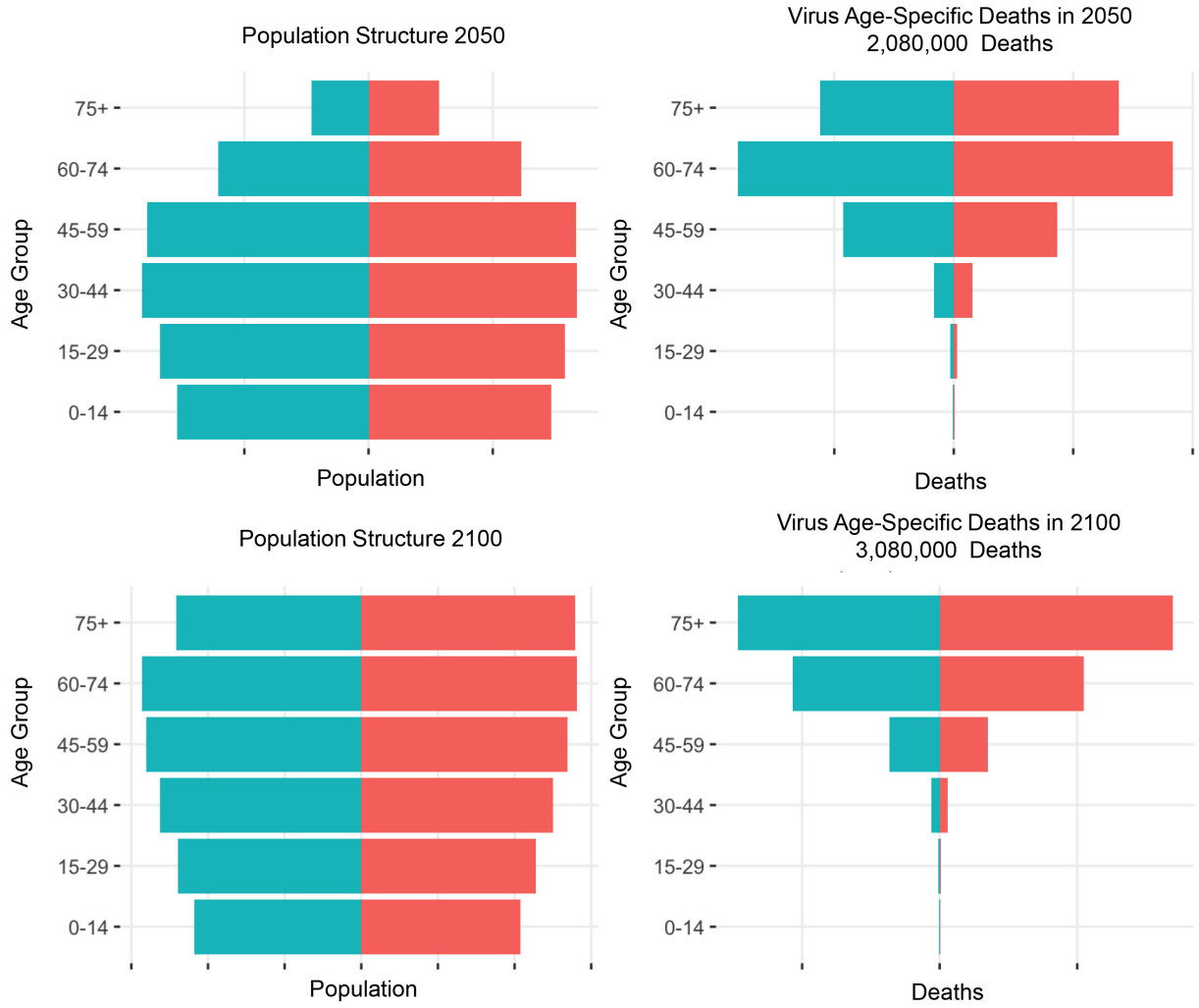


FIGURE 1
Population projections of age structure and virus age-specific deaths in India. United Nations. Department of Economic and Social Affairs, Population Division, World population prospects 2022, Source: https://population.un.org/wpp/Download/Standard/Population/.

Conclusion

A future pandemic is inevitable. A pandemic caused by a virus with mortality rates similar to severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), which kills proportionately more older people in a pattern that follows "Gompertz' Law" (28), would become increasingly deadly as India's falling fertility rates, negative net migration of mostly young people, and rapidly growing life expectancy (29) and rising longevity produce proportionally older populations over time (Figure 1). The same "COVID-19-like" virus scaled to kill 1 million people in India in 2022 would kill just over 2 million in 2050, and 3 million in 2100. Thus, changes in age structure alone would result in more deadly pandemics in the near and distant future. A potential limitation of this assessment, and of health information in India in general, is the lack of consistency and completeness in public health data and vital statistics. The contrast between the United States and India's excess mortality highlights these disparities in the public health infrastructure, limited medical care for critically ill patients in a large-scale health crisis, and a lack of vital recordkeeping capacities. Before the next pandemic, developing nations, such as India, need to prioritize investments in more resilient health systems that can sustain essential health services during crises, including stronger health information systems.

Author contributions

PPZ drafted Table 1. CS contributed to Figure 1. KLW and PPZ revised the commentary. PPZ, KLW, CS, and AP

contributed to the conception and writing of this commentary. All authors contributed to the article and approved the submitted version.

Acknowledgments

We thank Berkeley's Eight Annual Formal Demography Workshop for the inspiration for this commentary and for bringing workshop participants, PPZ and CS together. We thank Jefferson College of Population Health Dean, Billy Oglesby, Ph.D., MBA, MSPH, FACHE for bringing PPZ and KLW together and Prof. Sonalde Desai, Ph.D., M.A., University of Maryland for bringing PPZ and AP together.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Our World in Data. *Coronavirus (COVID-19) Deaths*. (2022). Available online at: https://ourworldindata.org/COVID-deaths (accessed October 15, 2022).

2. Cohen J. Is India's COVID-19 death paradox vanishing? *Science*. (2021) 372:552–3. doi: 10.1126/science.372.6542.552

3. Centers for Disease Control and Prevention (CDC). *Excess Deaths Associated with COVID-19*. (2022). Available online at: https://www.cdc.gov/nchs/nvss/vsrr/COVID19/excess_deaths.htm (accessed October 15, 2022).

4. Oke J, Heneghan C. *Oxford Cocid-19 Evidence Service CEBM Research*. Nuffield Department of Primary Health Care, Oxford (2020). Available online at: https://www.cebm.net/COVID-19/global-COVID-19-case-fatality-rates/ (accessed October 15, 2022).

5. United Nations. *World population prospects - population division*. United Nations(2022). Available online at: https://population.un.org/wpp/Download/Standard/Population/ (accessed September 25, 2022).

6. Health Equity Tracker. *Investigate Rates of COVID-19 in the United States*. (2022). Availoable online at: https://healthequitytracker.org/exploredata?gclid=CjwKCAjw-L-ZBhB4EiwA76YzOT0Ev7j1KkqLT81Or4Cbf5t3WX49xdOXjAEG0OFZF__xzg54JV8WXRoCLYQQAvD_BwEanddemo=sexanddt1=covid_deaths (accessed October 15, 2022).

7. The COVID-19 sex Disaggregated Data Tracker. *Global 50, Health 50*. International Center for Research on Women. African Population and Health Research Center (2020). Available online at: https://globalhealth5050.org/the-sex-gender-and-covid-19-project/the-data-tracker/?explore=country (accessed October 15, 2022).

8. Dehingia N, Raj A. Sex differences in COVID-19 case fatality: do we know enough? *Lancet Glob Health*. (2021) 9:e14–5. doi: 10.1016/S2214-109X(20)30464-2

9. Perianayagam A, Bloom D, Lee J, Parasuraman S, Sekher TV, Mohanty SK, et al. Cohort profile: the longitudinal ageing study in India (LASI). *Int J Epidemiol*. (2022) 51:e167–76. doi: 10.1093/ije/dyab266

10. World Health Organization. *Noncommunicable Diseases*. World Health Organization (2022). Available online at: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases (accessed October 15, 2022).

11. Clark A, Jit M, Warren-Gash C, Guthrie B, Wang HHX, Mercer SW, et al. Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. *Lancet Glob Health*. (2020) 8:e1003–17. doi: 10.1016/S2214-109X(20)30264-3

12. World Health Organization. *The Impact of the COVID-19 Pandemic on Noncommunicable Disease Resources and Services: Results of a Rapid Assessment*. Geneva: World Health Organization (2020).

13. Anderson G, John William Frank C, David N, Walter W, Patrick F. Using socioeconomics to counter health disparities arising from the COVID-19 pandemic *BMJ*. (2020) 369:m2149doi: 10.1136/bmj.m2149. Available online at: https://www.bmj.com/content/369/bmj.m2149 (accessed October 15, 2022).

14. Collins F. *COVID-19 Brings Health Disparities Research to the Forefront, Director's Blog*,. National Institute of Health (NIH) (2020). Available online at: https://directorsblog.nih.gov/2020/05/14/covid-19-brings-health-disparities-research-to-the-forefront/ (accessed October 15, 2022).

15. Hill CV, Pérez-Stable EJ, Anderson NA, Bernard MA. The National Institute on aging health disparities research framework. *Ethn Dis*. (2015) 25:245–54. doi: 10.18865/ed.25.3.245

16. National Institutes on Minority Health and Health Disparities Research Framework (2022). Available online at: https://www.nimhd.nih.gov/about/overview/research-framework/nimhd-framework.html (accessed October 15, 2022).

17. Social Determinants of Health. *Healthy People 2030. U.S. Department of Health and Human Services*. Office of Disease Prevention and Health Promotion (2022). Available online at: https://health.gov/healthypeople/priority-areas/social-determinants-health (accessed October 15, 2022).

18. Social Determinants of Health. *Robert Wood Johnson Foundation*. (2022). Available online at: https://www.rwjf.org/en/our-focus-areas/topics/social-determinants-of-health.html (accessed October 15, 2022).

19. Boersma P, Black LI, Ward BW. Prevalence of multiple chronic conditions among US adults 2018. *Prev Chronic Dis*. (2020) 17:E106. doi: 10.5888/pcd17.200130

20. *Fight Inequality, beat poverty. India: extreme inequality in numbers*. OXFAM International (2022). Available online at: https://www.oxfam.org/en/india-extreme-inequality-numbers (accessed October 15, 2022).

21. Adair T. Who dies where? Estimating the percentage of deaths that occur at home, *BMJ Global Health*. (2021) 6:e006766. doi: 10.1136/bmjgh-2021-006766

22. Jha P, Deshmukh Y, Tumbe C, Suraweera W, Bhowmick A, Sharma S, et al. COVID mortality in India: National Survey Data and Health Facility deaths. *Science*. (2022) 375:667–71. doi: 10.1126/science.abm5154

23. Basu JK, Adair T. Have inequalities in completeness of death registration between states in India narrowed during two decades of civil registration system strengthening? *Int J Equity Health*. (2021) 20:195. doi: 10.1186/s12939-021-01534-y

24. COVID-19 Excess Mortality Collaborators. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *Lancet*. (2022) 399:1513–36. doi: 10.1016/S0140-6736(21)02796-3

25. Gundlapalli A, Lavery A, Boehmer T, Beach M, Walke H, Sutton P, et al. Death Certificate–based ICD-10 diagnosis codes for COVID-19 Mortality surveillance

— United States, January–December 2020. *Morb Mortality Wkly Rep*. (2021) 70:523–7. doi: 10.15585/mmwr.mm7014e2

26. Stokes A, Lundberg D, Elo I, Hempstead K, Bor J, Preston H. Assessing the impact of the COVID-19 pandemic on US mortality: a county-level analysis. *medRxiv*. (2022). doi: 10.1101/2020.08.31.20184036. Available online at: https://www.medrxiv.org/content/10.1101/2020.08.31.20184036v5.article-info (accessed October 15, 2022).

27. World Health Organization. *14.9 million excess deaths associated with the COVID-19 pandemic in 2020 and 2021*. World Health Organization (2022). Available online at: https://www.who.int/news/item/05-05-2022-14.9-million-excess-deaths-were-associated-with-the-COVID-19-pandemic-in-2020-and-2021 (accessed October 15, 2022).

28. Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philos Trans R Soc Lond*. (1825) 115:513–83.

29. The World Bank. *India Data Bank*. (2019). Available online at: https://data.worldbank.org/country/india (accessed September 25, 2022).

frontiers | Frontiers in Public Health

TYPE Brief Research Report
PUBLISHED 03 November 2022
DOI 10.3389/fpubh.2022.1029196

OPEN ACCESS

EDITED BY
Shannon Monnat,
Syracuse University, United States

REVIEWED BY
Matthew Brooks,
McGill University, Canada
Tse-Chuan Yang,
University at Albany, United States

*CORRESPONDENCE
Wesley L. James
wljames1@memphis.edu

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 26 August 2022
ACCEPTED 14 October 2022
PUBLISHED 03 November 2022

CITATION
James WL, Brindley C, Purser C and Topping M (2022) Conceptualizing rurality: The impact of definitions on the rural mortality penalty. *Front. Public Health* 10:1029196. doi: 10.3389/fpubh.2022.1029196

Conceptualizing rurality: The impact of definitions on the rural mortality penalty

Wesley L. James[1]*, Claire Brindley[1], Christopher Purser[2] and Michael Topping[3]

[1]Department of Sociology, Center for Community Research and Evaluation, University of Memphis, Memphis, TN, United States, [2]Department of Politics, Justice, Law, and Philosophy, University of North Alabama, Florence, AL, United States, [3]Department of Sociology, Center for Demography and Ecology, University of Wisconsin-Madison, Madison, WI, United States

Background: In the U.S., inequality is widespread and still growing at nearly every level conceivable. This is vividly illustrated in the long-standing, well-documented inequalities in outcomes between rural and urban places in the U.S.; namely, the *rural mortality penalty* of disproportionately higher mortality rates in these areas. But what does the concept of "rural" capture and conjure? How we explain these geographic differences has spanned modes of place measurement and definitions. We employ three county-level rural-urban definitions to (1) analyze how spatially specific and robust rural disparities in mortality are and (2) identify whether mortality outcomes are dependent on different definitions.

Methods: We compare place-based all-cause mortality rates using three typologies of "rural" from the literature to assess robustness of mortality rates across these rural and urban distinctions. Results show longitudinal all-cause mortality rate trends from 1968 to 2020 for various categories of urban and rural areas. We then apply this data to rural and urban geography to analyze the similarity in the distribution of spatial clusters and outliers in mortality using spatial autocorrelation methodologies.

Results: The rural disadvantage in mortality is remarkably consistent regardless of which rural-urban classification scheme is utilized, suggesting the overall pattern of rural disadvantage is robust to any definition. Further, the spatial association between rurality and high rates of mortality is statistically significant.

Conclusion: Different definitions yielding strongly similar results suggests robustness of rurality and consequential insights for actionable policy development and implementation.

KEYWORDS

rural definitions, rural/urban, mortality, rural mortality penalty, health disparities, rural disadvantage, United States

Introduction

The relationship between health and place has been prominently featured in public health research for decades. One of the primary indicators of place in the U.S. is rurality, as demonstrated by the overwhelming evidence to a geographic penalty in health outcomes: the *rural mortality penalty* (1–3). The rural mortality penalty is historically relevant, as it represents a departure from a decades-long trend of urban disadvantage in mortality. As large swaths of the U.S. population began concentrating in large cities, morbidity and mortality rates exceeded those in rural places because of population density, substandard living conditions, spread of contagious diseases, inadequate sewage disposal, and poor water quality (4). This lasted until the mid-1900s with a period of relative equal mortality patterns due to improvements in public health infrastructure, vaccinations, physical examinations, and health education (5). Beginning in the 1980s, a dramatic reversal in place-based mortality disproportionately affecting rural places occurred, and has widened ever since.

Despite the evidence of rural disparities in health and mortality, it is not without nuanced findings from past research. In recent decades, various findings have shown that perhaps the age-structure of the population is responsible for higher rural mortality rates (6), or higher rates of age-sex-race adjusted mortality happens in urban areas (7). In more recent years, a flood of research on the rural disadvantage has emerged, suggesting that however health outcomes are measured, the rural disadvantage appears prominent. An often-overlooked feature of this body of work is the ambiguity of the term "rural" itself. The present work seeks to address this ambiguity by directly testing the impact of the three major definitions of rurality on mortality rates. We seek clarity on the extent to which the definitions of "rural" matter, first by analyzing a dichotomized definition of rural-urban as a baseline, and secondly by analyzing intra-rural definitions of varying conceptualizations of rurality. Our analysis reveals how robust place-based mortality rates are, expanding the conversation on how to address underlying disparities with policy solutions aimed at reducing the substantial gaps in rural-urban health and overall spatial inequalities.

Rural health disadvantage

Rural populations in the United States have long experienced worse health outcomes than major cities, and these patterns persist. Continuing experiential and statistical evidence of rural disadvantage is clear and overwhelming, from the opioid epidemic (8–11) to disproportionate mortality rates and life expectancy (4, 12–14), and even the disparate effects of multiple COVID-19 variants (15–18). The enduring gap in rural-urban mortality nationwide is especially concerning and is increasing each year (19). This phenomenon, known as the *rural mortality penalty* (RMP), is well documented and has identified tens of thousands of additional deaths compared to urban places (1–3). This inequality of outcomes is associated with numerous societal factors—rural populations typically have higher unemployment rates, percentages of poor and uninsured residents, and are more vulnerable than their urban counterparts to economic downturns due to more concentrated economic specialization, among many others (20–22). Despite this spatially anchored pattern of disparate outcomes, there remains a lack of clarity in what "rural" means. Rural is often conceptualized as simply "non-urban," but rural America is far from a homogenous collection of places. Though researchers have established various definitions of rural, we explore how variations in classifications matter, notably in how we understand overall inequality.

Given that current policy development and resource distribution depends heavily on institutional definitions, it is critical to understand how much our knowledge of rural disadvantage reflects reality or is an artifact of varying conceptualization and operationalization of "rural." Currently, the three major coding schemes are (1) Rural Urban Continuum Codes (RUCC), (2) Urban Influence Codes (UIC), and the National Center for Health Statistics (NCHS) Urban-Rural Classification Scheme for Counties. Using these conceptualizations, we conduct an examination of nationwide rural-urban mortality rates to determine whether varying definitions of "rural" produce the same level of rural mortality disadvantage.

Rural definitions

Prior research has shown how mortality and morbidity rates vary across rural-urban classification schemes. For example, when applying RUCCs, UICs, and NCHS codes to rural counties in Texas, one study found considerable variation in colorectal cancer incidence and mortality rates depending on the code examined (23). Other studies have shown similar varying results based on rural-urban definitions in cancer (24), access to hospitals and physicians (25), and all-cause mortality (1, 26). However, another effort based on county found little difference (27). In response to such variation, researchers have urged continued work to better understand the importance of rural and urban definitions, particularly as it pertains to health research. In fact, there has been a call for a nationwide classification study of all rural counties to further clarify how outcomes vary depending on the codes used (28). We attempt to answer this call by examining the three major rural-urban taxonomies, as the utility of these schemes may vary across regions and specific research aims.

The RUCC, UIC, and NCHS coding schemes undergird much of the social sciences, public health, and demographic literature on place in the U.S. Below we describe each of them:

- Rural-Urban Continuum Codes: These codes differentiate counties by population size and adjacency to metro areas (29). Codes 1 through 3 are urban, with population ranging from <250,000 to more than 1,000,000 people. Codes 4 through 9 indicate rural counties. The even-numbered codes (4, 6, and 8) are adjacent to metro areas, whereas the odd numbered codes (5, 7, and 9) are not adjacent to metro areas. Codes 4 and 5 have populations of over 20,000 people; codes 6 and 7 have populations ranging from 2,500 to 19,999, and codes 8 and 9 have fewer than 2,500 people.
- Urban Influence Codes: This is a twelve-code classification system of counties based on population size for metro areas, size of the largest city/town, and adjacency to metro and micro areas (30). Codes 1 and 2 are urban, stratified based on the county having more or fewer than 1,000,000 people. Codes 3 through 12 are categorized as rural and divided into two classifications: micropolitan (codes 3, 5, and 8) and noncore (codes 4, 6, 7, 9, 10, 11, and 12). The micropolitan codes are divided according to adjacency to large metro areas (code 3), small metro areas (code 5), and non-adjacency to metro areas (code 8). Likewise, the noncore areas are divided according to their adjacency to large metro (code 4), small metro (codes 6 and 7) and micro areas (codes 9, 10, 11, and 12), and those with a population of 2,500 people or fewer.
- NCHS Rural-Urban Classification Scheme for Counties: Six categories underscore urban distinctions by differentiating between central and fringe counties of large metro areas (26). The most urban category comprises the central counties of large metropolitan areas, and the most rural categories are noncore, nonmetropolitan counties. This means NCHS has more metropolitan levels (four) than micropolitan (two), largely because about 85% of the population lives in metropolitan areas (31).

While the preceding operationalization of rural areas utilizes ecological measures (e.g., population size and density) of the construct, these strategies neglect other approaches to delineating rural areas. Scholars have employed multidimensional conceptualizations of "rural," incorporating occupational and socio-cultural elements into their definitions (32). This invokes Weber's *verstehen*, as they find, among a sample of Pennsylvania residents, laypeople conceptualize rural as being comprised of socio-cultural/occupational elements, such as a fondness for agrarian lifestyles, love of the wilderness, and an active distaste for urban ideals. This is an example of the wide range of criteria used to describe, explain, and define rural. Another key detail is that the terms rural and urban are technically not the same as nonmetropolitan and metropolitan, even though they are often used interchangeably (33). In our work, for ease of interpretation, we use the terminology "rural" consistently as we refer to non-urban places.

Further, we emphasize that though the RUCC, UIC, and NCHS codes are each based on Office of Management and Budget's (OMB) delineations of metropolitan and non-metropolitan statistical areas, our analysis primarily hinges on disaggregating various levels of rurality in each scheme, rather than only the rural-urban binary common across the three. In particular, the classification schemes themselves have underlying differences beyond how they categorize different levels of urban and rural—their respective data sets show slight differences in total number of counties as well as the corresponding total populations, as well as data reliability for certain counties over the decades. This reality notwithstanding, the differing population counts are not a part of our analysis, though the overall rural-urban patterns and results are not changed in any meaningful way. We report these numbers to provide context of how many people (and %) live in rural and urban areas. In terms of the actual data provided from each classification scheme's website, we only utilize the codes themselves to attach to our county-specific mortality rates.

Methods

We use three classification schemes - RUCC, UIC, and NCHS—to assess robustness of mortality rates across these rural and urban distinctions. To address the changing classification codes that happens over time as a county increase or decreases in population, we implement a floating definition to all three. (We also tested the graphs with floating definitions of rurality against some using a fixed definition, and we conclude that our results are robust to both methods of fixed and floating definitions. The results are remarkably similar to those using the fixed definition, in terms of the overall 53-year pattern of each rural-urban designation, and the relative difference between definitions.) We assign schemes to years on a decade-by-decade basis. We then analyze a rural-urban dichotomy, combining all urban and rural subcategories together within each of the three schemes. This serves as our baseline understanding of broad temporal rural-urban patterns, and we delve further into degrees of intra-rural variation. Our primary focus is to determine if the rural disadvantage in mortality is similar regardless of the definition of rural. Second, we explore similar rural sub-categories across classification schemes to assess the magnitude of difference in mortality. This analysis of intra-rural variation highlights the level of robustness in rural disadvantage at a more precise level of measurement.

We use data from the Centers for Disease Control and Prevention (CDC) Compressed Mortality File (CMF). The CMF is a national population database that provides county-level data on U.S. mortality history. It measures all deaths by cause, age, race, sex, county of residence, and other characteristics recorded on death certificates by International Classification of Disease (ICD) codes (34). The mortality rates are measured in five-year

10.3389/fpubh.2022.1029196

TABLE 1 Number, percent, and population of counties by RUCC, NCHS, and UIC codes.

Rural code	# of counties	% of counties	Total population	% of population
RUCC				
4	214	6.9%	13,538,322	4.4%
5	89	2.9%	4,670,365	1.5%
6	593	19.1%	14,784,976	4.8%
7	425	13.7%	8,113,866	2.6%
8	219	7.0%	2,155,622	0.7%
9	408	13.1%	2,546,256	0.8%
NCHS				
5	637	20.5%	26,912,771	8.6%
6	1311	42.2%	18,854,621	6.1%
UIC				
3	130	4.2%	7,190,190	2.3%
4	149	4.8%	3,243,787	1.1%
5	242	7.8%	11,180,286	3.6%
6	344	11.1%	7,290,442	2.4%
7	161	5.2%	1,574,215	0.5%
8	265	8.5%	8,486,815	2.8%
9	184	5.9%	2,798,944	0.9%
10	187	6.0%	1,339,635	0.4%
11	115	3.7%	1,818,968	0.6%
12	171	5.5%	886,125	0.3%

averages and are age-adjusted to the 2000 Standard Million, per 100,000 people. The five-year averages provide stability for low-population rural counties. We analyze age-adjusted, all-cause mortality rates covering a period of 53 years (1968–2020), calculating, and graphing ten, five-year-averaged time periods and one three-year averaged time period, to assess rural-urban-specific trend lines by each classification scheme. Time periods are 1968–72, 1973–77, 1978–82, 1983–87, 1988–92, 1993–97, 1998–2002, 2003–07, 2008–12, 2013–2017, and 2018–2020. All-cause mortality trends are assessed by aggregated (dichotomous rural-urban classifications) and non-aggregated (intra-rural classifications) rural-urban status for the three major coding schemes discussed above (RUCC, UIC, NCHS).

We assigned RUCC, NCHS, and UIC codes to each of the 3,100+ U.S. counties and merged with 1968–2020 all-cause mortality data based on Federal Information Processing Standard (FIPS) codes. We assign coding schemes to years on a decade-by-decade basis as follows:

- NCHS: 1990 codes for 1968–1999, 2006 codes for 2000–2009, and 2013 codes for 2010–2020.
- RUCC: 1974 codes for 1968–1979, 1983 codes for 1980–1989, 1993 codes for 1990–1999, 2003 codes for 2000–2009, and 2013 codes for 2010–2020.
- UIC: 1993 codes for 1968–1999, 2003 codes for 2000–2009, and 2013 codes for 2010–2020.

The following two analyses are of: (1) the robustness in dichotomized rural-urban definitions, and (2) robustness in subcategories of intra-rural definitions. The classification of corresponding RUCC, UIC, and NCHS codes are presented in Table 1. Fifty-three-year trends in mortality within and across classification schemes are presented in the following section.

An additional analysis combines measures of association with geographically anchored spatial visualization techniques. We examined each grouping's data for spatial autocorrelation to test against the null hypothesis of spatial randomness (in which any mortality level is equally likely at any location), using global Moran's I (initial test for any patterns in mortality) and then local bivariate Moran's I for regional clusters (both rural and high mortality or both urban and low mortality) and outliers (rural and high mortality surrounded by the opposite or urban and low mortality surrounded by the opposite). Number and relation of neighbors was calculated using queen contiguity for spatial weights. The global Moran's I indicates the direction and magnitude of the spatial relationship between rurality and mortality, in the form of a coefficient between −1 and 1, along with a *p*-value. We then used bivariate local Moran's I to test significance of clusters and outliers in the relationship between rurality and mortality rates, comparing contiguous counties. This mapped the level of each county's spatial autocorrelation in terms of mortality to visualize statistically significant regional clustering of high rurality and high mortality rate, along with

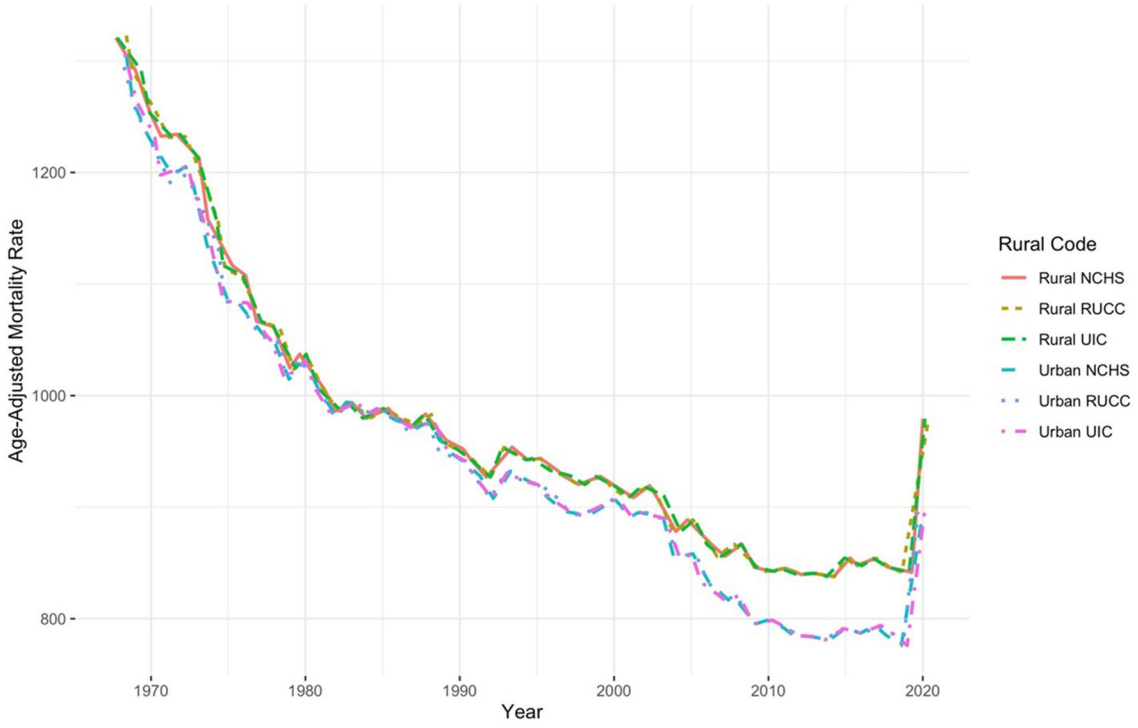


FIGURE 1
Dichotomized urban-rural comparisons of all-cause mortality, 1968–2020.

the opposite combinations thereof. The resulting maps illustrate the magnitude of association between mortality and each rural-urban coding scheme.

Results

The number, percent, and population of counties by RUCC, UIC, and NCHS codes using the downloadable data is shown in Table 1. From this, there are 3,109 contiguous counties common to all three coding schemes, with some disqualified because they did not have a code in at least one of the schemes or because their population data were unreliable. (When merging in the all-cause mortality data, the number of contiguous counties moves to 3,081 due to some of the counties' all-cause mortality rates being unreliable.) Of the 3,109 total counties by population, 1,161 (37.3%) are classified as urban across all three coding schemes. Rural counties make up the remaining 1,948 counties (62.7%). Although nearly two-thirds of counties are rural, only 15% of the American population resides there. The other 85% of the population live in urban counties (15).

An investigation into rural sub-categories provides insight into the heterogeneity that exists in rural places. For instance, RUCC 6, which are counties adjacent to metro areas with populations of 2,500 to 19,999, are the most common rural places ($N = 593$) in the United States. Nearly 15 million people reside in these counties, accounting for 4.8% of the total population. Alternatively, the stereotypical characterization of rural places, e.g., RUCC 9, remote counties with population below 2,500, are less common ($N = 408$). These remote locations are occupied by only 2.5 million people, comprising <1% of the total population. According to NCHS codes, there are roughly twice as many counties that classify as neither micropolitan nor metropolitan (NCHS 6) than there are micropolitan counties (NCHS 5). However, there are 8 million fewer people in NCHS 6 compared to NCHS 5. With UIC codes, the modal number of counties classified as rural are those with the code 6 classification (noncore areas adjacent to metro areas, with populations of at least 2,500). UIC 6 contains 7.3 million people and only 2.4% of the total population. These figures demonstrate a few examples of the variation in rural conceptualization which may affect the results of spatially oriented statistical analyses.

Figure 1 graphs comparisons of all-cause mortality between dichotomized urban and rural areas for the entire U.S. population from 1968 to 2020, for RUCC, UIC, and NCHS codes. The general trendlines are remarkably similar. All three urban definitions track precisely with one another throughout the time series, as do the three rural definitions. After nearly 20 years of no discernable rural-urban disparity, the rural

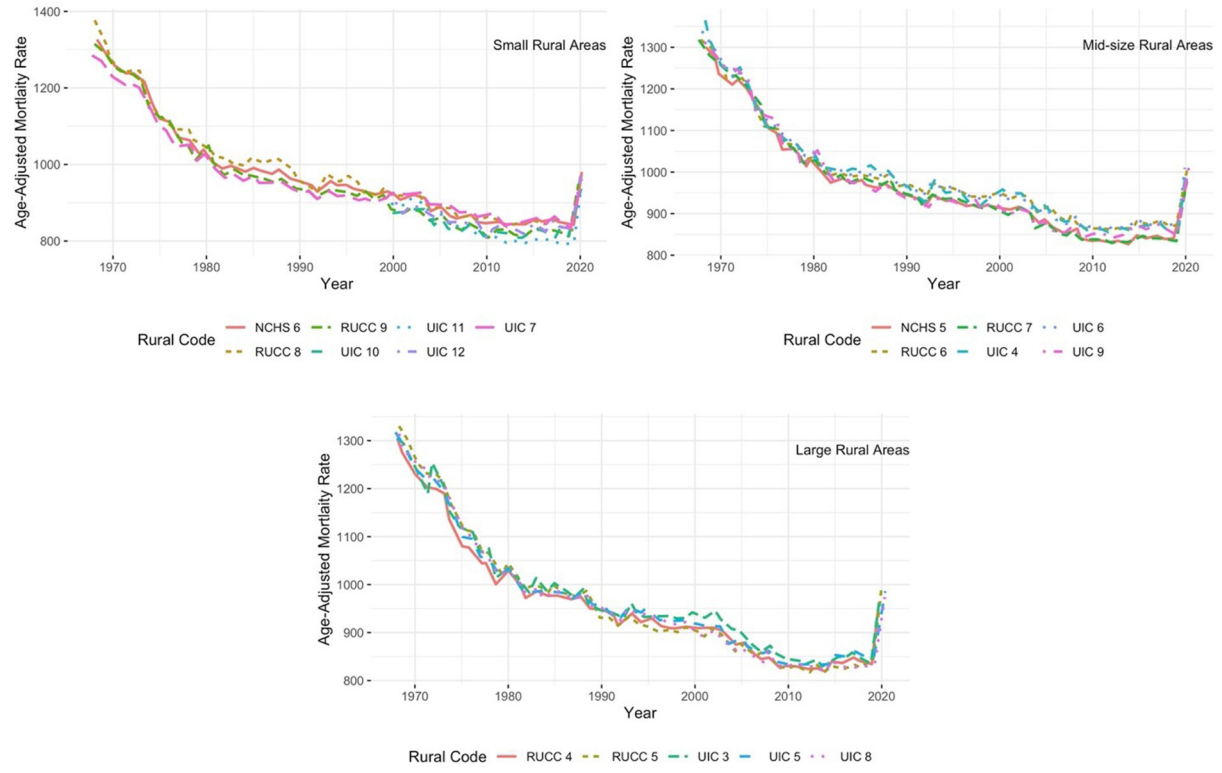


FIGURE 2
Intra-rural comparisons of all-cause mortality, 1968–2020.

disadvantage in mortality emerges in the mid-1980s and continues to grow through 2020, regardless of which rural-urban classification scheme is utilized, suggesting the overall pattern of rural disadvantage is robust to any definition. Pre-COVID, rural places exhibited a mortality rate of roughly 840 deaths per 100,000 compared to the urban rate of about 776 per 100,000. Another key observation consistent across classification schemes is the spike in mortality in rural and urban places in the last 2 years. This is largely the influence of COVID-19 substantially increasing mortality rates throughout the nation (35).

Figure 2 focuses exclusively on rural subcategories across RUCC, UIC, and NCHS schemes for small, mid-size, and large areas. The categories for each classification scheme are as follows: (1) small rural areas: RUCC codes 8 and 9, UIC codes 7, 10, 11, and 12; NCHS code 6; (2) midsize rural areas: RUCC codes 6 and 7, UIC codes 4, 6, and 9; NCHS code 5; and (3) large rural areas: RUCC codes 4 and 5 and UIC codes 3, 5, and 8. Overall, the definitions follow a similar trend through the decades. In small rural areas, NCHS 6 counties (classified as noncore, nonmetropolitan counties) reliably exhibit higher levels of mortality (about 987 per 100,000 in 2020) than any other category but track very closely with UIC 12 (970 per 100,000) counties (noncore not adjacent to metro- or micropolitan area) and UIC 7 (961 per 100,000) (noncore adjacent to small metro). The other four categories (RUCC 8-9 and UIC 10-11) show slightly lower levels of mortality than their other rural counterparts. On average, the combination of small rural areas exhibits a collective mortality rate just shy of 956 deaths per 100,000. For mid-size, UIC 6 and 9 (about 1,010 per 100,000 in 2020) and RUCC 6 (1,009 per 100,000) show consistently higher mortality levels than RUCC 7, NCHS 5, and UIC 4. The collective mortality rate of mid-sized rural places is near 996 deaths per 100,000. And in large rural areas, mortality trends are the most similar across time; however, RUCC 5 (about 990 per 100,000 in 2020) and UIC 3 (948 per 100,000) are generally the highest and lowest, respectively. The collective mortality rate for these places is between their small and mid-size counterparts, hovering just above 974 deaths per 100,000.

Finally, analysis of spatial clusters and outliers in rural-urban mortality rates show statistically significant concentrations of high mortality across the rural classification schemes, especially in the Southeast and Appalachia. An initial global Moran's I test for concentrations of high and low age-adjusted mortality rates shows positive autocorrelation—the mortality data does spatially cluster ($I \approx 0.5$). Then we add rurality: looking at local Moran's I for a bivariate model (rurality and mortality) of spatial clusters and outliers. Figure 3 shows a statistically significant

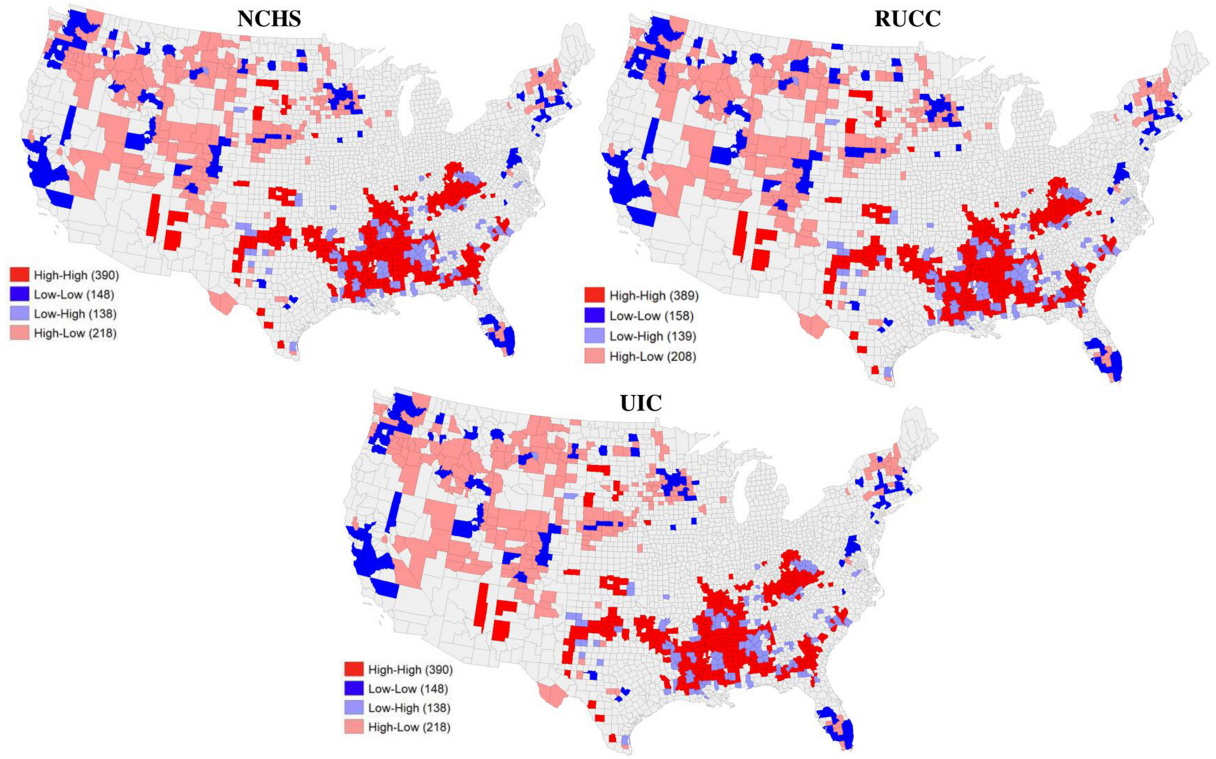


FIGURE 3
Spatial clusters of rurality and high age-adjusted mortality rates (all codes).

pattern to the clustering of mortality rates with rurality ($p < 0.05$). Bright red areas indicate county clusters that are both rural and are surrounded by counties with high mortality—with two-and-a-half times as many counties as the next cluster (bright blue areas: urban county areas surrounded by low death rates; gray indicates lack of significance). Being rural had a positive significant spatial correlation with high mortality rates, especially in Appalachia and the Southeast, which is consistent with the results of mortality alone. The rural spatial outliers (light red) reveal more counties with high mortality surrounded by urban spaces. Light-blue counties have low mortality but are surrounded by rural areas. Overall, Between the three classification schemes, similar numbers of counties are both rural and nested among high age-adjusted mortality rates.

Discussion

The association between rurality and mortality is consistent across the three major classification schemes. This is not to argue that how "rural" is defined does not matter unilaterally—no single definition could ever capture the diversity of rural spaces and the combined interdependence and heterogeneity between places. But when examining rural-urban all-cause mortality disparities, the RMP persists across classification schemes, not only when conceptualized as dichotomized rural or urban, but also to a very large degree with smaller rural sub-classifications. To this point, researchers have argued that dichotomized rural-urban definitions mask variability within rural areas (28). However, our relatively unique finding is that the standard rural subcategories, e.g., intra-rural classifications, offer a large degree of uniformity in mortality outcomes.

The conceptualizations of rural-urban are secondary to rural disadvantage, which is robust across definitions. This fits social theorist Emile Durkheim's "social facts"—these definitions transcend the individual while constraining the individual (36). This reflects the persistence of place-based mortality and social determinants, spatial inequalities of access and outcomes, and a continued basis for further disaggregation of data and analysis of specific forces affecting different groups of rural residents.

Social determinants of mortality—poverty, education, income, race, etc.—are more concentrated in rural areas. As rural communities are left further behind with each passing generation, the level of relative deprivation compared to urban places continues to grow (37, 38). This stark inequality is the culmination of the many disparities apparent in American life. That the RMP is significant and still growing illuminates the gap between health improvements in urban and rural places.

Mortality in rural places is indeed improving, but at a much slower rate than their urban counterparts (39), once again leaving rural Americans further behind. And as the COVID-19 pandemic has vividly shown, inequity in health is a collective, not individual, problem. Focusing on why these disparities exist and persist is a more powerful tool for solving them than centering the rural-urban definitional divide itself. Even where myriad inter-area variations may appear, the rural disadvantage still begets higher mortality.

Inequality is a destabilizing force (40), and grips the U.S. map from coast to coast. Discussion of its alleviation is incomplete without mention of space and the distribution of resources across places. Local, state, and national policymakers can use more targeted analysis to address imbalances in investment and facilitate equitable environments. Solutions are likely multi-faceted; for instance, policy change can address the scope of practice for nurse practitioners to enhance their ability to practice medicine independently, especially in places with a shortage of family doctors and specialists (41). State governments can attack problems through investment in struggling areas, a recent example being how ARPA funds are distributed by state and local governments (42). Lastly, local-level change can happen through program intervention aimed at addressing chronic disease in struggling areas through the implementation of telemedicine or other programs (43). Some of these policies and programs could be based on high mortality rates by place, rather than rural-urban classifications. Investing in both people and places better echoes the multidimensionality of this country's geography.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found in the following repositories: United States Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), Compressed Mortality File (CMF) on CDC WONDER Online Database (1968–78, 1979–1998, 1999–2020); 2020 U.S. Census Bureau, Cartographic Boundary Files.

Author contributions

WJ is primarily responsible for the research idea, conceptualizing the data analysis strategy, and drafting the manuscript. CP conducted the literature review. CB contributed to the data analysis and constructed the graphs, table, and maps. MT conducted initial data analysis and constructed a set of graphs. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Cosby AG, Neaves TT, Cossman RE, Cossman JS. James, WL Feierabend N, et al. Preliminary evidence for an emerging nonmetropolitan mortality penalty in the United States. *Am J Public Health.* (2008) 98:1470–2. doi: 10.2105/AJPH.2007.123778

2. Cossman JS, James WL, Cosby AG, Cossman RE. Underlying causes of the emerging non-metropolitan mortality penalty. *Am J Public Health.* (2010) 100:1417–9. doi: 10.2105/AJPH.2009.174185

3. James WL. All rural places are not created equal: revisiting the rural mortality penalty in the United States. *Am J Public Health.* (2014) 104:2122–9. doi: 10.2105/AJPH.2014.301989

4. Cosby AG, McDoom-Echebiri MM, James WL, Khandekar H, Brown W, Hanna HL. Growth and persistence of place-based mortality in the United States: the rural mortality penalty. *Am J Public Health.* (2019) 9:155–62. doi: 10.2105/AJPH.2018.304787

5. Haines MR. The urban mortality transition in the United States. *Ann Demogr Hist (Paris).* (2001) 101:33–64. doi: 10.3917/adh.101.0033

6. Miller MK, Stokes CS, Clifford WB. A comparison of the rural-urban mortality differential for deaths from all causes, cardiovascular disease and cancer. *J Rural Health.* (1987) 3:23–34. doi: 10.1111/j.1748-0361.1987.tb00165.x

7. McLaughlin DK. Stokes CS, Nonoyama A. Residence and income inequality: Effects on mortality among US counties. *Rural Sociol.* (2001) 66:579–98. doi: 10.1111/j.1549-0831.2001.tb00085.x

8. Singh GK, Siahpush M. Widening rural-urban disparities in all-cause mortality and mortality from major causes of death in the U.S.A., 1969-2009. *J Urban Health.* (2014) 91:272–92. doi: 10.1007/s11524-013-9847-2

9. Monnat SM. Trends in U.S. working-age non-Hispanic white mortality: Rural-urban and within-rural differences. *Popul Res Policy Rev.* (2020) 39:805–34. doi: 10.1007/s11113-020-09607-6

10. Kiang MV, Basu S, Chen J, Alexander MJ. Assessment of changes in the geographical distribution of opioid-related mortality across the United States by opioid type, 1999-2016. *JAMA Netw Open.* (2019) 2:e190040. doi: 10.1001/jamanetworkopen.2019.0040

11. Peters D J, Monnat SM, Hochstetler AL, and Berg MT. The opioid hydra: understanding overdose mortality epidemics and syndemics across the rural-urban continuum. *Rural Sociol.* (2020) 85:589–622. doi: 10.1111/ruso.12307

12. Chetty R, Stepner M, Abraham S, Lin S, Scuderi B, et al. The association between income and life expectancy in the United States, 2001-2004. *JAMA*. (2016) 315:1750–66. doi: 10.1001/jama.2016.4226

13. Garcia MC, Faul M, Massetti G, Thomas CC, Hong Y, Bauer UE, et al. Reducing potentially excess deaths from the five leading causes of death in the rural United States. *MMWR Surveill Summ.* (2017) 66:1–7. doi: 10.15585/mmwr.ss6602a1

14. Leider JP, Meit M, McCullough JM, Resnick B, Dekker D, Alfonso YN, et al. The state of rural public health: Enduring needs in a new decade. *Am J Public Health*. (2020) 110:1283–90. doi: 10.2105/AJPH.2020.305728

15. Dobis EA, Krumel Jr TP, Cromartie J, Conley KL, Sanders A, and Ortiz R. *Rural America at a Glance*. Washington, D.C.: U.S. Department of Agriculture, Economic Research Service (2021). p. 18.

16. Karim SA, Chen HF. Deaths from COVID-19 in rural, micropolitan, and metropolitan areas: A county-level comparison. *J Rural Health.* (2021) 37:124–32. doi: 10.1111/jrh.12533

17. Karmakar M, Lantz PM, Tipirneni R. Association of social and demographic factors with COVID-19 incidence and death rates in the US. *JAMA Netw Open.* (2021) 4:e2036462. doi: 10.1001/jamanetworkopen.2020.36462

18. Mueller JT, McConnell K, Burow PB, Pofahl K, Merdjanoff AA, and Farrell J. Impacts of the COVID-19 pandemic on rural America. *Proc Natl Acad Sci.* (2022) 118:2019378118. doi: 10.1073/pnas.2019378118

19. Hoffman A, Holmes M. *Regional Differences in Rural and Urban Mortality Trends*. US: The University of North Carolina at Chapel Hill, 2017. Available online at: https://www.shepscenter.unc.edu/wp-content/uploads/dlm_uploads/2017/08/Regional-Differences-in-Urban-and-Rural-Mortality-Trends.pdf (accessed on August 12, 2022).

20. Hassenger EW, Hobbs DJ. Rural society—the environment of rural health care. In: Straub LA, Walzer N (eds). *Rural Health Care: Innovation in a Changing Environment*. New York: McGraw-Hill (1992). p. 178–90.

21. Hart GL, Larson EH, Lishner DM. Rural definitions for health policy and research. *Am J Public Health.* (2005) 95:1149–55. doi: 10.2105/AJPH.2004.042432

22. Hines F, Brown D and Zimmer J. *The Social and Economic Characteristics of the Population in Metro and Nonmetro Counties, 1970*. Washington, D.C.: U.S. Department of Agriculture, Economic Research Service (1975).

23. Hall SA., Kaufman JS, and Ricketts TC. Defining urban and rural areas in U.S. epidemiological studies. *J Urban Health*. (2006) 83:162–75. doi: 10.1007/s11524-005-9016-3

24. Hawley ST, Change S, Risser D, and Zhang, Q. Colorectal cancer incidence and mortality in Texas 1990-1992: A comparison of rural classifications. *J Rural Health*. (2002) 18:536–46. doi: 10.1111/j.1748-0361.2002.tb00920.x

25. Doogan NJ, Roberts ME, Wewers ME, Tanenbaum ER, Mumford EA, Stillman FA. Validation of a new continuous geographic isolation scale: A tool for rural health disparities research. *Soc Sci Med.* (2018) 215:123–32. doi: 10.1016/j.socscimed.2018.09.005

26. Zhao J, Ameratunga S, Lee A, Browne M, and Exeter DJ. Developing a new index of rurality for exploring variations in health outcomes in Auckland and Northland. *Soc Indic Res.* (2019) 144:955–80. doi: 10.1007/s11205-019-02076-1

27. Hackbarth GM. *Report to the Congress: Medicare in Rural America. Medicare Payment Advisory Commission*. (2001). Available online at: http://67.59.137.244/documents/Jun01%20Entire%20report.pdf (accessed on August 12, 2022).

28. Cromartie J, Parker T. *What is rural? U.S. Department of Agriculture Economic Research Service*. (2019). Available online at: https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/what-is-rural.aspx (accessed on August 12, 2022).

29. Cromartie J. *Rural-Urban Continuum Codes. U.S. Department of Agriculture Economic Research Service*. (2020). Available online at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx (accessed on August 12, 2022).

30. Cromartie J. *Urban Influence Codes. U.S. Department of Agriculture Economic Research Service*. (2019). Available online at: https://www.ers.usda.gov/data-products/urban-influence-codes/ (accessed on August 12, 2022).

31. Ingram DD, Franco SJ. *2013 NCHS Urban-Rural Classification Scheme for Counties*. Washington, D.C.: U.S. Department of Health and Human Services, Office of Analysis and Epidemiology (2014). 81 p.

32. Willits FK, Bealer RC, Timbers VL. Popular images of "rurality": Data from a Pennsylvania survey. *Rural Sociol.* (1990) 55:559–78. doi: 10.1111/j.1549-0831.1990.tb00697.x

33. Ratcliffe M, Burd C, Holder K, Fields A. Defining rural at the US Census Bureau. *Am Commun Surv.* (2016) 1:1–8.

34. *Compressed Mortality File (CMF) on CDC WONDER Online Database (1968-2020)*. Washington, D.C.US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics (2021). Available online at: https://wonder.cdc.gov/cmf-icd10.html (accessed August 17, 2022).

35. Sabo S, Johnson S. *Pandemic disrupted historical mortality patterns, caused largest jump in deaths in 100 years*,Suitland-Silver Hill, MD: United States Census Bureau (2022). Available online at: https://www.census.gov/library/stories/2022/03/united-states-deaths-spiked-as-covid-19-continued.html (accessed on August 21, 2022).

36. Durkheim E. What is a social fact? In: Lukes S ed. *The Rules of Sociological Method*. New York: Simon and Schuster (1982). p. 50–9.

37. Abrams LR, Myrskylä M, Mehta NK. The growing rural–urban divide in US life expectancy: contribution of cardiovascular disease and other major causes of death. *Int J Epidemiol.* (2021) 50:1970–197. doi: 10.1093/ije/dyab158

38. Richman L, Pearson J, Beasley C, Stanifer J. Addressing health inequalities in diverse, rural communities: An unmet need. *SSM Popul Health.* (2019) 9:100398. doi: 10.1016/j.ssmph.2019.100398

39. James W, J Cossman, and J Wolf. Persistence of death in the United States: The remarkably different mortality patterns between America's Heartland and Dixieland. *Demogr Res.* (2018) 39:897–910. doi: 10.4054/DemRes.2018.39.33

40. Birdsall N. Why inequality matters: Some economic issues. *Ethics Int Affairs.* (2001) 15:3–28. doi: 10.1111/j.1747-7093.2001.tb00356.x

41. Nielsen M, D'Agostino D, Gregory P. Addressing rural health challenges head on. *Mo Med*. (2017) 114:363–6.

42. Berube A. *How Local Governments are Combating Economic Disadvantage with American Rescue Plan Funds*. Washington, D.C.: Brookings Institute. (2022). Available online at: https://www.brookings.edu/blog/the-avenue/2022/08/23/how-local-governments-are-combatting-economic-disadvantage-with-american-rescue-plan-funds/ (accessed on August 12, 2022).

43. Butzner M, Cuffee Y. Telehealth interventions and outcomes across rural communities in the United States: Narrative review. *J Med Internet Res.* (2021) 23:e29575. doi: 10.2196/29575

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 03 November 2022
DOI 10.3389/fpubh.2022.995585

OPEN ACCESS

EDITED BY
Irma Elo,
University of Pennsylvania,
United States

REVIEWED BY
Claudia Holzman,
Michigan State University,
United States
Florencia Torche,
Stanford University, United States

*CORRESPONDENCE
Marielle Côté-Gendreau
marielle.cote-gendreau@princeton.edu
Katie Donnelly Moran
kfmoran@princeton.edu

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 16 July 2022
ACCEPTED 10 October 2022
PUBLISHED 03 November 2022

CITATION
Côté-Gendreau M and Donnelly Moran K (2022) Geographic heterogeneity in Black-white infant mortality disparities. *Front. Public Health* 10:995585. doi: 10.3389/fpubh.2022.995585

Geographic heterogeneity in Black-white infant mortality disparities

Marielle Côté-Gendreau[1*†] and Katie Donnelly Moran[1,2*†]

[1]Office of Population Research, Princeton University, Princeton, NJ, United States, [2]Department of Sociology, Princeton University, Princeton, NJ, United States

Despite recent decreases in Black infant mortality, racial disparities persist, motivating continued research into factors related to these inequalities. While the inverse association between education and infant mortality has been documented across races, less is known about its geographic heterogeneity. Using vital statistics from the National Center for Health Statistics, this study considers Black-white disparities in infant mortality for births occurring between 2011 and 2015 across regions and metropolitan status of maternal residence. With logistic regressions, we investigate heterogeneity in maternal educational gradients of infant mortality by geographic residence both within and between races. Beyond confirming the well-known relationship between education and infant mortality, our findings document a slight metropolitan advantage for infants born to white mothers as well as lower returns to education for infants born to Black mothers residing in nonmetropolitan counties. We observe a metropolitan advantage for infants born to Black mothers with at least a bachelor's degree, but a metropolitan disadvantage for infants born to Black mothers with less than a high school degree. The South is driving this divergence, pointing to particular mechanisms limiting returns to education for Southern Black mothers in nonmetropolitan areas. This paper's geographic perspective emphasizes that racial infant health disparities are not uniform across the country and cannot be fully understood through individual and household characteristics.

KEYWORDS

infant mortality, geography, metropolitan, region, race, maternal education

1. Introduction

Infant mortality is a key indicator of population health as it reflects the mother's health, environmental context, and access to socioeconomic resources and healthcare (1). Within countries, infant mortality gives important insight into the health of population subgroups and brings to light long-lasting inequalities (2). Infant mortality in the United States far exceeds infant mortality in peer European countries, with geographic heterogeneity and racial disparities contributing to these higher rates (3, 4). Black-white disparities have persisted over time, and in 2015, the infant mortality rate for Black infants was more than double that of white infants (4). This racial gap is also present between mothers of similar socioeconomic status (5–7), which points to the importance of examining

health disparities along with related dimensions such as education, economic circumstances, geography, and nativity status.

The heterogeneity of disparities in infant mortality across the country remain not well-understood. In this paper, we look at geographic patterns in infant mortality in the United States through their intersection with racial and educational inequalities. We bring together socioeconomic status, race, geography, and timing of death, which have been identified by previous papers as essential dimensions for understanding infant mortality in the United States. Using the National Center for Health Statistics' complete linked birth/infant death datasets from 2011 to 2015 and 1998 to 2002, we investigate the role of region and metropolitan status on infant mortality for children born to non-Hispanic white and non-Hispanic Black mothers across categories of educational attainment. Spatial approaches to health research have important implications for public policy through informing appropriate targeting of resources and intervention strategies. Additionally, better understanding the intricacies behind Black-white disparities provides insight into the systemic inequalities that Black communities face and their broad impacts on health more generally.

Below we briefly outline related literature before discussing our data and methods. We then present our findings and conclude with a discussion of our results and their significance in the broader context of racial and geographic disparities in health.

2. Literature review

In the infant mortality literature, socioeconomic status is frequently proxied with maternal education. Education is expected to lower the probability of infant mortality through two main pathways. On one hand, women are expected to leverage higher education into improved social and economic conditions, such as better housing, financial stability, quality healthcare, and adequate nutrition. On the other hand, education—through increased knowledge—has an effect on the adoption of individual behaviors that impact health, such as smoking, exercising, assigning importance to nutrition, and seeking early prenatal care (8).

Many have documented group disparities in both infant and adult health and mortality along educational gradients. With data from 2007 to 2010, Fishman et al. (7) found that the Black-white gap in infant mortality cannot be accounted for by differential educational attainment, with college-educated Black mothers experiencing higher infant mortality than white mothers with at most a high school degree. Rather, they identified gestational length as a meaningful factor for explaining the racial gap, suggesting that educational attainment does not reflect comparable life experiences for Black and white mothers. Moreover, Green and Hamilton (6) estimated infant mortality by race/ethnicity and nativity in the United States across educational attainment categories between 1998 and 2002. They observed higher mortality of infants born to Black mothers and lower relative returns to education compared to those born to white mothers. Within each ethnoracial group, they also emphasized the lower levels of infant mortality and flatter educational gradients of immigrant mothers compared to their U.S.-born counterparts, reflecting immigrants' well-documented health advantage (9, 10), part of which gets transferred to their descendants (11–14). Explanations emphasize the positive selection of immigrants with respect to health and protective cultural habits. Yet, neither Green and Hamilton (6) nor Fisherman et al. (7) investigated how returns to education with respect to infant mortality vary across the country.

The important role of geography in infant mortality in the United States has been described in numerous studies. With regard to region, the West and Northeast have historically had lower probabilities of infant mortality than the South and Midwest (3, 6, 15). However, little research has addressed how geographic differences manifest across levels of education. Although Montez and Berkman (16) observed similar gradients in adult mortality across races and census regions and trends in these gradients over time, it is unclear whether these findings would hold in the context of infant mortality and across levels of rurality. In recent years, Rossen et al. (17) reported variations in county-level racial disparities in infant mortality, with the Great Lakes, mid-Atlantic, and parts of Florida having the largest Black-white gap. Additionally, Sparks et al. (18), Yao et al. (19), and Luo and Wilkins (20) emphasized the persistent disadvantage of rural infants in their first year of life across multiple contexts. This disadvantage is linked to lower socioeconomic status and limited access to healthcare services and resources in rural areas. However, after controlling for availability of physicians and neonatal care as well as socioeconomic and other local conditions at the county level, Sparks et al. (18) found that rural counties generally have lower neonatal mortality (mortality within the first 4 weeks) than counties located in large metropolitan areas. Yet, the rural disadvantage persisted with regard to postneonatal mortality (mortality in the remainder of the first year of life).

Consistent with Sparks et al. (18), multiple studies have found that mortality in the neonatal and postneonatal periods is associated with distinct causes and has different associations with maternal and neighborhood socioeconomic characteristics. Notably, neonatal mortality tends to have a weaker association with maternal education than postneonatal mortality (6). Neonatal deaths are most likely to derive from pregnancy- and delivery-related factors—such as congenital malformations, prematurity, very low birth weight, and delivery complications (21, 22)—as well as issues related to access to and quality of neonatal care. Low birth weight and short gestational length have been identified as the most important predictors of neonatal mortality: in a study of California-born infants in 1995–1997,

the Black-white gap in neonatal mortality was entirely explained by Black mothers' higher rates of low birth weight and preterm birth (23). Mortality in the postneonatal period, on the other hand, reflects the continuing effects of pregnancy-related complications as well as the role of environment- and household-related factors, with congenital abnormalities, accidents, and sudden infant death syndrome ranking as the leading causes of death (21, 22).

The timing of infant death across neonatal and postneonatal periods has important implications for understanding inequalities in infant mortality. Chen et al. (3) documented that the American disadvantage in infant mortality compared to peer European countries in the period 2000–2005 is driven by postneonatal mortality. Within the United States, regional differences are also primarily explained by differences in postneonatal mortality rather than neonatal mortality. Thus, Chen et al. (3) concluded that the United States' higher infant mortality is due to a steeper socioeconomic gradient as well as large regional differences. These findings speak to the need to better document how the United States' heterogeneity manifests in mortality, and the importance of decomposing infant mortality into its neonatal and postneonatal components.

In light of this review of the literature, this study brings together socioeconomic status, race, geography, and timing of death to further our understanding of the disparities in infant mortality across the United States. In particular, we examine the differences in educational gradients in infant mortality for native-born non-Hispanic Black and white mothers across metropolitan residence and region. First, in line with existing literature, we aim to confirm the negative association between education and infant mortality. Then, the following research questions guide our main analysis:

1. *Heterogeneity within race.* For both Black and white mothers, do returns to education in infant mortality vary across metropolitan status and region?
2. *Heterogeneity in patterns across races.* Do the geographic patterns in returns to education in infant mortality differ between Black and white mothers?

In order to consider the distinct causes of infant mortality throughout the first year of life, we examine the above questions across both the neonatal and postneonatal periods.

Identifying geographic patterns in infant mortality and determining whether Black and white populations share similar patterns is important for two main reasons: first, it informs appropriate allocation of resources; second, it assesses whether geographic heterogeneity contributes to the overall racial gap in infant mortality. We expect significant overlap between our results and the robust body of literature on infant mortality in the U.S., which has documented the persistence of racial disparities and educational gradients. Beyond these well-known characteristics of U.S. infant mortality, this study seeks to shed light on the understudied intersection of race, education, and geography. Our approach of estimating mortality across educational categories is in line with recent research drawing attention to the role of compositional differences in education in time-space comparisons of mortality (24, 25).

3. Data and methods

3.1. Data

This paper uses data from the National Vital Statistics Birth Cohort Linked Birth/Infant Death Data (LBID) from the National Center for Health Statistics (NCHS) (26). This dataset contains almost all infants born in the United States during a given year linked with information from their death record if death occurred within the first 365 days. Coverage of births is quasi-exhaustive, and we use the restricted data in order to have complete data on mothers' county of residence. We first look at the years 2011–2015, which correspond to the five most recent years for which these data are available. However, the coverage of our data is incomplete due to the adoption of a revised birth certificate form in 2003. While some states immediately switched to the new form, it was not until 2016 that all states were using it. In 2011, 14 states were not using the revised form, and in 2015, two states had still not made the change.

The revised birth certificate form affected the recording of maternal educational attainment and race. The revised birth certificate issued in 2003 reports educational attainment in terms of highest educational level completed, rather than years of schooling which was collected by the 1989 birth certificate format. With regard to race, the revised 2003 birth certificate allows for the reporting of multiple racial identities, whereas the 1989 birth certificate form only allowed one race to be recorded for each parent. With the revised form, the NCHS created a *bridged* race variable transforming multiple race responses into one single race, allowing for continuity with the older records. We use the NCHS bridged race in this analysis. In both time periods, we use NCHS's imputation when racial information was missing from the birth certificate. Additionally, in 2011, NCHS began only releasing maternal education and race data for births that were recorded with the revised birth certificate form. Therefore, in the 2011–2015 time period, we only observe these maternal characteristics for mothers who gave birth in states that had adopted the 2003 revised birth certificate.

Thus, while the LBID has quasi-exhaustive coverage, we are limited to analyzing births that were recorded with the revised form. Because of our incomplete coverage of states in the period 2011–2015 and our focus on geography, we reproduce our analysis in two ways. First, we turn to the years 1998–2002, which are the 5 years that immediately precede the adoption of the revised birth certificate. Although the main purpose of this robustness check is to examine consistency and continuity in geographic patterns, this also allows us to consider long run

trends in levels and compare our findings to previous work, notably Green and Hamilton (6). Next, restricting to states that had adopted the revised birth certificate before 2011, we examine whether patterns are stable across the full and restricted set of states in both time periods.

Our main geographic variable of interest is maternal county of residence's metropolitan status. This is a binary variable defined as metropolitan or nonmetropolitan according to whether or not the county was located in a metropolitan statistical area (MSA) as of 2005, i.e., an urban cluster with a population of at least 50,000. More precisely, we use the 2006 NCHS Urban-Rural Classification Scheme for Counties, which places U.S. counties into six categories: large central metropolitan, large fringe metropolitan, medium metropolitan, small metropolitan, micropolitan, and non-core. Large central metropolitan and large fringe metropolitan counties form the core and peripheral counties of MSAs of at least 1,000,000 people. All counties in MSAs with populations in the intervals 250,000–999,999 and 50,000–249,999 are categorized, respectively, as medium and small metropolitan counties. Micropolitan counties belong to urban clusters with population of less than 50,000 and, along with non-core—or rural—counties, form the nonmetropolitan counties in our analysis. We use the 2006 classification, because it is a midpoint between the two time periods under study.

Across our analysis, we restrict the dataset to singletons born to U.S.-born mothers between the ages of 18 and 46 living in the United States at the time of giving birth and who reported a racial/ethnic identity of either non-Hispanic white or non-Hispanic Black[1] and have non-missing information on maternal education and the covariates of interest. Green and Hamilton (6) show that foreign-born mothers have lower levels of infant mortality and display a weaker association between infant mortality and educational attainment. Interpreting these differences in educational gradients is challenging, because it is unknown from the data how long foreign-born mothers have lived in the U.S. and where they were educated. The inclusion of foreign-born mothers in our analysis posed challenges to the interpretation of our findings. For these reasons, we restrict our analysis to U.S.-born mothers. In the two time periods, we pool births from the 5 years.

3.2. Methods

3.2.1. Main analysis

To examine the relationship between infant mortality, education, and metropolitan county residence, we run separate logistic regressions on infant death for infants born to non-Hispanic white and non-Hispanic Black mothers between 2011 and 2015:

$$Logit(mortality_i|C_i, E_i, X_i) = \beta_0 + \beta_1 C_i + \beta_2 E_i + \beta_3 C_i E_i + \mathbf{X_i'}\boldsymbol{\alpha} \quad (1)$$

where $mortality_i$ is infant mortality, C_i is county of residence's metropolitan status, E_i is mother's educational attainment, C_iE_i is the interaction between metropolitan residence and educational attainment, and $\mathbf{X_i'}$ is the vector of controls. We follow Green and Hamilton's (6) model specifications and control for mother's age, age squared, marital status, first trimester prenatal care, child's sex, birth order, birth year, and U.S. region of birth (Northeast, Midwest, South, and West). We also produce robust standard errors clustered by U.S. census region to reflect that regional characteristics likely cause observations in the same region to be correlated, and we present 95% confidence intervals.[2] For each race, we run both a baseline model and a fully specified model. In the baseline model (Model 1), we only include maternal age and maternal age squared in the vector of controls $\mathbf{X_i'}$, recognizing that maternal age is a main driver of infant mortality and has a non-linear relationship with infant death (28). In the fully specified model (Model 2), we include the full set of controls as listed above. To consider the different underlying causes of infant mortality, we also conduct the above analysis separately for neonatal and postneonatal mortality where $mortality_i$ is infant death in days 0–27 and in days 28–364, respectively.

Then, to address the question of how the interaction between race and education varies over geography, we conduct this analysis broken out by U.S. region. The model specification is

1 Analysis was also conducted on those born to non-Hispanic Asian (including Pacific Islanders) or Hispanic (regardless of race) mothers. Results are available upon request. Although we have looked at four ethnoracial groups (non-Hispanic white, non-Hispanic Black, non-Hispanic Asian, and Hispanic), we chose to focus this paper on infants born to white and Black mothers, acknowledging the continued importance of Black-white disparities in the U.S. context. Moreover, large proportions of Hispanic and Asian mothers are first-generation immigrants who were dropped from this analysis, and a considerable number of U.S.-born individuals from these communities are second- or third-generation immigrants. In light of the health selection of immigrants (27) and evidence suggesting partial intergenerational transmission of the healthy immigrant advantage (11–14), the educational gradient in infant mortality for Hispanic and Asian mothers can be assumed to derive from distinct mechanisms than for white and Black mothers, who largely belong to communities that have been in the United States for multiple generations.

2 The code for this paper's analysis was adapted from Green and Hamilton (6), who made replication material available on the website of *Demographic Research*.

as follows:

$$Logit(mortality_i|C_i, R_i, E_i, X_i) = \beta_0 + \beta_1 C_i + \beta_2 R_i + \beta_3 E_i + \beta_4 C_i R_i + \beta_5 C_i E_i + \beta_6 R_i E_i + \beta_7 C_i R_i E_i + \mathbf{X_i'}\boldsymbol{\alpha} \quad (2)$$

where variables are as defined above with the inclusion of the complete interaction between county of residence's metropolitan status, maternal educational attainment, and U.S. region of birth (Northeast, Midwest, South, and West).

In figures, we present the total predicted probabilities of infant, neonatal, and postneonatal mortality. These predicted probabilities are computed by averaging over predicted probabilities for each mother in the dataset, using the regression coefficients. Thus, the estimates shown in the figures below do not eliminate differences in distributions of covariates between subgroups and, rather, represent the average predicted probability of infant mortality for mothers in these subgroups.

When logistic regression is used to model rare events, there are potential concerns related to the low number of events observed in the data (29). Past research has suggested that thresholds as low as 10 (30) or even 5 (31) events per variable included in the model produce valid estimates. Because our coverage of infant mortality in the LBID is quasi-exhaustive, we observe 7,080 instances of mortality (323 per predictor) in our most restrictive model with the largest number of coefficients (Black mothers, postneonatal mortality, Equation 1, Model 2). The bias in logistic regression models fitted by maximum likelihood has been found to be minimal with much smaller sample sizes and fewer events per variable than in the LBID (32–36). Therefore, while infant mortality is a rare event, it is not rare enough in our data to bias our logistic regression estimates.

3.2.2. Supplemental analyses

We conduct two supplemental analyses. First, to alleviate data quality concerns with the incomplete coverage of births in the 2011–2015 data and consider long run trends, we repeat the above analysis for births that occurred between 1998 and 2002. These are the 5 years that directly preceded the adoption of the revised birth certificate and are thus the most recent years for which data are complete with respect to our variables of interest. Moreover, this allows for direct comparison with Green and Hamilton (6), who used the 1998–2002 period for the same reasons. Second, we repeat the main analysis on both time periods (2011–2015 and 1998–2002) restricting to the subset of states that had switched to the 2003 revised birth certificate form by 2011 in order to examine the possibility that the missing and non-missing states are fundamentally different and restricting the analysis to a subset of states is driving some of the results.

4. Findings

4.1. Description of the sample

Our main analysis is conducted on 10,343,382 births which occurred between 2011 and 2015. Before restricting our sample, we have record of 19,849,690 infant births. Restricting to singletons born to U.S.-born non-Hispanic white or non-Hispanic Black mothers between the ages of 18 and 46 residing in the U.S. at the time of birth brings this figure down to 11,786,983. Before making this restriction, mother's country of birth and race were missing from 0.3% and 4.3% of records, respectively. We keep records with imputed maternal race; 18.85% of our final sample have an imputed value for race. In terms of our main independent variables of interest, we observe maternal county of residence—and thus metropolitan status—as well as U.S. region of birth for all births to mothers residing in the U.S. Due to the changes in the birth certificate format discussed above, we do not observe maternal education for over 1 million (8.67%) records. While unobserved maternal education is not correlated with metropolitan county status, it is not evenly distributed geographically: whereas 20.87% births in the Northeast are missing maternal education, almost all education is observed in the Midwest (99.36%). Of our covariates—maternal age, marital status, prenatal care, child's sex, birth order, birth year, and U.S. region of birth—we are only missing values for birth order (0.54%) and prenatal care (11.53%). Restricting to observations for which we observe maternal education and have non-missing covariates brings us to the final analytic sample of 10,343,382 births.

Table 1 summarizes our key variables by race and maternal metropolitan status. Of the births that we analyze, 81% are to white mothers. A larger proportion of white mothers reside in nonmetropolitan counties than Black mothers—22% and 10%, respectively. Infant mortality—at any point in the first year—occurs more among infants born to Black mothers. white mothers living in metropolitan counties have the largest proportion with at least a bachelor's degree, whereas only 7% of Black mothers living in the nonmetropolitan South have at least a bachelor's degree. While a majority of metropolitan Black mothers live in the South (55%), 89% of nonmetropolitan Black mothers live in the South. white mothers in metropolitan counties are on average older than white mothers in nonmetropolitan counties and Black mothers regardless of county of residence. Fewer Black mothers are married than white mothers. A disproportionate number of births in our analytic sample come from births that occurred in 2014 and 2015 because of states' staggered adoption of the 2003 revised birth certificate form. Table 2 summarizes the same variables for the sample of 12,303,635 births that occurred between 1998 and 2002. The 1998–2002 data are used to consider long run trends in infant mortality and the robustness

TABLE 1 Summary statistics by race and metropolitan county of residence for births between 2011 and 2015.

	All	White		Black	
		Metro	Nonmetro	Metro	Nonmetro
Infant mortality (per 1,000 births)					
Infant mortality	5.10	3.80	5.06	9.46	9.80
Neonatal mortality	3.08	2.29	2.95	5.86	5.92
Postneonatal mortality	2.02	1.52	2.12	3.62	3.90
Maternal educational attainment					
< High school	0.09	0.06	0.12	0.15	0.17
High school	0.25	0.20	0.30	0.34	0.41
Some college (no degree, associate)	0.33	0.31	0.36	0.37	0.35
Bachelor's degree +	0.33	0.42	0.22	0.14	0.07
Region of birth					
Northeast	0.14	0.15	0.11	0.11	0.02
Midwest	0.29	0.29	0.40	0.24	0.08
South	0.40	0.35	0.37	0.55	0.89
West	0.17	0.21	0.13	0.09	0.01
Infant characteristics					
Male	0.51	0.51	0.51	0.51	0.51
Maternal characteristics					
Age (mean)	28.12	29.00	26.99	26.34	25.29
(SD)	(5.58)	(5.43)	(5.35)	(5.65)	(5.23)
Married	0.61	0.72	0.64	0.24	0.20
Prenatal care in the 1st trimester	0.78	0.82	0.77	0.66	0.65
No previous births	0.40	0.42	0.38	0.37	0.35
1 previous birth	0.33	0.33	0.33	0.29	0.30
2 previous births	0.16	0.15	0.18	0.18	0.19
3+ previous births	0.11	0.09	0.12	0.16	0.16
Year of birth					
2011	0.19	0.19	0.18	0.18	0.16
2012	0.19	0.19	0.19	0.19	0.16
2013	0.20	0.20	0.20	0.20	0.21
2014	0.21	0.21	0.22	0.21	0.24
2015	0.22	0.21	0.22	0.22	0.24
n	10,343,382	6,556,951	1,837,069	1,747,672	201,690

Values reported are proportions unless otherwise noted.

of our results due to the unobserved births in 2011–2015 in states that had not yet adopted the revised form.

4.2. Confirmation of educational gradients in infant mortality

Figure 1 presents the educational gradients of infant mortality by metropolitan status of county of residence for births occurring between 2011 and 2015, separately for white and Black mothers. These predicted probabilities of infant mortality—and all predicted probabilities presented in our figures—are calculated from the regression coefficients of the fully specified model (Table 3, Model 2). From the regression results in Table 3, maternal age and education are negatively associated with the probability of infant mortality in the baseline model (Model 1). However, the effect of age is eliminated after controlling for additional maternal and infant characteristics in the fully specified model (Model 2). In Model 2, for both white and Black mothers, being married and starting prenatal care in the first trimester are negatively associated with the probability of infant mortality. In all models, region is significantly associated with infant

TABLE 2 Summary statistics by race/ethnicity and metropolitan county residence for births between 1998 and 2002.

	All	White		Black	
		Metro	Nonmetro	Metro	Nonmetro
Infant mortality (per 1,000 births)					
Infant mortality	5.83	4.33	5.54	11.49	11.98
Neonatal mortality	3.61	2.69	3.25	7.27	7.69
Postneonatal mortality	2.23	1.65	2.30	4.25	4.32
Maternal educational attainment					
< 12 years	0.12	0.09	0.15	0.20	0.23
12 years	0.34	0.30	0.40	0.42	0.50
13–15 years	0.25	0.25	0.26	0.26	0.21
≥16 years	0.29	0.36	0.19	0.12	0.06
Region of birth					
Northeast	0.17	0.19	0.11	0.14	0.01
Midwest	0.27	0.26	0.37	0.23	0.05
South	0.39	0.34	0.40	0.53	0.93
West	0.17	0.20	0.12	0.09	0.01
Infant characteristics					
Male	0.51	0.51	0.51	0.51	0.51
Maternal characteristics					
Age (mean)	27.66	28.58	26.45	25.62	24.43
(SD)	(5.86)	(5.76)	(5.57)	(5.77)	(5.36)
Married	0.70	0.80	0.75	0.31	0.27
Prenatal care in the 1st trimester	0.87	0.90	0.86	0.76	0.73
No previous births	0.39	0.41	0.38	0.34	0.33
1 previous birth	0.34	0.35	0.35	0.31	0.33
2 previous births	0.17	0.16	0.17	0.19	0.20
3+ previous births	0.10	0.08	0.10	0.16	0.14
Year of birth					
1998	0.20	0.20	0.20	0.20	0.20
1999	0.20	0.20	0.20	0.20	0.20
2000	0.20	0.20	0.20	0.20	0.21
2001	0.20	0.20	0.20	0.20	0.20
2002	0.20	0.20	0.20	0.20	0.19
n	12,303,635	8,004,620	2,089,123	1,937,565	272,327

Values reported are proportions unless otherwise noted.

mortality, with the Northeast having the lowest probability of infant mortality.

The negative association between education and infant mortality is observed in Figure 1: across racial groups, more-educated mothers face a lower probability of infant death than their less-educated counterparts. Additionally, Figure 1 demonstrates the racial disparities in infant mortality. Children born to white mothers with less than a high school education living in nonmetropolitan counties have about the same predicted probability of infant mortality as children born to Black mothers with at least a bachelor's degree living in metropolitan counties. Both of these findings are in line with previous research and are presented here as confirmation of general trends before examining geographic heterogeneity.

4.3. Metropolitan gradients

Consistent with previous research, at each level of education, infant mortality tends to be higher in nonmetropolitan counties. Figure 1 demonstrates that the negative association between education and infant mortality is present across metropolitan county status for infants born to both white and Black mothers. However, there is important heterogeneity in this association.

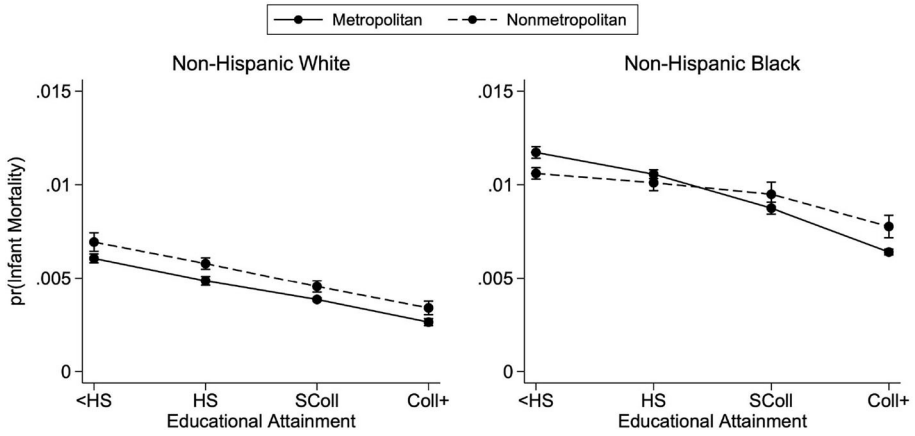


FIGURE 1
Predicted probabilities of infant mortality by metropolitan residence and race, 2011–2015. Predicted probabilities are from logistic regressions controlling for mother's age, mother's marital status, first trimester prenatal care, child's sex, birth order, birth year, U.S. region of birth, and mother's county of residence's metropolitan status. See Table 3 for full regression results. Births occurred between the years 2011 and 2015. Educational attainment is observed only from the revised birth certificates which had not yet been adopted by every state.

For white mothers, the educational gradients are downward sloping for both metropolitan and nonmetropolitan residence, and the probability of infant mortality is consistently higher in nonmetropolitan counties, across all levels of educational attainment; the lines are parallel. For Black mothers, however, this relationship is less consistent. The metropolitan gradient is much steeper than the nonmetropolitan gradient, resulting in nonmetropolitan residence predicting lower probabilities of infant mortality for mothers with less than a high school education, whereas the reverse is true for college-educated mothers.

4.4. Regional trends

In this section, we break out the above analysis by region of birth; thus we present educational gradients of infant mortality for white and Black mothers living in metropolitan and nonmetropolitan counties by U.S. region of birth (Northeast, Midwest, South, and West). Before considering these findings, it is important to note that white and Black mothers are not evenly distributed across regions and metropolitan counties (Table 1). The majority of Black mothers live in the South; of all Black mothers who reside in nonmetropolitan counties, 89% live in the South. In contrast, of all white mothers who reside in nonmetropolitan counties, 37% live in the South and 40% live in the Midwest. For both white and Black mothers, a majority live in metropolitan counties. Additionally, it is important to mention that there is geographic variation within the metropolitan and nonmetropolitan categories. Table 4 presents the distribution of births for each of the six NCHS Urban-Rural county classifications by region and race. Across each region (including the South), a higher proportion of Black mothers live in large central metropolitan counties, whereas larger proportions of white mothers live outside of large central metropolitan counties. Thus, while we conduct our analysis along the distinction of metropolitan and nonmetropolitan counties, there is heterogeneity within these categories with white and Black mothers tending to live in different types of metropolitan counties.

Figure 2 presents the predicted probabilities of infant mortality, breaking down educational gradients by both metropolitan county status and U.S. region according to the model specifications of Equation 2 (see Supplementary Table 1 for complete regression results). For white mothers, there are only small regional differences in levels and slopes of predicted probability of infant mortality, in both metropolitan and nonmetropolitan counties. Overall, the regional differences are more distinct for Black mothers. In metropolitan counties, the educational gradients of the four U.S. regions have similar slopes and a clear ordering. The Midwest is associated with the highest predicted probability of infant mortality followed by the South. The Northeast and the West have the lowest levels of predicted infant mortality for Black mothers residing in metropolitan counties. The panel for nonmetropolitan Black mothers is noisier, given that sample sizes are very small in the West and the Northeast and that 89% of births occurred in the South. Nevertheless, the remarkable flatness of the nonmetropolitan Southern Black mothers' educational gradient indicates that the flatness observed in Figure 1 is driven by Southern states. This finding does not extend to metropolitan counties. Thus, the South is driving both the nonmetropolitan *advantage* for Black mothers with low education and, on the contrary, the nonmetropolitan *disadvantage* for Black mothers with high education.

TABLE 3 Logistic regression models of infant mortality stratified by race, 2011–2015.

	Whites		Blacks	
	Model 1	Model 2	Model 1	Model 2
Maternal demographic characteristics				
Nonmetropolitan	0.059	0.069*	−0.093***	−0.091***
	(−0.020, 0.138)	(−0.005, 0.142)	(−0.154, −0.032)	(−0.127, −0.056)
Maternal age	−0.140***	−0.106***	−0.049***	−0.003
	(−0.179, −0.100)	(−0.139, −0.073)	(−0.066, −0.032)	(−0.024, 0.017)
Maternal age^2	0.002***	0.002***	0.001***	0.000***
	(0.002, 0.003)	(0.001, 0.002)	(0.001, 0.001)	(0.000, 0.001)
Mother married		−0.317***		−0.099***
		(−0.375, −0.259)		(−0.165, −0.034)
Maternal education				
High school	−0.325***	−0.223***	−0.106***	−0.106***
	(−0.377, −0.272)	(−0.281, −0.166)	(−0.126, −0.086)	(−0.129, −0.083)
Some college	−0.635***	−0.457***	−0.300***	−0.296***
	(−0.658, −0.611)	(−0.498, −0.416)	(−0.334, −0.266)	(−0.357, −0.236)
College +	-1.162***	−0.837***	−0.631***	−0.610***
	(−1.255, −1.069)	(−0.930, −0.744)	(−0.691, −0.572)	(−0.650, −0.571)
Nonmetro * Educ				
Nonmetro * HS	0.060*	0.040	0.061*	0.057
	(−0.003, 0.124)	(−0.020, 0.100)	(−0.008, 0.129)	(−0.012, 0.127)
Nonmetro * Some college	0.060***	0.032***	0.192***	0.184***
	(0.039, 0.082)	(0.011, 0.052)	(0.066, 0.318)	(0.057, 0.311)
Nonmetro * College +	0.168***	0.118**	0.316***	0.297***
	(0.054, 0.282)	(0.010, 0.226)	(0.217, 0.415)	(0.170, 0.423)
Child characteristics				
Child male		0.210***		0.179***
		(0.178, 0.241)		(0.141, 0.218)
1 prior birth		−0.046***		−0.243***
		(−0.076, −0.016)		(−0.312, −0.173)
2 prior births		0.107***		−0.238***
		(0.060, 0.153)		(−0.326, −0.150)
3 or more prior births		0.263***		−0.140***
		(0.168, 0.357)		(−0.240, −0.040)
1st trimester prenatal care		−0.296***		−0.157***
		(−0.333, −0.260)		(−0.178, −0.136)
Region of birth				
Midwest		0.207***		0.254***
		(0.197, 0.218)		(0.243, 0.265)
South		0.234***		0.170***
		(0.228, 0.241)		(0.166, 0.174)
West		0.131***		−0.017***
		(0.126, 0.135)		(−0.022, −0.012)
Year of birth				
2012		−0.010		0.012*
		(−0.042, 0.021)		(−0.001, 0.025)

(Continued)

TABLE 3 (Continued)

	Whites		Blacks	
	Model 1	Model 2	Model 1	Model 2
2013		−0.013		−0.027*
		(−0.057, 0.031)		(−0.058, 0.004)
2014		0.016		0.013
		(−0.004, 0.035)		(−0.004, 0.030)
2015		0.012		0.039***
		(−0.028, 0.052)		(0.015, 0.063)
Constant	-2.967***	-3.531***	-3.958***	-4.702***
	(−3.521, −2.413)	(−4.044, −3.018)	(−4.152, −3.765)	(−4.978, −4.426)
Observations	8,394,020	8,394,020	1,949,362	1,949,362

Robust ci in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

TABLE 4 Distribution of births between 2011 and 2015 across 2006 NCHS Urban-Rural Classification Scheme by region and race.

	Northeast		Midwest		South		West	
	White	Black	White	Black	White	Black	White	Black
Metropolitan								
Large central metro	0.20	0.64	0.14	0.53	0.17	0.31	0.33	0.64
Large fringe metro	0.33	0.19	0.24	0.17	0.25	0.21	0.17	0.17
Medium metro	0.24	0.13	0.19	0.18	0.22	0.21	0.23	0.15
Small metro	0.07	0.02	0.14	0.08	0.13	0.11	0.13	0.03
Nonmetropolitan								
Micropolitan	0.12	0.02	0.17	0.03	0.14	0.10	0.10	0.01
Non-core	0.04	0.00	0.12	0.01	0.09	0.05	0.05	0.00
n	1,200,707	199,271	2,601,143	439,337	2,953,750	1,144,138	1,638,420	166,616

Values reported are proportions of all births in our analytic sample that were born to mothers living in each of the six county classifications from the 2006 NCHS Urban-Rural Classification Scheme. Proportions are broken out by race and region of birth.

4.5. Neonatal and postneonatal trends

The next analysis considers how the educational gradients by metropolitan residence and race vary by timing of infant death. Figures 3, 4 break out Figure 1 by whether the infant died in the first 28 days of life or between days 28 and 364, respectively. For white mothers, the neonatal and postneonatal educational gradients look remarkably similar. In Figures 3, 4, nonmetropolitan county residence is associated with a slightly higher probability of infant mortality—at either time range—for white mothers, and the slopes are very similar between metropolitan and nonmetropolitan counties. There is a greater distinction between the neonatal and postneonatal educational gradients for infants born to Black mothers. Their predicted probabilities of neonatal mortality do not differ between metropolitan and nonmetropolitan counties at any maternal education level besides less than high school (Figure 3). Black mothers with less than a high school education living in metropolitan counties have a higher probability of neonatal mortality than their counterparts residing in nonmetropolitan counties. In terms of postneonatal mortality for infants born to Black mothers, the divergence occurs at the other end of maternal educational attainment (Figure 4). While there is no difference in predicted probability of postneonatal mortality for infants born to Black mothers with lower levels of education, there is a higher probability of postneonatal mortality for infants born to mothers with at least a bachelor's degree and who live in nonmetropolitan counties. Contrary to what was observed in previous studies (6), we find no marked difference in slope between the neonatal and postneonatal gradients.

4.6. Robustness checks

4.6.1. Temporal comparison

Next, we have repeated our analysis looking at educational gradients by metropolitan residence status for an earlier time period where all states could be included. Figure 5 presents

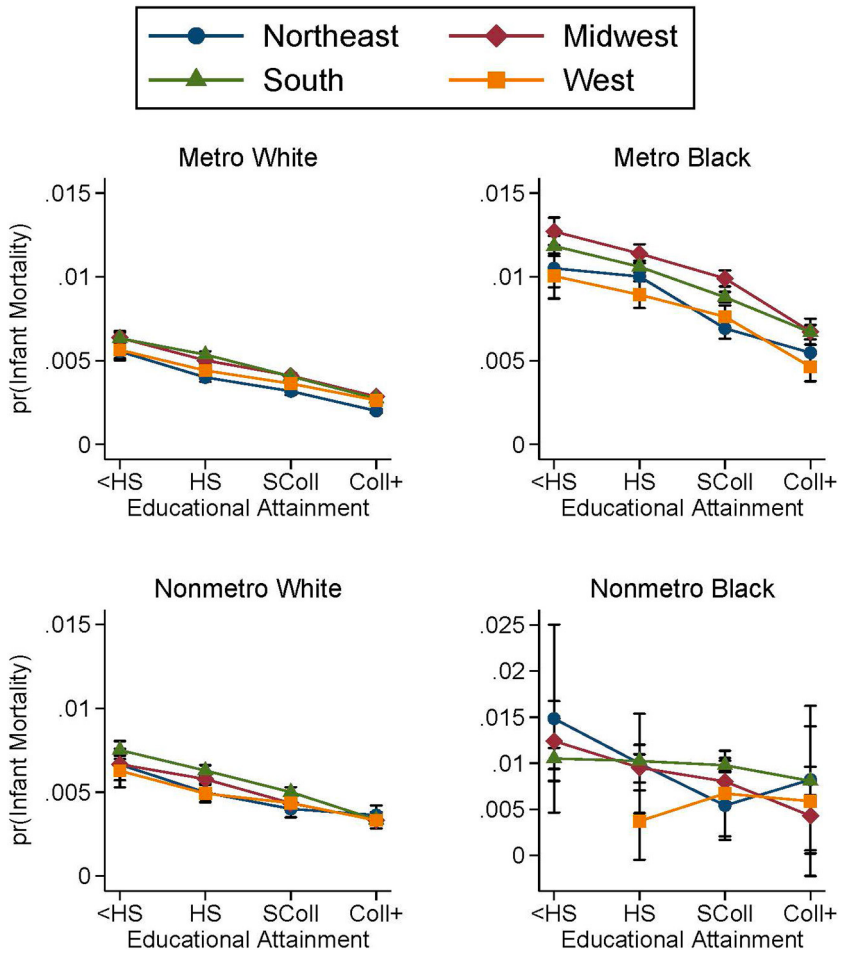


FIGURE 2
Predicted probabilities of infant mortality by metropolitan residence and region for white and Black mothers, 2011–2015. Predicted probabilities are from logistic regressions controlling for mother's age, mother's marital status, first trimester prenatal care, child's sex, birth order, birth year, U.S. region of birth, and mother's county of residence's metropolitan status as well as the complete interaction between metropolitan status, region, and educational attainment. See Supplementary Table 1 for full regression results. Births occurred between the years 2011 and 2015. Educational attainment is observed only from the revised birth certificates which had not yet been adopted by every state. The top panels plot the educational gradients of infant mortality for white and Black mothers living in metropolitan counties. The bottom panels plot the educational gradients of infant mortality for white and Black mothers living in nonmetropolitan counties.

the predicted probabilities of infant mortality by race and metropolitan residence status for births that occurred between 1998 and 2002 and suggests largely similar patterns. Again, higher education is associated with lower infant mortality, though there is heterogeneity across race and metropolitan residence status. For infants born between 1998 and 2002 to white mothers, there is a metropolitan advantage with infants born to mothers residing in metropolitan counties having a lower predicted probability of infant mortality at any level of educational attainment. Black mothers' gradients are also negatively sloped and predict higher levels of infant mortality than white mothers. Similar to the 2011–2015 period, the educational gradient is flatter for nonmetropolitan mothers than metropolitan mothers. At the two highest levels of maternal education, there is a higher predicted probability of infant mortality for infants born to Black mothers living in nonmetropolitan counties.

As suggested by recent work, there was a sharp decline in infant mortality for infants born to Black mothers across levels of maternal education and metropolitan residence between 1998–2002 and 2011–2015. Despite this progress, infants born to Black mothers continue to face higher probabilities of infant mortality at any level of education when compared to infants born to white mothers. This time comparison shows a persistence in racial, geographic, and educational patterns in infant mortality, which provides support for our results despite incomplete data. However, it is not possible to more directly compare the two time periods, because, as noted above, the

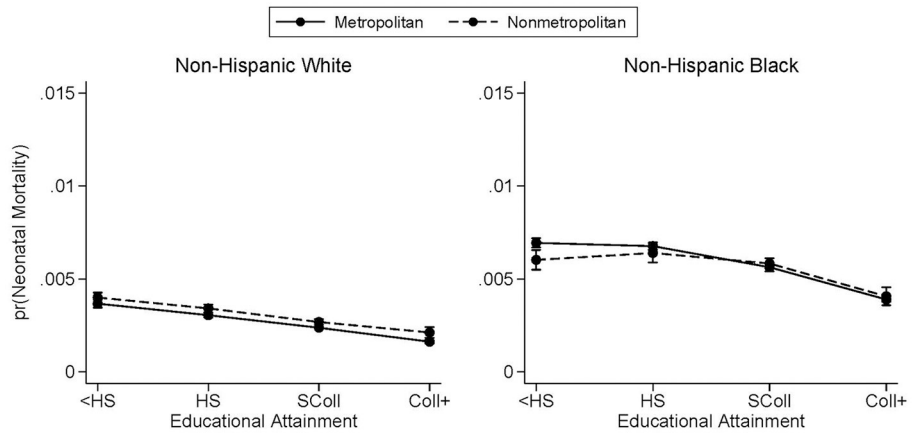


FIGURE 3
Predicted probabilities of neonatal mortality by metropolitan residence and race. Neonatal mortality is infant death occurring between days 0 and 27. Predicted probabilities are from logistic regressions controlling for mother's age, mother's marital status, first trimester prenatal care, child's sex, birth order, birth year, U.S. region of birth, and mother's county of residence's metropolitan status. See Supplementary Table 2 for full regression results. Births occurred between the years 2011 and 2015. Educational attainment is observed only from the revised birth certificates which had not yet been adopted by every state.

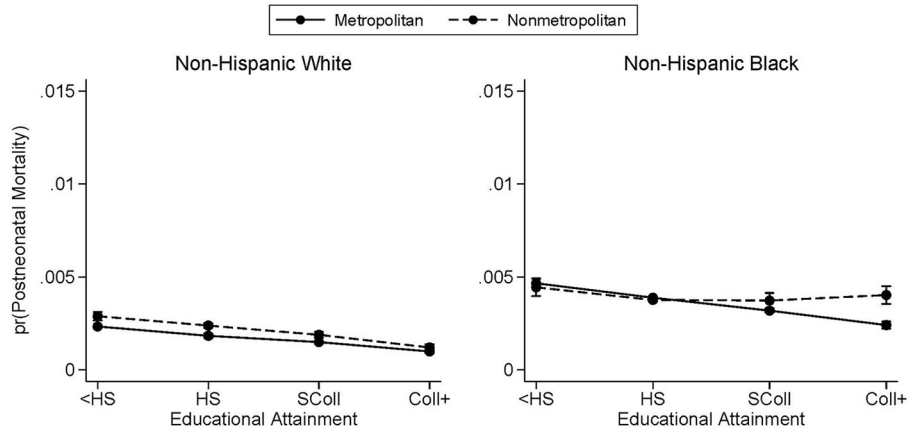


FIGURE 4
Predicted probabilities of postneonatal mortality by metropolitan residence and race. Postneonatal mortality is infant death occurring between days 28 and 364. Predicted probabilities are from logistic regressions controlling for mother's age, mother's marital status, first trimester prenatal care, child's sex, birth order, birth year, U.S. region of birth, and mother's county of residence's metropolitan status. See Supplementary Table 3 for full regression results. Births occurred between the years 2011 and 2015. Educational attainment is observed only from the revised birth certificates which had not yet been adopted by every state.

revised birth certificate format changed how maternal education was measured.

4.6.2. Robustness of states included

Due to the staggered adoption of the revised birth certificate form, we do not observe maternal education in all states between 2011 and 2015. To address the possibility that the absence of some states in the more recent data is driving some of the changes between the two time periods, we also run the analysis for both time periods restricting the samples to the U.S. states that switched to the revised birth certificate before 2011 (i.e., states for which we have data on maternal education for all years). Figure 6 plots the 14 states that had not yet adopted the revised birth certificate form before 2011. We conduct the above analysis on the 36 states that had revised their form as well as the District of Columbia. Figures 7, 8 plot the educational gradients of infant mortality by race and metropolitan residence for births occurring between 2011–2015 and 1998–2002, respectively, restricting to states that had revised their birth certificates before 2011. We find very similar results in both time periods with this restricted sample. Thus, we conclude that the absent states in the

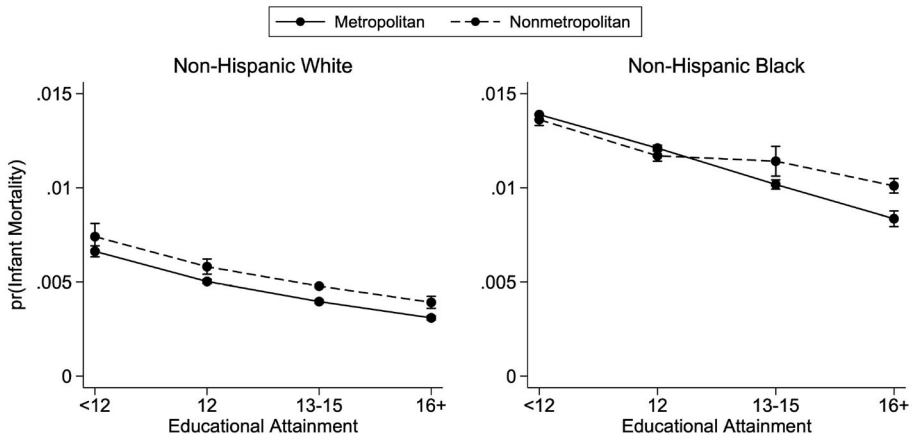


FIGURE 5
Predicted probabilities of infant mortality by metropolitan residence and race—1998–2002. Predicted probabilities are from logistic regressions controlling for mother's age, mother's marital status, first trimester prenatal care, child's sex, birth order, birth year, U.S. region of birth, and mother's county of residence's metropolitan status. See Supplementary Table 4 for full regression results. Births occurred between the years 1998 and 2002.

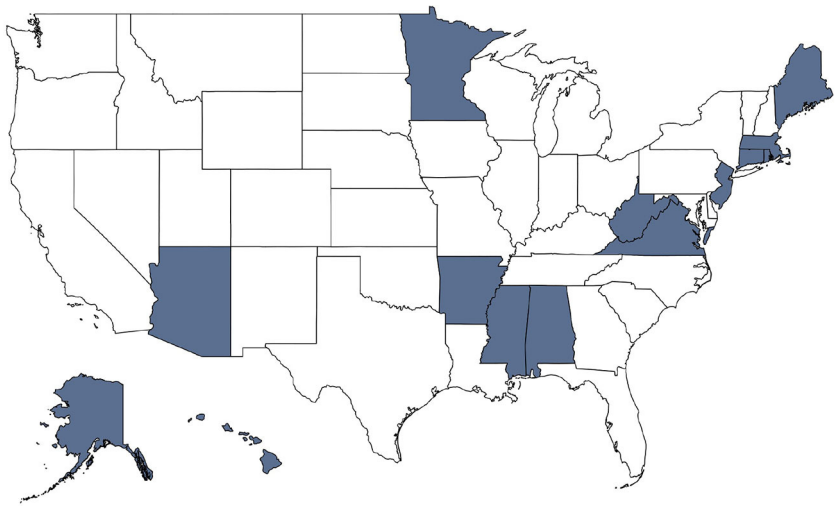


FIGURE 6
States' adoption of revised birth certificate form by 2011. This map indicates with blue shading the 11 states that had not begun recording births using the 2003 revised birth certificate form before 2011.

newer years are not driving the observed trends and focus our discussion on the analysis using the full sample of infants for whom we observe maternal education.

5. Discussion

Through our findings, we have documented the negative association between maternal education and infant death as well as within-race heterogeneity in the association between education and infant mortality, across both metropolitan status and region. These differences are subtle for infants born to white women but substantial and meaningful for infants born to Black women. The main finding of this paper is the remarkable flatness of the educational gradient of nonmetropolitan Black mothers, driven by Southern states and across both the neonatal and postneonatal periods. This observation could not have been made without this paper's approach looking at the intersection of education and geography. Below we discuss our main findings in more detail for white and Black mothers and some potential mechanisms before outlining this study's limitations and contributions.

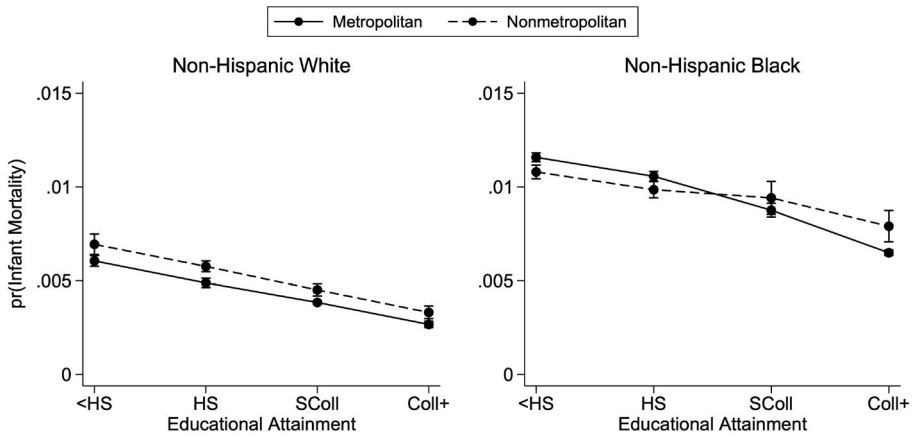


FIGURE 7
Predicted probabilities of infant mortality by metropolitan residence and race, 2011–2015, state subsample. Educational gradients of infant mortality by race and metropolitan residence for states that had adopted the 2003 revised birth certificate form before 2011. Births occurred between 2011 and 2015. Supplementary Table 5 has full regression results from which the predicted probabilities are calculated. See Figure 1 for complete figure notes and Figure 6 for states included in subsample.

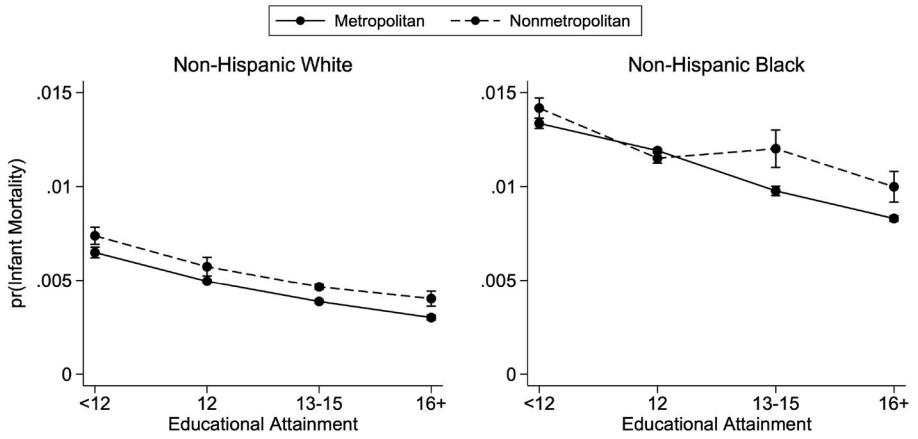


FIGURE 8
Predicted probabilities of infant mortality by metropolitan residence and race, 1998–2002, state subsample. Educational gradients of infant mortality by race and metropolitan residence for states that had adopted the 2003 revised birth certificate form before 2011. Births occurred between 1998 and 2002. Supplementary Table 6 has full regression results from which the predicted probabilities are calculated. See Figure 5 for complete figure notes and Figure 6 for states included in subsample.

We find that geographic differences in infant mortality exist more starkly for infants born to Black mothers than to white mothers. The small regional differences for infants born to white mothers and the consistent, yet small, difference between metropolitan and nonmetropolitan counties suggest that, for white mothers, the detrimental effect of living in a nonmetropolitan county is fairly constant across educational levels at the national scale and for each of the four regions. This finding relates to a broader literature on rural disadvantage that has emphasized difficulties in accessing quality healthcare (37) and also reflects lower maternal age at birth and lower proportions of married mothers, factors that are suggestive of single motherhood as well as higher instability and stress (38).

The relationship between education and geographic residence for infants born to Black mothers, however, requires a more nuanced discussion. Black mothers display distinct patterns of infant mortality when considered through the intersection of both education and geography. Although a marked racial gap exists in all four census regions, the metropolitan Midwest and South contribute the most to the overall Black-white gap in infant mortality. The flat educational gradient observed for nonmetropolitan mothers

at the national level is driven by the South, where the vast majority of nonmetropolitan Black mothers live. The difference between metropolitan and nonmetropolitan counties is not constant across levels of maternal education, as evidenced by the crossover between the gradients. At the lowest level of maternal educational attainment, nonmetropolitan mothers fare better than their metropolitan counterparts. This is in contrast to the metropolitan advantage observed for infants born to white mothers as well as the metropolitan advantage that is present for infants born to Black mothers with at least a bachelor's degree.

This geographic pattern has persisted in the past two decades despite the overall decline in infant mortality and the reduction of the absolute gap in infant mortality between Black and white infants, as documented by previous studies (4) and our comparison of the time periods 1998–2002 and 2011–2015. The relative flatness of the educational gradient in infant mortality for nonmetropolitan Black mothers does not explain the overall racial gap. However, it can be interpreted as evidence of low returns to education for infants born to Black mothers in nonmetropolitan Southern counties, which challenges the well-known role of education as a powerful lever for improving health.

These findings point to potentially crucial mechanisms of rural and micropolitan Black health in the United States, with important policy implications. It can be thought that this evidence of low returns to education for the nonmetropolitan Black population in the South, as it manifests through the first year of infancy, stems from more fundamental sources that likely affect a variety of outcomes for which an educational gradient is expected. This finding could be symptomatic of a wider phenomenon impacting multiple facets of life. Below we explore four potential mechanisms that could contribute to this finding. Rather than explanations, these should be taken as avenues for future research.

First, it is possible that the Southern nonmetropolitan counties where these Black mothers live—which includes the Black Belt region known for its high proportions of African Americans and persistent disadvantage through legacies of slavery—offer few opportunities for highly educated mothers to leverage education into higher income and better living situations. Lower observed or perceived quality of education and racial employment discrimination could be a barrier to employment opportunities for college-educated Black mothers. Unfortunately, education is the only measure of socioeconomic status available on U.S. birth certificates, which prevents us from directly testing the hypothesis that the flat educational gradient in infant mortality for nonmetropolitan Black mothers derives from a flatter educational gradient in income.

Second, the selection effect of migration to metropolitan areas could contribute to the Black infant mortality gap between metropolitan and nonmetropolitan areas. Positive selection of migrants with respect to health is generally observed with young internal migrants tending to be healthier than their non-migrant peers. Migration of young adults to metropolitan areas also signals motivation to access better opportunities and living conditions, which could also reflect in higher degrees of health consciousness. However, opportunity-motivated migration comes with potentially detrimental consequences, such as higher stress and weaker social networks (39–41). These mechanisms and their implications for the health gradients of metropolitan and nonmetropolitan populations could be differentiated by race. However, it remains unclear how urban-rural migration within the United States would affect our finding of a flatter educational gradient for Black mothers in nonmetropolitan counties.

Third, given the residential segregation that persists in the U.S. South between Black and white populations, it is possible that nonmetropolitan Black mothers, regardless of education, have worse access to healthcare than their white counterparts, and that health institutions present in predominantly Black counties are of lower quality. The history of racial discrimination in healthcare may also cause Black mothers to receive poorer services or be more reluctant to trust medical professionals even when they do have access to healthcare. For example, Pathman et al. (42) report that Black adults in the rural South experience more dissatisfaction and barriers to care than whites, and research has shown that the racial gap in adverse birth outcomes is connected with levels of racial prejudice in both the county of residence and the county of birth (43). However, access to healthcare cannot account for the entirety of observed trends, in part because it does not explain the steeper educational gradient in metropolitan counties and the nonmetropolitan advantage in neonatal mortality for Black mothers without a high school degree. Landrine and Corral (44) note that residential segregation can have other impacts on health, through differential exposure to environmental conditions, pollutants, and toxins, as well as disparities in the built environment shaping access to and use of fast food, grocery stores, and recreational facilities.

Lastly, it can be thought that education has a weaker association with the adoption of beneficial health behaviors for Black mothers living in nonmetropolitan areas. This hypothesis is partly supported by research suggesting that education is a significant predictor of health consciousness for whites and Hispanics, but not for Blacks (45). This could arise from lower quality education or result from different social dynamics, social network structures and discourse around health in nonmetropolitan Black communities.

These last two potential explanations—access to healthcare and adoption of beneficial health behaviors—are related to our analysis of neonatal and postneonatal mortality insomuch as timing of infant mortality indicates different underlying patterns of causes of death. Because the nonmetropolitan advantage occurs during the neonatal period for Black mothers with less than a high school education, this could suggest

that, contrary to initial expectations, it might be harder for mothers with low educational attainment to access quality obstetric and neonatal healthcare in metropolitan counties. Additionally, in all four census regions, Black mothers are much more likely than white mothers to reside in large central metropolitan counties (Table 4). Higher neonatal mortality for the least educated Black women in these areas might result from detrimental contextual conditions, rather than geographical access to healthcare. Moreover, the very similar probabilities of postneonatal death across educational attainment for nonmetropolitan Black mothers suggest that behaviors and environmental conditions may not vary across socioeconomic status. As noted earlier, these factors, which could include child nutrition, parental supervision, smoking, and use of appropriate indoor and outdoor recreational spaces, may be shaped by the built environment.

Although our analysis highlights that nonmetropolitan residence limits Black mothers' returns to education with respect to infant mortality—particularly in the South—, this paper can only hypothesize about the mechanisms underlying this finding. Additional research is needed to shed light on these processes. This paper only provides a partial picture of infant mortality in the U.S. given the sample restrictions, excluding immigrants, Hispanics, Asians, Pacific Islanders, Indigenous peoples, and other ethnoracial groups. We also acknowledge that the data limit us to considering race through categorical identities, despite the complexity of this concept and the fuzzy boundaries between racial groups, especially for multiracial individuals.

A further limitation to this study is the potential for intracategorical variation, both in terms of metropolitan status and education. While our analysis focuses on the distinction between metropolitan and nonmetropolitan counties, the distribution of mothers across the six NCHS urban-rural codes (see Table 4) suggests differences in the locations within the categories of metropolitan and nonmetropolitan where Black and white mothers reside. However, infant mortality is a rare enough occurrence that the data become too sparse to be able to meaningfully compare educational gradients between more fine-grained geographic areas. We also cannot consider intracounty differences in the types of neighborhoods where Black and white mothers live. Thus, further research is needed to continue to understand the heterogeneity within the categories of metropolitan and nonmetropolitan. There is also potential for intracategorical variation in the four-level scale of educational attainment across races and metropolitan county status. For example, we expect higher proportions of college-educated mothers with additional degrees in metropolitan areas, where most research institutions are located, which could contribute to lower levels of postneonatal mortality for college-educated Black mothers in metropolitan counties. Lastly, we want to reiterate that education is not a perfect proxy for socioeconomic status and does not explain all the life course differences between Black and white mothers. Education is acquired with a range of intentions, challenges and results; it also fails to reflect the continuing effects of past socioeconomic status during childhood and other life course experiences. This is especially salient in the health context, as behaviors and health are shaped through childhood.

This paper contributes to the literature on the ways through which mothers' situation and context impact infants' health and mortality. Beyond confirming the persistence of the well-known racial gap in infant mortality and the negative association with maternal educational attainment, we have documented the smaller returns to education for Black mothers living in nonmetropolitan counties, a pattern that is observed across the periods 1998–2002 and 2011–2015. This finding offers a new axis for research and policy intervention focusing on issues relating to limited returns to education for the nonmetropolitan Black population living in the U.S. South. While metropolitan residence and region cannot account for the overall racial gap, the fact that geographic variations are much more salient for Black mothers than for white mothers suggests broader issues related to availability and quality of healthcare and education as well as persistent social stress and discrimination in Black communities. This points to the importance of looking at infant health in a holistic perspective, beyond individual and household characteristics. Considering the geographic dimensions of these dynamics and their persistence over time helps to understand the systemic and ingrained nature of disparities.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: the datasets analyzed for this study are restricted and can be requested from NCHS. Public versions of these data, which do not include geographic information, can be downloaded from the National Bureau of Economic Research website. Requests to access these datasets should be directed to https://www.cdc.gov/nchs/nvss/nvss-restricted-data.htm.

Author contributions

MCG wrote the first draft of the manuscript. KDM completed the statistical analysis. MCG and KDM contributed to the conception, manuscript writing and revision, and read and approved the submitted version.

Funding

This publication was supported by the Princeton University Library Open Access Fund.

Acknowledgments

The authors acknowledge the feedback from Yu Xie, Tod Hamilton, Neil Mehta, the participants of Princeton Sociology's Stratification Workshop, classmates in SOC 504, and two reviewers, and thank the National Center for Health Statistics for providing access to restricted-use vital statistics data.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.995585/full#supplementary-material

References

1. Reidpath D, Allotey P. Infant mortality rate as an indicator of population health. *J Epidemiol Commun Health.* (2003) 57:344–6. doi: 10.1136/jech.57.5.344

2. Wise PH, Pursley DM. Infant mortality as a social mirror. *N Engl J Med.* (1992) 326:1558–60. doi: 10.1056/NEJM199206043262309

3. Chen A, Oster E, Williams H. Why is infant mortality higher in the United States than in Europe? *Amer Econ J Econ Policy.* (2016) 8:89–124. doi: 10.1257/pol.20140224

4. Riddell CA, Harper S, Kaufman JS. Trends in differences in US mortality rates between Black and White infants. *JAMA Pediatr.* (2017) 171:911–3. doi: 10.1001/jamapediatrics.2017.1365

5. Schoendorf KC, Hogue CJR, Kleinman JC, Rowley D. Mortality among infants of Black as compared with White college-educated parents. *N Engl J Med.* (1992) 326:1522–6. doi: 10.1056/NEJM199206043262303

6. Green T, Hamilton T. Maternal educational attainment and infant mortality in the United States: does the gradient vary by race/ethnicity and nativity? *Demogr Res.* (2019) 41:713–52. doi: 10.4054/DemRes.2019.41.25

7. Fishman SH, Hummer RA, Sierra G, Hargrove T, Powers DA, Rogers RG. Race/ethnicity, maternal educational attainment, and infant mortality in the United States. *Biodemography Soc Biol.* (2021) 66:1–26. doi: 10.1080/19485565.2020.1793659

8. Currie J, Moretti E. Mother's education and the intergenerational transmission of human capital: evidence from college openings. *Quart J Econ.* (2003) 118:1495–532. doi: 10.1162/003355303322552856

9. Acevedo-Garcia D, Bates LM, Osypuk TL, McArdle N. The effect of immigrant generation and duration on self-rated health among US adults 2003–2007. *Soc Sci Med.* (2010) 71:1161–72. doi: 10.1016/j.socscimed.2010.05.034

10. Vable AM, Cohen AK, Leonard SA, Glymour MM, Duarte CdP, Yen IH. Do the health benefits of education vary by sociodemographic subgroup? Differential returns to education and implications for health inequities. *Ann Epidemiol.* (2018) 28:759.e5–66.e5. doi: 10.1016/j.annepidem.2018.08.014

11. Balcazar AJ, Grineski SE, Collins TW. The Hispanic health paradox across generations: the relationship of child generational status and citizenship with health outcomes. *Publ Health.* (2015) 129:691–7. doi: 10.1016/j.puhe.2015.04.007

12. Perreira KM, Ornelas IJ. The physical and psychological well-being of immigrant children. *Future Child.* (2011) 21:195–218. doi: 10.1353/foc.2011.0002

13. Hendi AS, Mehta NK, Elo IT. Health among Black children by maternal and child nativity. *Amer J Publ Health.* (2015) 105:703–10. doi: 10.2105/AJPH.2014.302343

14. Van Hook J, Landale NS, Hillemeier MM. *Is the United States Bad for Children's Health? Risk and Resilience Among Young Children of Immigrants.* Washington, DC: Migration Policy Institute (2013).

15. Allen DM, Buehler JW, Hogue CJ, Strauss LT, Smith JC. Regional differences in birth weight-specific infant mortality, United States, 1980. *Publ Health Rep.* (1987) 102:138–45.

16. Montez JK, Berkman LF. Trends in the educational gradient of mortality among US adults aged 45 to 84 years: bringing regional context into the explanation. *Amer J Publ Health.* (2014) 104:e82–90. doi: 10.2105/AJPH.2013.301526

17. Rossen LM, Khan D, Schoendorf KC. Mapping geographic variation in infant mortality and related Black-White disparities in the US. *Epidemiology.* (2016) 27:690–6. doi: 10.1097/EDE.0000000000000509

18. Sparks PJ, McLaughlin DK, Stokes CS. Differential neonatal and postneonatal infant mortality rates across US counties: the role of socioeconomic conditions and rurality. *J Rural Health.* (2009) 25:332–41. doi: 10.1111/j.1748-0361.2009.00241.x

19. Yao N, Matthews SA, Hillemeier MM. White infant mortality in appalachian states, 1976–1980 and 1996–2000: changing patterns and persistent disparities. *J Rural Health.* (2012) 28:174–82. doi: 10.1111/j.1748-0361.2011.00385.x

20. Luo ZC, Wilkins R. Degree of rural isolation and birth outcomes. *Paediatr Perin Epidemiol.* (2008) 22:341–9. doi: 10.1111/j.1365-3016.2008.00938.x

21. Anderson RN. Deaths: leading causes for 2000. *Natl Vital Stat Rep.* (2002) 50:1-85. Available online at: https://www.cdc.gov/nchs/data/nvsr/nvsr50/nvsr50_16.pdf

22. Heron M. Deaths: leading causes for 2013. *Natl Vital Stat Rep.* (2016) 65:1-95. Available online at: https://www.cdc.gov/nchs/data/nvsr/nvsr65/nvsr65_02.pdf

23. Hessol NA, Fuentes-Afflick E. Ethnic differences in neonatal and postneonatal mortality. *Pediatrics.* (2005) 115:e44–51. doi: 10.1542/peds.2004-0478

24. Hendi AS. Trends in U.S. life expectancy gradients: the role of changing educational composition. *Int J Epidemiol.* (2015) 44:946–55. doi: 10.1093/ije/dyv062

25. Elo IT, Hendi AS, Ho JY, Vierboom YC, Preston SH. Trends in non-Hispanic White mortality in the United States by metropolitan-nonmetropolitan status and region, 1990–2016. *Popul Dev Rev.* (2019) 45:549–83. doi: 10.1111/padr.12249

26. National Center for Health Statistics. *Birth-Cohort Linked Births/Infant Deaths – All Counties (1998–2002, 2011–2015) as Compiled from Data Provided by the 57 Vital Statistics Jurisdictions Through the Vital Statistics Cooperative Program.* (2022).

27. Riosmena F, Kuhn R, Jochem WC. Explaining the immigrant health advantage: self-selection and protection in health-related factors among five major national-origin immigrant groups in the United States. *Demography.* (2017) 54:175–200. doi: 10.1007/s13524-016-0542-2

28. Kozuki N, Lee AC, Silveira MF, Sania A, Vogel JP, Adair L, et al. The associations of parity and maternal age with small-for-gestational-age, preterm, and neonatal and infant mortality: a meta-analysis. *BMC Publ Health.* (2013) 13:S2. doi: 10.1186/1471-2458-13-S3-S2

29. King G, Zeng L. Logistic regression in rare events data. *Polit Anal.* (2001) 9:137–63. doi: 10.1093/oxfordjournals.pan.a004868

30. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol.* (1996) 49:1373–9. doi: 10.1016/s0895-4356(96)00236-3

31. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and cox regression. *Amer J Epidemiol.* (2007) 165:710–8. doi: 10.1093/aje/kwk052

32. Courvoisier DS, Combescure C, Agoritsas T, Gayet-Ageron A, Perneger TV. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure. *J Clin Epidemiol.* (2011) 64:993–1000. doi: 10.1016/j.jclinepi.2010.11.012

33. Wynants L, Bouwmeester W, Moons KGM, Moerbeek M, Timmerman D, Van Huffel S, et al. A simulation study of sample size demonstrated the importance of the number of events per variable to develop prediction models in clustered data. *J Clin Epidemiol.* (2015) 68:1406–14. doi: 10.1016/j.jclinepi.2015.02.002

34. van Smeden M, de Groot JAH, Moons KGM, Collins GS, Altman DG, Eijkemans MJC, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Med Res Methodol.* (2016) 16:163. doi: 10.1186/s12874-016-0267-3

35. Austin PC, Steyerberg EW. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. *Stat Methods Med Res.* (2017) 26:796–808. doi: 10.1177/0962280214558972

36. van Smeden M, Moons KG, de Groot JA, Collins GS, Altman DG, Eijkemans MJ, et al. Sample size for binary logistic prediction models: beyond events per variable criteria. *Stat Methods Med Res.* (2019) 28:2455–74. doi: 10.1177/0962280218784726

37. James WL. All rural places are not created equal: revisiting the rural mortality penalty in the United States. *Am J Public Health.* (2014) 104:2122–9. doi: 10.2105/AJPH.2014.301989

38. Bennett T, Braveman P, Egerter S, Kiely JL. Maternal marital status as a risk factor for infant mortality. *Fam Plann Perspect.* (1994) 26:252–6, 271.

39. Halliday TJ, Kimmitt MC. Selective migration and health in the USA, 1984–93. *Popul Stud.* (2008) 62:321–34. doi: 10.1080/00324720802339806

40. Darlington F, Norman P, Gould M. Health and internal migration. In: Smith DP, Finney N, Halfacree K, Walford N. *Internal Migration: Geographical Perspectives and Processes*. Ashgate Publishing, Ltd. (2015). pp. 113–28, 129–44. Available online at: https://www.taylorfrancis.com/chapters/edit/10.4324/9781315589275-15/health-internal-migration-fran-darlington-paul-norman-myles-gould.

41. Acevedo-Garcia D, Sanchez-Vaznaugh EV, Viruell-Fuentes EA, Almeida J. Integrating social epidemiology into immigrant health research: a cross-national framework. *Soc Sci Med.* (2012) 75:2060–8. doi: 10.1016/j.socscimed.2012.04.040

42. Pathman DE, Fowler-Brown A, Corbie-Smith G. Differences in access to outpatient medical care for Black and White adults in the rural South. *Med Care.* (2006) 44:429–38. doi: 10.1097/01.mlr.0000207487.85689.a8

43. Orchard J, Price J. County-level racial prejudice and the Black-White gap in infant health outcomes. *Soc Sci Med.* (2017) 181:191–8. doi: 10.1016/j.socscimed.2017.03.036

44. Landrine H, Corral I. Separate and unequal: residential segregation and Black health disparities. *Ethnic Dis.* (2009) 19:179–84. Available online at: https://www.ethndis.org/priorarchives/ethn-19-02-179.pdf

45. Nevarez L, Hovick SR, Enard KR, Lloyd SM, Kahlor LA. Race/Ethnic variations in predictors of health consciousness within the cancer prevention context. *Amer J Health Promot.* (2020) 34:740–6. doi: 10.1177/0890117120904000

frontiers | Frontiers in Public Health

TYPE Brief Research Report
PUBLISHED 17 November 2022
DOI 10.3389/fpubh.2022.1003117

OPEN ACCESS

EDITED BY
Irma Elo,
University of Pennsylvania,
United States

REVIEWED BY
Neil Mehta,
University of Texas Medical Branch at Galveston, United States
Atheendar Venkataramani,
University of Pennsylvania,
United States

*CORRESPONDENCE
Douglas A. Wolf
dawolf@syr.edu

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 25 July 2022
ACCEPTED 28 October 2022
PUBLISHED 17 November 2022

CITATION
Wolf DA (2022) State policies that promote, and that inhibit, improved public health: An exploratory analysis of paid sick leave.
Front. Public Health 10:1003117.
doi: 10.3389/fpubh.2022.1003117

State policies that promote, and that inhibit, improved public health: An exploratory analysis of paid sick leave

Douglas A. Wolf*

Aging Studies Institute, Syracuse University, Syracuse, NY, United States

The United States has no national requirement that employers provide paid sick leave (PSL) to their employees, despite the many established public health benefits of PSL access. Many states, and some localities, have passed laws requiring PSL within their jurisdictions. Past studies have shown that these PSL mandates are effective in promoting increased PSL access. However, past studies have not considered two other commonly-used state policy initiatives—PSL preemption and right-to-work laws—that could hypothetically influence employers' decisions to provide PSL. During the past few decades, all possible combinations of these policy interventions can be found in one or more U.S. states. This study estimates the combined associations of these 3 policies with PSL access. The estimates support recent research on the positive effects of PSL mandates, but also suggest that PSL preemption and right-to-work laws may have offsetting effects. Failure to take account of these additional policies may lead to an over-estimate of the effectiveness of PSL mandates.

KEYWORDS

paid sick leave, preemption, right-to-work laws, public health, labor unions

Introduction

Access to employer-provided paid sick leave (PSL) has been shown to be beneficial to employees and employers, and to improve public health. Workers with PSL are less likely to show up for work when ill (1, 2), are less likely to delay seeking medical care (2), and increase their use of outpatient care (3) while reducing their usage of emergency department care (4). Workers with PSL access have also been found to have higher levels of retirement savings (5). Employers that offer PSL experience reduced rates of occupational injuries and illness (6) and of overall leave-taking (7) among their workforces. Some studies have identified specific forms of illness—influenza (8) and food-borne illness (9), for example—whose prevalence is diminished by employer-provided PSL. Employers also experience benefits in the form of lower rates of both employee separation (10) and a more general indicator—employee turnover—that encompasses hires as well as separations (11). Finally, PSL has been shown to reduce both all-cause mortality and mortality from specific causes among working-age adults (12, 13).

Along with these benefits, providing PSL also imposes costs on employers, who must pay workers for the accrued sick time that they take; also, the aggregate amount of employee time lost to illness may be greater when PSL is provided than when it is not. In March 2020 it was estimated that the average cost of PSL to employers was $0.45 per employee hour worked (14). These costs, in turn, are equal to 2.5–3.3% of employees' wage compensation (15).

Until recently, private employers have not faced any legal requirements to offer PSL benefits. Private-sector fringe benefits, of which PSL is one component, were rare prior to World War II (16), but have grown rapidly since then. In the absence of legal requirements, employers' provision of benefits has been ascribed to bargaining by labor union, a firm's stability and profitability, religious and ethical concerns of corporate leaders, and institutional factors such as the professionalization of human resources personnel (17). The U.S. Labor Department reported on the prevalence of employee benefits for the first time in 1979, at which time 56% of full-time employees in private-sector industries had access to PSL (18). By 2020 this figure had grown to 78% (19).

However, there is no Federal law requiring PSL coverage in the private sector, and as a consequence there remains great variation in PSL coverage by occupation, region, industry, and wage level, among other factors. As a way of broadening PSL access and reducing inequality in access to it, a growing number of U.S. states have begun to pass laws that mandate a minimum (or "floor") level of PSL coverage (20). The first such PSL mandate was adopted in Washington DC in 2008, and by 2021, 14 states had such laws in place. Several local-level governments—counties or cities—have also passed laws mandating PSL coverage for workers within their jurisdictions (21).

Two recent studies have investigated the effectiveness of PSL mandates with respect to increased employee access. Maclean et al. (15) used restricted-access individual job-level data from the US Labor Department's annual Employee Compensation Survey (ECS) for 2009 through 2017 to estimate the impact of state PSL mandates on PSL access. Using a difference-in-differences (DD) methodology, with additional controls for a worker's union membership and full-time status, they find that on average, access to PSL for workers in PSL mandate states is 12.8% higher than for their counterparts in non-mandate states, a difference that is statistically significant. A second paper, by Callison and Pesko (22), also used restricted-access individual-level data, taken from the National Health Interview Survey for 2005–2018, and considered both state- and local-level PSL mandates. Their DD estimates also find statistically significant increases in PSL coverage attributable to the PSL mandates, ranging from about 8 to over 20 percentage points, in various model specifications.

Both of these recent studies provide strong evidence that PSL mandates produce an increase in PSL coverage. However, neither study takes into account other state-level policies that might influence PSL access, and that might even counteract the positive effects of the mandates on access. Variation within both the treatment-group and the comparison-group jurisdictions along relevant policy dimensions could undermine the validity of the estimated PSL mandate effects. This study considers two such policy domains: PSL preemption and so-called "right-to-work" (RTW) laws.

State PSL preemption laws, which restrict the ability of lower-level governments to impose PSL requirements, are one manifestation of a larger and growing phenomenon whereby states restrict their constituent governments' actions in areas such as environmental, health, and public-safety domains (23). During the period 2009 through 2021, the number of states with PSL preemption laws grew from 1 to 24 (20). PSL mandates can be characterized as a type of preemption, because they establish a floor level of benefits below which local governments cannot depart. Alternatively, states can pass "ceiling" preemption laws that prevent local governments from requiring even higher levels of PSL. The ceiling is often set at zero, effectively ruling out a government-mandated PSL requirement statewide. A few states have passed a PSL mandate (a floor) while simultaneously imposing ceiling preemption, thereby establishing a floor level of PSL benefit within the state but also preventing lower-level governments from going above that level. PSL mandates and ceiling preemption can be viewed either as two separate policy domains, or can be interacted so as to identify three policy regimes (mandate without ceiling, ceiling without mandate, or mandate with ceiling).

Paid sick leave preemption surely inhibits growth in the prevalence of PSL access, but it also may actually reduce the prevalence of PSL access if it induces employers that might otherwise add to their fringe benefit package not to do so. PSL preemption might also contribute to a decline in PSL access, if employment growth is greater in "business-friendly" states—states that have passed preemption laws—than in states with more stringent regulatory requirements. Differential employment growth could result from existing businesses' decisions to relocate, or from the location decisions of new enterprises.

Right-to-work laws, which prohibit workplace contracts that require employees who are not union members to contribute to the costs of union representation, are aimed at reducing union strength and are associated with reduced overall wage levels and with lower levels of employee benefits, including access to paid sick leave (24, 25). Therefore, right-to-work laws could contribute to diminished growth in, or to actual reductions in, the prevalence of PSL access.

The possible effects of PSL preemption and of RTW laws on PSL access, whether individually or in combination with PSL mandates, does not appear to have been studied. This paper presents an exploratory analysis of the three policies, viewing each of the three state-level policies as "treatments." It uses a

straightforward empirical approach and readily available public data sources.

Methods

Data

Data from several online sources were combined for this analysis. The outcome variable, PSL coverage, is taken from annual tables published by the US Department of Labor (26). These tables show, for 2009–2021, the percentage of civilian workers with access to PSL in each of 9 Census Divisions. The Divisions contain from 3 to 9 states (27). Thus, there are 9 (divisions) times 13 (years) = 117 observations in the data file.

Online sources provided, for each state, the implementation year for PSL (floor) mandate laws (21), PSL (ceiling) preemption laws (28), and RTW laws (29). For each policy variable (mandate, ceiling, and RTW, respectively) a series of annual indicators was created, where a value of 1 indicates that the policy was present that year, and a zero indicates that the policy was not present. These state-by-year policy variables were aggregated to the Census Division level, using as weights the size of each state's civilian employed population in the relevant year. Population counts came from the Bureau of Labor Statistics' Local Area Unemployment series, which includes counts of employed people (30). The entire data set is included in the Supplementary material.

Analysis

The data used in this study do not support the widely used DD approach to inferring causality with non-experimental data. Each observation used here represents a collection of states for which, in any given year, some may have, while others have not, implemented one or more of the policies studied. For each division-year observation the three policy variables fall into the 0, 1 interval (inclusive). Thus, there is not an evident "pre-treatment" nor an evident "post-treatment" period for any of the Census divisions. This problem is further complicated by the fact that three distinct treatments are considered, and that in any given year, several combinations of the treatments may have been adopted.

As a consequence, a simple weighted least-squares regression approach is used here. The outcome, PSL coverage, is regressed on three measures of district-level presence of each of the three policies, that is PPSLMAN (proportion covered by a mandate), PPSLPRE (proportion covered by ceiling preemption), and PRTW (proportion covered by a RTW law). In alternate specifications, all possible combinations (i.e., interactions) of the mandate and ceiling variables are entered, as is an interaction between ceiling and RTW In all cases, the treatments were coded to begin in the calendar year after the respective law was implemented, to allow for lags in employers' responses to the policy changes. All regressions also include fixed effects for Census divisions as well as calendar year dummy variables. It should be noted that in appearance, this regression looks just like the two-way fixed effects regression widely used in evaluation research; it is the data, and not the estimating equation, that departs from the usual DD setup. Division-year observations are weighted by the size of the civilian employed population in the division that year. Inference is based on confidence sets and p-values obtained using the wild cluster bootstrap algorithm (31). With the highly aggregated data used here, most of the within-division (i.e., between-state) variability in both outcomes and regressors is lost, and there is a precipitous loss of degrees of freedom. As a consequence, the potential for obtaining statistically significant results is greatly diminished. Therefore, the analysis must be viewed as exploratory, and the results as suggestive rather than as definitive.

Results

Figure 1 illustrates the spatial pattern of the three policy domains for 2021, in the form of a tile grid map. In this map policies are coded according to their year of implementation. Prior to 2009, just one state (Washington DC) had adopted a PSL mandate, but 21 states had already adopted a RTW law, mostly during the 1940s and 1950s. Additional PSL mandates began to appear in 2012, while PSL preemption first appeared in 2012, with both types of policies spreading thereafter. Five more states also adopted RTW laws beginning in 2012. By 2021, a eight possible combinations of the three binary policy indicators occurred at least once.

Several sets of regression results are reported in Table 1. The first four regressions consider only one of the three state-level policies, in turn; the remaining models address PSL mandates in combination with the other two policies. Model (1) includes only the mandate treatment, and uses only years 2009–2017, the same years used in Maclean et al. (15). The point estimate of the PSL mandate effect suggests an increase in PSL coverage of 8.7 percentage points, a result reasonably close to that obtained in Maclean et al. (15), and well within the 95% confidence interval of the earlier paper's estimate. However, the confidence set for this estimate includes zero. Model (2) uses all available years −2009-2021—and finds a much larger, and statistically significant, mandate effect, a nearly 16 percentage point increase in PSL coverage. The large increase in mandate effects is most likely due to the fact that when adding the more recent years, several additional states have now implemented PSL mandates, and past research has shown that the effect of the mandate on coverage is largest in the first few years of its existence.

Models (3) and (4) investigate the effects on PSL coverage of ceiling preemption and of RTW laws, respectively, in each case

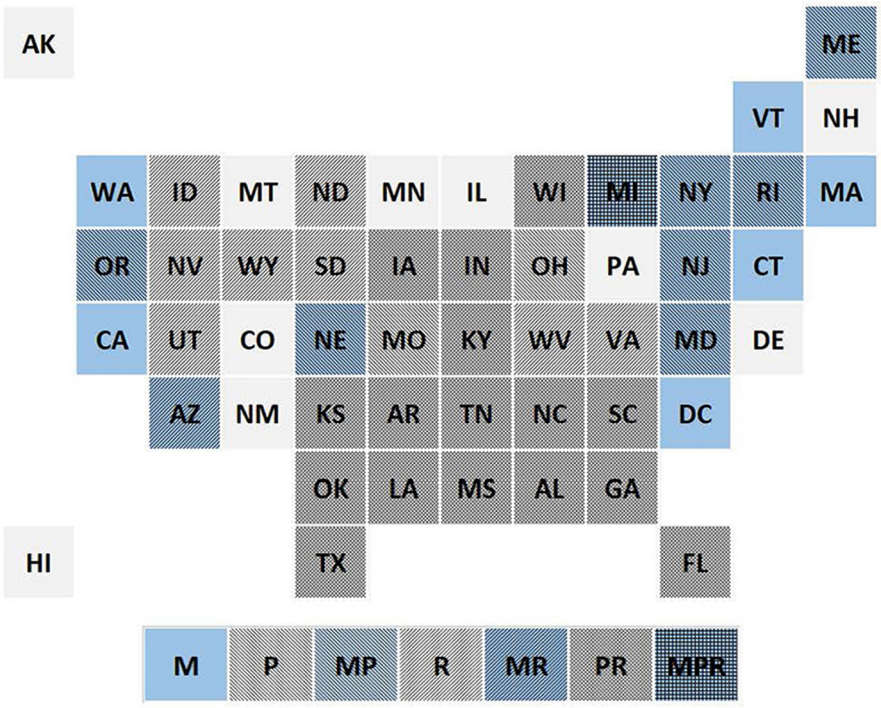


FIGURE 1
Tile grid map showing the pattern of PSL mandate, PSL preemption, and right-to-work laws, in 2021 (M, mandate; P, preemption; R, right-to-work).

TABLE 1 Estimated effects of policy variables on paid sick leave coverage, various specifications.

Variable	(1)[a]	(2)	(3)	(4)	(5)	(6)
Mandate	0.087	0.155*				
	[−0.078, 0.389]	[0.019, 0.331]				
Ceiling			−0.114			
			[−0.246, 0.026]			
Right to work (RTW)				−0.064		0.013
				[−0.881, 0.796]		[−0.163, 0.205]
Mandate, no ceiling					0.140	0.141
					[−0.101, 0.183]	[−0.105, 0.184]
Mandate with ceiling					−0.026	−0.004
					[−0.169, 0.282]	[−0.141, 0.305]
Ceiling, no mandate					−0.042	0.025
					[−0.076, 0.001]	[−0.075, 0.208]
Ceiling X RTW						−0.083
						[−0.301, 0.023]

Bootstrapped confidence sets in square brackets.
*Bootstrapped p-value < 0.05.
[a]Equation (1) uses data for 2009–2017; all other equations use data for 2009–2021.

without controlling for the other two policies. In both cases the regression coefficients have negative signs, suggesting that both ceiling preemption and RTW laws reduce workers' access to PSL. Both coefficients are, however, imprecisely estimated.

However, it is clear from Figure 1 that the three policies are not independent of each other, indicating that they should be considered jointly. Model (5) distinguishes the three possible combinations of mandate (or floor preemption) and

ceiling preemption. It appears that only when a mandate is not accompanied by ceiling preemption does it increase PSL coverage. Ceiling preemption in the absence of a mandate—that is, a statewide prohibition on governmental requirements that employers provide PSL—appears to reduce PSL coverage (in this case, with a borderline-significant *p*-value of 0.085).

Equation (6) contains interaction effects involving all three policy domains. This regression suggests that a mandate not accompanied by a ceiling has the largest positive effect on PSL coverage, while the combination of ceiling preemption and a RTW law has the largest negative effect on PSL coverage. The other variables included in (6) produce estimates very close to zero, although it must be acknowledged that zero is included in the confidence sets of all the estimated coefficients.

Discussion

The results reported here support those from two recent papers (15, 22) that found statistically significant increases in workers' access to PSL as a result of state-level PSL mandates. Using aggregated data from the Labor Department's Employee Compensation Survey (ECS) for 2009–2017, the mandate effect is close to the estimate reported in Maclean et al. (15), which used the same survey, in its original disaggregated form, for the same years. The comparability of my estimated PSL mandate to those reported in Callison and Pesko (22) cannot be determined due to non-overlapping years studied.

Using the full 13 years of data and a specification analogous to that in Maclean et al. (15), the effect of a PSL mandate on PSL coverage is positive and statistically significant. However, due to data limitations I am unable to use the specifications, and perform the sensitivity tests, and carry out other robustness checks that have become best practice for supporting claims of causality with non-experimental data. Yet the fact that my results are so close to those previously reported, despite these limitations, is reassuring. This analysis also displays the potential for using easily accessed online data to explore issues that have, to date, required expensive and burdensome procedures to make use of restricted-access data (to use the ECS data, a researcher must first be approved by the Labor Department and, upon approval, travel to one of a limited set of data enclaves to carry out the analysis).

The main point of this analysis, however, is that state adoption of PSL mandates has occurred along with the adoption of other policies—preemption of lower-level governments' ability to mandate PSL provision, and the adoption of right-to-laws—that are expected to have their own, possibly countervailing, consequences for PSL access. If PSL ceiling preemption and right-to-work laws are each considered as an individual "treatment," each appears to reduce workers' PSL access. In the models that examine each policy in isolation [i.e., (2), (3), and (4)] a PSL mandate appears to have a larger positive effect on access than either of other policies' negative effects. This is to be expected, since the mandates are targeted directly at expanding PSL access, whereas PSL preemption and right-to-work laws have only indirect consequences for the spread of PSL coverage. When all three policies are considered jointly [in equation (6)], a mandate with no ceiling, and a ceiling in combination with a RTW law, appear to have the largest consequences for PFL coverage. However, as already noted, data limitations dictate that these results be viewed as exploratory, and, at best, suggestive. They can, however, serve as a possible roadmap for follow-up research based on individual- or state-level data.

The present analysis also reminds us of a familiar shortcoming of observational studies, namely their need to deal with omitted-variables biases. In this case, a study that addresses whether states' PSL mandates produce increases in workers' access to PSL, but fails to account for either PSL preemption or RTW laws, appears to be subject to omitted variables biases in its estimate of policy impacts. The present study, of course, is not exempt from this problem, inasmuch as there could be additional factors that vary across states, that are correlated with any or all of the three policy variables used here, and that have their own effects on PSL coverage.

In view of the substantial body of evidence supporting the public health benefits of PSL, and the additional evidence that state PSL mandates lead to higher rates of PSL access, activists will presumably want to direct their efforts toward the further spread of these state-level mandates. It is concerning, however, that PSL preemption, which appears to hinder the growth of, and even reduce the prevalence of, PSL access, has been adopted in more states than PSL mandates have been. Faced with this situation, greater attention might be paid to encouraging adoption of a national PSL requirement. It is noteworthy that the U.S. is the only country, among 22 wealthy nations, that lacks a national PSL law (32). This, in turn, is likely to be just one of several factors explaining the fact that the U.S. continues to lag behind many other countries in various public health indicators, including mortality (33).

Data availability statement

The original contributions presented in the study are included in the Supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

DW was responsible for all aspects of the research.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.1003117/full#supplementary-material

References

1. Schneider D. Paid sick leave in Washington State: evidence on employee outcomes, 2016–2018. *Am J Pub Health.* (2020) 110:499–504. doi: 10.2105/AJPH.2019.305481

2. DeRigne L, Stoddard-Dare P, Quinn L. Workers without paid sick leave less likely to take time off for illness or injury compared to those with paid sick leave. *Health Aff.* (2016) 35:520–7. doi: 10.1377/hlthaff.2015.0965

3. Cook WK. Paid sick days and health care use: an analysis of the 2007 National Health Interview Survey data. *Am J Ind Med.* (2011) 54:771–9. doi: 10.1002/ajim.20988

4. Bhuyan SS, Wang Y, Bhatt J, Dismuke SE, Carlton EL, Gentry D, et al. Paid sick leave is associated with fewer ED visits among US private sector working adults. *Am J Emergency Med.* (2016) 34:784–9. doi: 10.1016/j.ajem.2015.12.089

5. Stoddard-Dare P, DeRigne L, Collins C, Quinn L. Retirement savings among US older adult male workers by paid sick leave, flexible work, and vacation benefit status. *Community Work Fam.* (2021) 24:357–73. doi: 10.1080/13668803.2019.1677557

6. Hawkins D, Zhu J. Decline in the rate of occupational injuries and illnesses following the implementation of a paid sick leave law in Connecticut. *Am J Industrial Med.* (2019) 62:859–73. doi: 10.1002/ajim.23028

7. Stearns J, White C. Can paid sick leave mandates reduce leave-taking? *Labor Econ.* (2018) 51:227–46. doi: 10.1016/j.labeco.2018.01.002

8. Asfaw A, Rosa R, Pana-Cryan R. Potential economic benefits of paid sick leave in reducing absenteeism related to the spread of influenza-like illness. *J Occup Env Med.* (2017) 59:822–9. doi: 10.1097/JOM.0000000000001076

9. Hsuan C, Ryan-Ibarra, DeBurgh K, Jacobson DM. Association of paid sick leave laws with foodborne illness rates. *Am J Prev Medicine.* (2017) 53:609–15. doi: 10.1016/j.amepre.2017.06.029

10. Hill HD. Paid sick leave and job stability. *Work Occup.* (2013) 40:143–73. doi: 10.1177/0730888413480893

11. Wething H. *Does Paid Sick Leave Affect Job Turnover? Working Paper Series. Washington Center for Equitable Growth.* (2022). Available online at: https://equitablegrowth.org/working-papers/does-paid-sick-leave-affect-job-turnover/ (accessed July 8, 2022).

12. Kim D. Paid sick leave and risks of all-cause and cause-specific mortality among adult workers in the USA. *Int J Env Res Pub Health.* (2017) 14:1247–56. doi: 10.3390/ijerph14101247

13. Wolf DA, Montez JK, Monnat SM. U.S. State preemption laws and working-age mortality. *Am J Prev Med.* (2022) 63:681–8. doi: 10.1016/ampre.2022.06.005

14. U.S. Department of Labor, Bureau of Labor Statistics. *Paid Sick Leave: What is Available to Workers? (Factsheet).* (2020). Available online at: https://www.bls.gov/ncs/ebs/factsheet/paid-sick-leave.htm (accessed June 6, 2022).

15. Maclean C, Pichler S, Ziebarth NR. *Mandated Sick Pay: Coverage, Utilization, and Welfare Effects. Washington Center for Equitable Growth Working Paper Series.* (2022). Available online at: https://equitablegrowth.org/working-papers/mandated-sick-pay-coverage-utilization-and-welfare-effects/ (accessed January 1, 2022).

16. Chen Y-P. The growth of fringe benefits: implications for social security. *Mon Labor Rev.* (1981) 104:3–10.

17. Licht W. Fringe benefits: a review essay on the American workplace. *Int Labor Work Class Hist.* (1998) 53:164–78. doi: 10.1017/S0147547900013715

18. US Department of Labor, Bureau of Labor Statistics. *Employee Benefits in Industry: A Pilot Survey.* (1980). Report 615. Available online at: https://www.bls.gov/ncs/ebs/sp/ebbl0040.pdf (accessed June 9, 2022).

19. US Department of Labor, Bureau of Labor Statistics. *Paid Sick Leave: What is Available to Workers?* (2021). Available online at: https://www.bls.gov/ncs/ebs/factsheet/paid-sick-leave.htm

20. Pomeranz JL, Silver D, Lieff SA, Pagán JA. State paid sick leave and paid sick-leave preemption laws across 50 U.S. States, 2009–2020. *Am J Prev Med.* (2022) 62:688–95. doi: 10.1016/j.amepre.2021.11.018

21. National Partnership for Women & Families, *"Paid Sick Days Statutes".* (2022). Available online at: https://www.nationalpartnership.org/our-work/resources/economic-justice/paid-sick-days/paid-sick-days-statutes.pdf (accessed August 5, 2022).

22. Callison K, Pesko MF. The effect of paid sick leave mandates on coverage, work absences, and presenteeism. *J Hum Resour.* (2022) 57:1178–208. doi: 10.3368/jhr.57.4.1017-9124R2

23. Pomeranz JL, Pertschuk M. State preemption: a significant and quiet threat to public health in the United States. *Am J Public Health.* (2017):107:900–2. doi: 10.2105/AJPH.2017.303756

24. Gould E, Shierholz H. *The Compensation Penalty of "Right-to-Work" Laws. Economic Policy Institute, EPI Briefing Paper #299.* (2011). Available online at: https://www.epi.org/publication/bp299/ (accessed June 8, 2022).

25. Van Heuvelen T. The right to work, power resources, and economic inequality. *AJS.* (2020) 125:1255–302. doi: 10.1086/708067

26. U.S. Bureau of Labor Statistics, National Compensation Survey (NCS). *Employee Benefits in the United States.* Available online at: www.bls.gov/ncs (accessed November 21, 2021).

27. U.S. Census Bureau. *Census Bureau Regions and Divisions with State FIPS Codes.* Available online at: https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf (accessed January 8, 2022).

28. Economic Policy Institute. *Workers' Rights Preemption in the U.S.* Available online at: https://www.epi.org/preemption-map/ (accessed June 8, 2022).

29. Ballotpedia. *Right-to-Work Laws.* Available online at: https://ballotpedia.org/Right-to-work_laws (accessed December 11, 2021).

30. U.S. Bureau of Labor Statistics, Local Area Unemployment Statistics. Available online at: https://www.bls.gov/lau/ (accessed July 10, 2021).

31. Roodman D, MacKinnon JG, Nielsen MØ, Webb MD. Fast and wild: bootstrap inference in Stata using boottest. *Stata J.* (2019) 19:4–60. doi: 10.1177/1536867X19830877

32. Rho HJ, Fremstad S, Gaby-Biegel J. *Contagion Nation 2020: United States Still the Only Wealthy Nation Without Paid Sick Leave.* Washington, DC: Center for Economic Policy Research (2020). Available online at: https://cepr.net/wp-content/uploads/2020/03/2020-03-Contagion-Nation-2020.pdf

33. Mehta NK, Myrskylä M. An introduction to the supplemental issue on why does health in the US continue to lag behind? *J Gerontol B.* (2022) 77 (Supp 2):S113–6. doi: 10.1093/geronb/gbac050

frontiers | Frontiers in Public Health

TYPE Brief Research Report
PUBLISHED 22 November 2022
DOI 10.3389/fpubh.2022.1007053

OPEN ACCESS

EDITED BY
Shannon Monnat,
Syracuse University, United States

REVIEWED BY
Courtney Boen,
University of Pennsylvania,
United States
Alicia Riley,
University of California, Santa Cruz,
United States

*CORRESPONDENCE
Tyson H. Brown
tyson.brown@duke.edu

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 29 July 2022
ACCEPTED 20 October 2022
PUBLISHED 22 November 2022

CITATION
Brown TH, Kamis C and Homan P (2022) Empirical evidence on structural racism as a driver of racial inequities in COVID-19 mortality.
Front. Public Health 10:1007053.
doi: 10.3389/fpubh.2022.1007053

Empirical evidence on structural racism as a driver of racial inequities in COVID-19 mortality

Tyson H. Brown[1]*, Christina Kamis[2] and Patricia Homan[3]

[1]Department of Sociology, Duke University, Durham, NC, United States, [2]Center for Demography of Health and Aging, University of Wisconsin-Madison, Madison, WI, United States, [3]Department of Sociology, Florida State University, Tallahassee, FL, United States

Objective: This study contributes to the literature by empirically testing the extent to which place-based structural racism is a driver of state-level racial inequalities in COVID-19 mortality using theoretically-informed, innovative approaches.

Methods: CDC data are used to measure cumulative COVID-19 death rates between January 2020 and August 2022. The outcome measure is a state-level Black-White (B/W) ratio of age-adjusted death rates. We use state-level 2019 administrative data on previously validated indicators of structural racism spanning educational, economic, political, criminal-legal and housing to identify a novel, multi-sectoral latent measure of structural racism (CFI = 0.982, TLI = 0.968, and RMSEA = 0.044). We map B/W inequalities in COVID-19 mortality as well as the latent measure of structural racism in order to understand their geographic distribution across U.S. states. Finally, we use regression analyses to estimate the extent to which structural racism contributes to Black-White inequalities in COVID-19 mortality, net of potential confounders.

Results: Results reveal substantial state-level variation in the B/W ratio of COVID-19 death rates and structural racism. Notably, regression estimates indicate that the relationship between the structural racism and B/W inequality in COVID-19 mortality is positive and statistically significant ($p < 0.001$), both in the bivariate model (adjusted $R^2 = 0.37$) and net of the covariates (adjusted $R^2 = 0.54$). For example, whereas states with a structural racism value 2 standard deviation *below* the mean have a B/W ratio of approximately 1.12, states with a structural racism value 2 standard deviation *above* the mean have a ratio of just above 2.0.

Discussion: Findings suggest that efficacious health equity solutions will require bold policies that dismantle structural racism across numerous societal domains.

KEYWORDS

structural racism, COVID-19 mortality, geographic inequality, racial inequality, health and mortality, measurement

Introduction

More than one million Americans have died from COVID-19. Notably, the impact of the pandemic has been unequally distributed across the color line. Racial inequities in COVID-19 mortality are well documented, with people racialized as Black experiencing much higher mortality rates than their White counterparts (1). Indeed, as of October 2022, the cumulative age-adjusted mortality rate for Blacks is 63% higher than it is for Whites (2). Consequently, Black Americans have experienced especially high levels of pandemic-related excess death (3, 4). The well-established disproportionate impact of the pandemic among Black the community has led to a great deal of discussion about the causes of these inequities.

Much of the scientific literature on Black-White COVID-19 mortality inequities has focused on the impact of *proximal causes,* such as racial inequalities in underlying health conditions, health care, and socioeconomic resources. A growing body of research, however, points to the role of upstream "causes of the causes" (5–7) undergirding the unequal toll of the pandemic along racial lines. In particular, numerous scholars have hypothesized that racial inequalities in COVID-19 mortality are driven by structural racism–i.e., a multi-sectoral, interrelated, system of racial oppression and exclusion from power, resources, opportunities, and well-being (8–11). This conceptualization aligns with accumulating evidence that discriminatory environments undermine the health of minoritized populations and contribute to racialized health outcomes (12–17). Theory highlights how structural racism indirectly harms the health of Black people because it leads to unequal access to salubrious resources and exposure to health risks (11, 18, 19). In the context of the COVID-19 pandemic, structural racism is thought to be an upstream cause of the downstream proximal causes (e.g., racial inequalities in underlying health conditions, economic and social deprivation, toxic living and working conditions, political exclusion, exposure to stressors, constrained autonomy and freedom, and inadequate health care) of Black-White inequities in COVID-19 mortality (7, 8, 20–22). Although a plethora of conceptual essays have hypothesized that structural racism is a driver of Black-White inequality in COVID-19 mortality (20, 23, 24), very few empirical studies have tested this proposition.

Robust empirical evidence of a relationship between areal structural racism and racial inequality in COVID-19 mortality would require at least three conditions:

1) geographic variation in racial inequalities in COVID-19 mortality.
2) geographic variation in structural racism.
3) a statistically significant relationship between structural racism and racial inequalities in COVID-19 mortality, net of likely confounders.

Below we summarize the evidence base for these three conditions, with a focus on limitations in prior research and how this study uses innovative approaches to improve our understanding of the extent to which structural racism is a driver of Black-White inequities in rates of COVID-19 mortality.

With respect to the first condition, prior research suggests that there is substantial variation in racial inequalities in COVID-19 mortality at both the county and state levels (21, 25–30). U.S. states are a particularly important geographic unit of analysis because, as Siegel and colleagues (2022) note, "*Understanding racial disparities at the state level is imperative because states have the primary responsibility for implementing policies related to the prevention, control, and response to COVID-19 and therefore are directly responsible for the emergence of, and amelioration of, racial disparities related to COVID-19.*" (30). Only a handful of quantitative studies have examined state-level variation in Black-White inequalities in COVID-19 mortality rates, and they are limited in several respects. For example, studies have often used crude death rates rather than rates that are age-adjusted (25, 29). Relying on crude death rates is problematic given the greater COVID-19 mortality risk among older adults in tandem with state differences in age distributions, as well as the younger population age profiles among Black Americans relative to their White counterparts. The few studies that have adjusted for age have often used indirect age standardization (24, 25), which is an inferior approach relative to direct age standardization because estimates based on indirect standardization are imprecise and are often not comparable across states (2). We are aware of only one published study on the topic that uses direct age standardization; findings show that not adjusting for age leads to severe underestimation of Black-White inequalities in COVID-19 mortality (30).

There is also growing evidence of state-level variation in structural racism. In fact, several studies have shown that indicators of structural racism—operationalized as Black-White inequities in societal domains such as housing, education, economics, politics, and the criminal-legal system—vary considerably across states, with levels of structural racism being particularly high in the Midwest and Northeast (31–34). These findings are consistent with the view that states are racialized institutional actors that shape the discriminatory, inequitable distribution of a plethora of social determinants of health along racial lines (35).

Regarding the third condition, a recent empirical study by Siegel and colleagues (2022) is the only one we are aware of that explored the association between state-level structural racism and Black-White inequities in COVID-19 mortality. Consistent with theory and hypotheses from a number of conceptual commentaries (7, 8), findings indicated that higher levels of structural racism—across multiple domains of society—were predictive of larger Black-White inequities (30). This was a very insightful contribution to the literature, yet the

study had several limitations and there remain important gaps in our understanding of the extent to which structural racism is a driver of racial inequities in COVID-19 mortality. First, the Siegel et al. study includes information on deaths due to COVID-19 only through November of 2020—and thus does not capture the vast majority of deaths attributed to COVID-19 throughout the pandemic as it has continued to the present. Second, the study by Siegel and colleagues relies on bivariate associations that do not account for potential confounders of the relationship between structural racism and mortality due to COVID-19 (30). Third, the study uses a summative index of structural racism across societal domains rather than a latent variable approach, which has a number advantages for measuring structural racism (described below).

We aim to extend prior research and address these gaps in the literature by using a theoretically-informed, innovative approach to measuring state-level structural racism and its impact on racial inequalities in COVID-19 mortality. Specifically, we use up-to-date data on (directly) age-standardized COVID-19 deaths (through August 20, 2022), adjust for potential confounders, and develop a novel, multi-sectoral latent measure of structural racism. This latent variable approach has several advantages including 1) capturing the multifaceted, interconnected and systemic nature of the complex and often hidden phenomena of structural racism, 2) allowing for variance in factor loadings (rather than assuming monolithic weights for each of the observed indicators), 3) permitting covariances specified between observed indicator variables, and 4) minimizing measurement error (14, 35–37). Collectively, findings suggest that these approaches have considerable utility for population health research, and that state-level structural racism is a driver of place-based Black-White inequalities in COVID-19 mortality. This is consistent with a growing literature pointing to population health as a mirror reflecting societal arrangements.

Methods

Age-adjusted mortality rates

CDC WONDER data are used to measure racial inequality in cumulative COVID-19 death rates between January 1st 2020 and August 20th 2022. The outcome measure is a state-level (Non-Hispanic) Black-White (B/W) ratio of age-adjusted death rates (AADR), which are calculated using the direct method[1]. CDC WONDER calculates age-adjusted death rates using direct standardization with the "2000 U.S. standard" as the standard population (for more information see CDC WONDER data documentation) (41). Age-adjusted death rates are preferable over crude death rates (CDR) because age is linked to COVID-19 mortality risk and because there are racial differences in the age profiles of the population. Consistent with other studies on state-level structural racism, this study excludes 13 states, producing a sample of 37 U.S. States (30, 31). The 13 states excluded have insufficient information on the state's Black population due to a low proportion of Black residents (<4.6%) and/or a low total population of Black residents (<50 k residents). The 37 states included in the study represent 99% of the U.S. Black population and 93% of the U.S. white population.

State-level structural racism indicators and latent scale

Consistent with research noting that U.S. states are racialized institutional actors shaping population health, and that structural racism involves multiple, interconnected societal domains (9, 10, 32, 35, 42), we utilize state-level 2019 administrative data on seven indicators of structural racism spanning educational, economic, political, criminal-legal and housing sectors. The indicators include: W/B ratios of Bachelor's degree, B/W ratios of poverty, W/B ratios of homeownership, B/W ratios of unemployment, W/B ratios of voting rates (in 2016 election), B/W ratios of incarceration, and the dissimilarity index of racial residential segregation (calculated at the state-level). A majority of these measures are derived from the U.S. Census Bureau's Current Population Study (CPS), with the exception of the measures of state-level residential segregation (data from America's Health Ranking) and incarceration (data from Bureau of Justice Statistics). Additionally, total population values were gathered from the American Community Survey 1-year estimates and used in the calculation of incarceration rates. Importantly, these seven indicators have been developed and validated in prior research (30, 32, 34, 35).

We use these validated measures to identify a novel, multi-sectoral latent measure of structural racism. Utilizing a latent measure of structural racism aligns with race theories positing that structural racism is systemic and often unobserved. We use

1 In assessing the quality of COVID, there is some evidence of unevenness in reporting COVID deaths across states (38, 39). One of the benefits of using a ratio measure of B:W COVID mortality (rather than overall or single race-specific) for our outcome is that even if reporting irregularities are systematic across states, the same "noise" may be present for both Black and White deaths within the state, in which case the validity of the measure of racial inequities in COVID would be unbiased. Furthermore, to the extent that there are racial differences in reporting, it would likely be under-reporting of Black deaths which would lead to conservative estimates of inequities (39, 40). Thus, we are unaware of any evidence to suggest that our findings of a relationship between structural racism and B-W COVID mortality inequities are biased due to data irregularities.

TABLE 1 Descriptive statistics of U.S. States ($N = 37$).

	Mean (or %)	SD	Range	Source
Cumulative age-adjusted death rates (AADR; 1/2020-8/20/2022)				
B/W AADR	1.59	0.27	[1.14, 2.08]	CDC Wonder
State-level structural racism indicators (2019)				
B/W incarceration rates	5.74	2.58	[2.51, 12.26]	BJS; ACS 1-year estimate; Author's Calculations
W/B college degree completion	1.74	0.38	[1.01, 2.96]	CPS ASEC; Author's Calculations
B/W unemployment rates	2.36	0.87	[0.77, 4.77]	CPS ASEC; Author's Calculations
B/W poverty rates	2.70	0.87	[1.12, 4.97]	CPS ASEC; Author's Calculations
W/B homeownership rates	1.93	0.53	[1.35, 3.62]	CPS ASEC; Author's Calculations
W/B voting rates	1.14	0.24	[0.86, 1.90]	CPS Voting Supplement; Author's Calculations
B/W segregation	57.57	8.28	[42.00, 72.00]	America's Health Ranking
State-level structural racism (2019)				
Latent structural racism	0.00	0.41	[-0.51, 1.11]	
Covariates				
Logged total population (2019)	15.61	0.82	[13.79, 17.49]	ACS 1-year estimate
Percentage NHB (2019)	15.53%		[4.59%, 38.58%]	CPS ASEC
Gini (2019)	0.47	0.02	[0.44, 0.51]	ACS 1-year estimate
Percentage below the poverty line (2019)	11.75%		[7.40%, 19.57%]	CPS ASEC
Region				U.S. Census
Northeast	16.22%			
Midwest	27.03%			
South	43.24%			
West	13.51%			

confirmatory factor analysis (CFA) to estimate a series of latent constructs with varying specifications. We first examine a model in which each structural racism dimension is loaded onto a single factor. We then allow for errors to be correlated for several dimensions in subsequent models, based on an assessment of the correlation matrix and driven by theoretical considerations. Fit was assessed using chi-square, BIC, RMSEA, CFI, and TLI.

The first model, which includes each structural racism indicator loaded onto a single factor with no correlated errors, had a moderate fit. The Chi-square was non-significant, but the RMSEA was over .05 and the CFI/TLI were both below .9. Permitting the error terms for the inequity in incarceration and inequity in unemployment to correlate improved fit, but the RMSEA was still over .05 and the TLI was still under .9. Adding an additional term that allowed for the error terms for the inequity in homeownership and inequity in voting to correlate had a much better fit, (CFI = 0.982, TLI = 0.968, and RMSEA = 0.044). Additional specifications were considered (such as allowing all errors for economic measures to be correlated), but they did not produce substantive changes in fit and had higher BIC values, therefore we proceed with the model that includes each structural racism indicator loaded onto a single factor with correlated errors between incarceration and employment inequities and between voting and homeownership inequities. See Supplementary Figure S1 for a diagram of our measurement model with factor loadings and correlated errors. We note that analyses using the latent variable produced by the base model without any correlated errors produced similar results to those presented here, despite it's relatively worse model fit[2]. In addition to the latent measure, we considered a composite index that standardized and summed each of the individual

2 We have assessed additional models that specified the relationship between structural racism and B/W disparities in cumulative COVID-19 AADR as quadratic and one that include structural racism as a categorical measure (quartiles of structural racism). The linear model (our current final model) resulted in a better fit for the data (lower BIC), providing support for a linear relationship.

indicators of structural racism. However, the latent variable model provided a better fit and a higher adjusted R-squared, indicating that it explained 24% more variation in our outcome (see the Supplementary Table S1 for additional details). While using a latent structural racism variable is the best approach for this study, it is possible that alternative approaches to measuring structural racism would be appropriate in other cases. Ultimately, the measurement of structural racism should be informed by research questions, logic, spatial and temporal contexts, feasibility, and data availability and fit.

Covariates

To minimize the risk of biased estimates, this study accounts for a range of potential confounders. Consistent with prior studies, regression estimates control for several state-level factors, including: population size (logged), percentage of the population that is NH Black, Gini coefficient, poverty rate, and region (31, 32, 34, 35).

Analyses

We begin by mapping B/W inequalities in COVID-19 mortality as well as the latent measure of structural racism in order to understand their spatial distribution across U.S. states. Next, we link the latent structural racism measure to CDC COVID data, and use Ordinary Least Squared (OLS) regression analyses to estimate the relationship between structural racism and Black-White inequalities in COVID-19 mortality. Multivariable analyses adjust for the covariates described above.

Results

Table 1 provides descriptive statistics and information about the sources for each of the study variables. The average ratio of B/W COVID-19 mortality suggests that for U.S. states, there are more Black deaths than white deaths. There are also B/W inequities across all measures of structural racism that indicate a larger burden on Black populations. Figure 1 includes maps showing substantial state-level variation in the B/W ratio of age-adjusted COVID-19 death rates (Figure 1A) and structural racism (Figure 1B), respectively. B/W ratios of COVID-19 mortality range from 1.14–2.08, with the greatest inequalities in upper midwestern and northeastern states. This means that in all states, Black COVID-19 death rates were substantially higher than white death rates. Similarly, mapping the spatial distribution of structural racism reveals that, despite its ubiquity, it tends be especially elevated in midwestern and northeastern states. This is consistent with an emerging

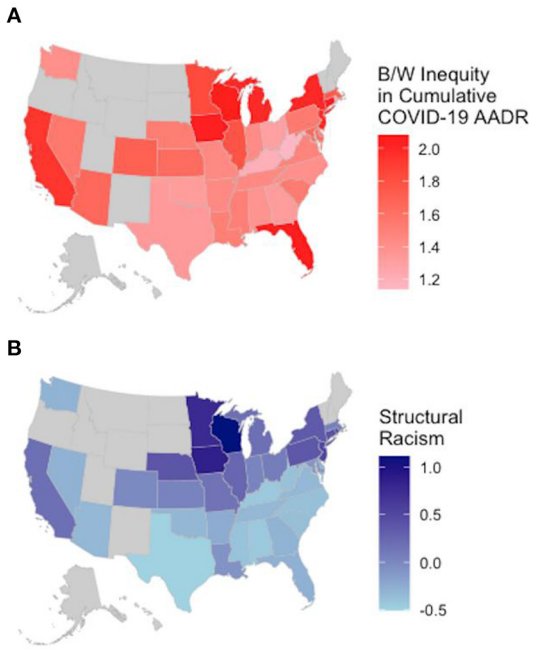


FIGURE 1
Geography of Black–White inequalities in cumulative COVID−19 mortality **(A)** and structural racism **(B)**, across U.S. States. **(A)** Shows B/W inequities in cumulative age–adjusted COVID−19 mortality (January 2020 to August 2022). **(B)** Shows a latent scale of state–level structural racism in 2019 spanning educational, economic, political, criminal–legal and housing domains.

body of literature on the spatial distribution of structural racism across U.S. states (30, 31, 35, 43–45). Although the historical and modern roots of state differences in structural racism are not fully understood, scholars have posited that elevated levels of contemporary structural racism—manifest in discriminatory institutional contexts—in the Midwestern and Northeastern states stem, in part, from institutionalized policies and practices of social control through racialized exclusion and subordination such as resource hoarding, redlining, racial covenants and discriminatory policing. These white supremacy tactics were increasingly deployed in response to the Great Migration because Northern Whites perceived the growing Black population as a threat (33, 35, 43, 44, 46).

Regression estimates in Table 2 indicate that the relationship between the structural racism and B/W inequality in COVID-19 mortality is statistically significant ($p < 0.001$), both in the bivariate model (Model 1; adjusted $R^2 = 0.37$) and net of the covariates (Model 2; adjusted $R^2 = 0.54$). Figure 2 graphically illustrates the predicted values of Black-White inequality in COVID-19 deaths as a function of structural racism using estimates from Table 2, Model 2 and holding all other covariates at their mean values. The figure shows that, higher levels of structural racism predict larger B/W ratios of COVID-19 death

TABLE 2 OLS regression predicting B/W inequities in cumulative COVID-19 Age-Adjusted Death Rates (AADR) by state-level structural racism ($N = 37$, U.S. States).

	Model 1 Coef (SE)	Model 2 Coef (SE)
Latent structural racism	0.402***	0.563***
	(0.086)	(0.136)
Logged total population (2019)		0.046
		(0.046)
Percentage NHB (2019)		0.247
		(0.471)
Gini (2019)		8.710*
		(3.833)
Percentage below the poverty line (2019)		−3.713*
		(1.603)
Region		
South (ref.)		
Northeast		−0.334
		(0.174)
Midwest		−0.095
		(0.132)
West		0.086
		(0.122)
Constant	1.591***	−2.748
	(0.035)	(1.367)
BIC	−4.327	1.254
Adjusted R–squared	0.369	0.537

*p < 0.05; **p < 0.01; ***p < 0.001.

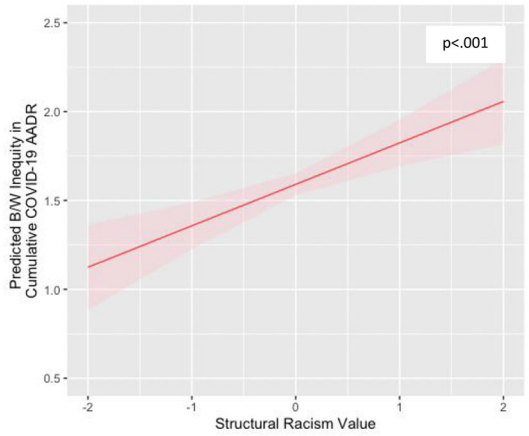


FIGURE 2
Relationship between state–level structural racism and Black-White inequities in COVID-19 mortality. Estimates account for potential confounders, including population size, percentage NH Black, Gini coefficient, poverty rate, and region.

rates. For example, whereas states with a structural racism value 2 standard deviations *below* the mean have a B/W ratio of approximately 1.12 (for every one White death, there are 1.12 Black deaths), states with a structural racism value 2 standard deviations *above* the mean have a ratio of just above 2.0 (for every one white death, there are just above 2 Black deaths). States with the average structural racism value have a B/W ratio of 1.6.

Discussion

Racial inequality in mortality is an enduring hallmark of the U.S. population health landscape. For as long as U.S. mortality data have been collected, Black people have experienced higher rates of mortality than their white counterparts (4, 47). Mortality rates during the COVID-19 pandemic are no exception. While the COVID-19 pandemic has led to significant excess deaths across all racial groups in the U.S., its deadly effects have not been spread over a level playing field (4, 8, 30). Numerous scholars have hypothesized that structural racism is the root cause of the disproportionately high rates of COVID-19 mortality among Black people (7, 8, 48).

The vast majority of studies on the role of structural racism in driving racial inequalities in COVID-19 mortality rates have been conceptual, leading to a dearth of empirical evidence on the topic. This study contributes to the literature by empirically testing the extent to which place-based structural racism undergirds state-level racial inequities in COVID-19 mortality using innovative approaches. Our theoretically-informed latent measure of structural racism allowed us to better capture the multifaceted, interconnected and systemic nature of racism, providing a more robust picture of its health consequences. In addition to an innovative approach to measuring structural racism, this study extends prior research by analyzing up-to-date mortality data (through June of 2022) and adjusting for potential confounding factors. We found that while all states had higher rates of Black COVID-19 mortality than white COVID-19 mortality, higher levels of structural racism were associated with larger Black-White inequalities in COVID-19 mortality. In other words, the more racism imbedded in state-level institutions the worse Black residents fared, relative to their white counterparts. Taken together, our findings provide empirical support for research theorizing a connection between racism and COVID-19 outcomes, and add to a growing literature documenting harmful health consequences of structural racism (22, 37, 49, 50).

Evidence that structural racism is a driver of racial inequalities in COVID-19 mortality is critical for shifting

the focus from untenable cultural deficit explanations—which blame the victims of White supremacy—toward the upstream root causes of the mortality inequities. Examining how unequal exposure to health-damaging social contexts, in general, and discriminatory environments in particular, aligns with prominent conceptual frameworks (e.g., Fundamental Cause Theory; Ecosocial Theory; the WHO Structural Determinants of Health framework) (11, 18, 51), as well as an emerging body of empirical research on the topic (21, 30). It is becoming more and more clear that Black-White inequities in population health reflect racialized societal arrangements across many sectors of society, including educational, economic, housing, political, and criminal-legal domains (9, 10, 42, 52).

As political, legal, administrative units, U.S. states play a key role in shaping the unequal distribution of social determinants of health (53–55). Moreover, findings from this study—in tandem with a nascent but growing body of research (30–32, 34, 35) —point to the importance of conceptualizing states as *racializing* institutional actors that shape population health. While structural racism is embedded in all states, results from this study reveal that states vary in their degrees of structurally racist contexts see also Siegel et al. (30). This is consistent with Bruch and colleagues' (57:163) contention that, "The state in which one resides has significant consequences for one's opportunities and life conditions and… for the structure of racial relations one must traverse" (56). Indeed, since the founding of the country states have been influential in sanctioning, exacerbating and alleviating racial oppression—from the historical roles they played with respect to policies on slavery, Jim Crow, and anti-miscegenation to their contemporary "race-neutral" policies that perpetuate racial domination such as voter disenfranchisement, gerrymandering, welfare state contraction and criminal sentencing laws (45, 57–59).

Our study has a number of limitations that point to fruitful avenues for future research. First, our study only contains state-level data. Although states are clearly a vital unit of analysis for understanding the mortality effects of structural racism, future research should seek to incorporate multilevel data to allow for the examination of individual-level exposures and outcomes, as well as structural racism at organizational, neighborhood, county, state and regional levels. Second, our data do not permit testing of the more proximal mechanisms connecting structural racism to COVID-19 deaths. To the extent that rich multilevel data become available, research should examine the theorized pathways through which racism is expected to increase risk of COVID-19 death, including: chronic health conditions, economic and social deprivation, toxic living and working conditions, political exclusion, inadequate health care, and psychosocial factors (e.g., social stressors, lack of autonomy, and stigma) (7, 11, 18, 30). Third, while we have employed a relatively comprehensive measure of structural racism across multiple institutional domains, it does not represent an exhaustive analysis of all the ways systemic racism shapes health. In addition to the institutional aspects of structural racism we examined, future research should also investigate the health effects of historical and contemporary discriminatory laws and policies, as well as anti-black cultural orientations and ideologies (60–62). Fourth, although this study is focused on understanding the dramatic Black-White inequities in COVID-19 deaths observed in the US, it also important for future research to examine race-specific COVID-19 death rates and whether there is evidence that elevated levels of structural racism are universally harmful. Studies examining other types of health outcomes have tended to find no effect of state-level structural racism among whites (31, 34, 35), but at least one has found evidence of a health benefit for whites (32). Finally, we focused on anti-Black structural racism because it has been a central and enduring feature of American society, but there is a need for future research to examine the impact of additional forms of structural racism on an array of racialized groups.

The COVID-19 Pandemic is shedding light on U.S. mortality inequities across the color line, leading to a growing understanding that structural racism is the root cause. Fundamental cause theory describes how societal forces (such as structural racism) shape the distribution of a multitude of health-relevant risks and resources and are therefore consistently linked to multiple disease outcomes through an array of mechanisms (18). Resources are flexible and can be leveraged to avoid disease even under changing circumstances, such as the COVID-19 pandemic. Thus, interventions to reduce racialized health inequities will be ineffective if they focus primarily on "proximal causes" of disease–which prove to be transient over time—rather than addressing structural racism as the more distal, fundamental cause. As new health threats emerge in the future—whether they are infectious diseases, environmental or climate related hazards, or even political-legal barriers to accessing necessary healthcare—we will continue to experience the same type of dramatic racial inequities we have seen during the COVID-19 pandemic unless we find ways to dismantle structural racism. Healthcare plays an important role in treating health problems and supporting population health, yet it is also critical to create social conditions that prevent (not just treat) health problems that disproportionately burden Black people in the US. Our study findings point to equity-promoting policies in social, economic, and political systems as necessary for creating conditions to achieve racial health equity. Research showing that Black-White inequalities in COVID-19 mortality (as well as other health outcomes) are a function of a multi-sectoral and reciprocal system of structural racism suggests that incremental policies that focus on a single domain are unlikely to substantially reduce racial inequalities (10, 18, 19, 30, 42, 50). Thus, efficacious health equity solutions will require bold policies that dismantle structural racism across numerous societal domains such as criminal justice reform, shoring up voting rights and eliminating felony disenfranchisement, implementing baby bonds to reduce the racial wealth gap and

a federal jobs guarantee to close employment and earnings gaps, and reforming the public education finance system to promote racial equity in schools (15, 63, 64).

While the COVID-19 pandemic is an unprecedented global public health emergency, the racial inequities that have emerged in the United States are following well-known and predictable patterns. People racialized as Black continue to bear a disproportionate burden of disease and death. This represents an enormous amount of unnecessary and unequal human suffering that demands redress.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Ethics statement

The studies involving human participants were reviewed and approved by Duke IRB. The patients/participants provided their written informed consent to participate in this study.

Author contributions

TB identified the research question, developed the research plan, provided input on the analyses, and took the lead role in writing the manuscript. CK led on the data analysis and visualizations, and both CK and PH contributed to writing the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research received support from grants P30 AG034424 (awarded to the Center for Population Health and Aging at Duke University by the National Institute on Aging), T32 AG00129 (awarded to the Center for Demography of Health and Aging at the University of Wisconsin-Madison by the National Institute on Aging), and 2R24AG045061-06 (awarded to the Network on Life Course Health Dynamics and Disparities in 21st Century America by the National Institute on Aging).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.1007053/full#supplementary-material

References

1. Elo IT, Luck A, Stokes AC, Hempstead K, Xie W, Preston SH. Evaluation of age patterns of COVID-19 mortality by race and ethnicity from March 2020 To October 2021 in the US. *JAMA Netw Open.* (2022) 5:e2212686. doi: 10.1001/jamanetworkopen.2022.12686

2. American Public Media. *Color of Coronavirus: COVID-19 Deaths Analyzed by Race and Ethnicity*. APM Research Lab. Available online at: https://www.apmresearchlab.org/covid/deaths-by-race (accessed January 11, 2022).

3. Chen R, Aschmann HE, Chen YH, Glymour MM, Bibbins-Domingo K, Stokes AC, et al. Racial and ethnic disparities in estimated excess mortality from external causes in the US, March to December 2020. *JAMA Intern Med.* (2022) 182:776–8. doi: 10.1001/jamainternmed.2022.1461

4. Wrigley-Field E. US racial inequality may be as deadly as COVID-19. *Proc Natl Acad Sci.* (2020) 117:21854–6. doi: 10.1073/pnas.2014750117

5. Bambra C, Smith KE, Pearce J. Scaling up: the politics of health and place. *Soc Sci Med.* (2019) 232:36–42. doi: 10.1016/j.socscimed.2019.04.036

6. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. *Public Health Rep.* (2014) 129(Suppl. 2):19–31. doi: 10.1177/00333549141291S206

7. Garcia MA, Homan PA, García C, Brown TH. The color of COVID-19: structural racism and the disproportionate impact of the pandemic on older black and latinx adults. *J Gerontol Ser B.* (2021) 76:e75–80. doi: 10.1093/geronb/gbaa114

8. Bailey ZD, Moon JR. Racism and the political economy of COVID-19: will we continue to resurrect the past? *J Health Polit Policy Law.* (2020) 45:937–50. doi: 10.1215/03616878-8641481

9. Bonilla-Silva E. Rethinking racism: toward a structural interpretation. *Am Sociol Rev.* (1997) 62:465–80. doi: 10.2307/2657316

10. Gee GC, Ford CL. Structural racism and health. inequities. *Bois Rev Soc Sci Res Race.* (2011) 8:115–32. doi: 10.1017/S1742058X11000130

11. Krieger N. Discrimination and health inequities. *Int J Health Serv Plan Adm Eval.* (2014) 44:643–710. doi: 10.2190/HS.44.4.b

12. Brown TH, Homan PA. *The Future of Social Determinants of Health: Looking Upstream to Structural Drivers*. (2023).

13. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *Lancet Lond Engl.* (2017) 389:1453–63. doi: 10.1016/S0140-6736(17)30569-X

14. Hardeman RR, Homan PA, Chantarat T, Davis BA, Brown TH. Improving the measurement of structural racism to achieve antiracist health policy. *Health Aff (Millwood).* (2022) 41:179–86. doi: 10.1377/hlthaff.2021.01489

15. Braveman PA, Arkin E, Proctor D, Kauh T, Holm N. Systemic and structural racism: definitions, examples, health damages, and approaches to dismantling. *Health Aff (Millwood).* (2022) 41:171–8. doi: 10.1377/hlthaff.2021.01394

16. Lantz P. *The Tenets of Critical Race Theory Have a Long-Standing and Important Role in Population Health Science*. The Millbank Quarterly. Available online at: https://www.milbank.org/quarterly/opinions/the-tenets-of-critical-race-theory-have-a-long-standing-and-important-role-in-population-health-science/ (accessed June 10, 2022).

17. HIng A. The right to vote, the right to health: voter suppression as a determinant of racial health disparities. *J Health Disparities Res Pract.* (2020) 12:5.

18. Phelan JC, Link BG. Is racism a fundamental cause of inequalities in health? *Annu Rev Sociol.* (2015) 41:311–30. doi: 10.1146/annurev-soc-073014-112305

19. Williams DR, Lawrence JA, Davis BA. Racism and health: evidence and needed research. *Annu Rev Public Health.* (2019) 40:105–25. doi: 10.1146/annurev-publhealth-040218-043750

20. Johnson-Agbakwu CE, Ali NS, Oxford CM, Wingo S, Manin E, Coonrod DV. Racism, COVID-19, and health inequity in the USA: a call to action. *J Racial Ethn Health Disparities.* (2022) 9:52–8. doi: 10.1007/s40615-020-00928-y

21. Tan SB. deSouza P, Raifman M. Structural racism and COVID-19 in the USA: a County-level empirical analysis. *J Racial Ethn Health Disparities.* (2022) 9:236–46. doi: 10.1007/s40615-020-00948-8

22. Michener J, LeBrón AMW. Racism, health, and politics: advancing interdisciplinary knowledge. *J Health Polit Policy Law.* (2022) 47:111–30. doi: 10.1215/03616878-9517149

23. Abraham P, Williams E, Bishay AE, Farah I, Tamayo-Murillo D, Newton IG. The roots of structural racism in the United States and their manifestations during the COVID-19 pandemic. *Acad Radiol.* (2021) 28:893–902. doi: 10.1016/j.acra.2021.03.025

24. Yearby R, Mohapatra S. Law, structural racism, and the COVID-19 pandemic. *J Law Biosci.* (2020) 7:lsaa036. doi: 10.1093/jlb/lsaa036

25. Boserup B, McKenney M, Elkbuli A. Disproportionate impact of COVID-19 pandemic on racial and ethnic minorities. *Am Surg.* (2020) 86:1615–22. doi: 10.1177/0003134820973356

26. Cheng KJG, Sun Y, Monnat SM. COVID-19 Death rates are higher in rural counties with larger shares of blacks and hispanics. *J Rural Health.* (2020) 36:602–8. doi: 10.1111/jrh.12511

27. Goldstein JR, Atherwood S. Improved measurement of racial/ethnic disparities in COVID-19 mortality in the United States. *medRxiv.* (2020):2020.05.21.20109116. doi: 10.1101/2020.05.21.20109116

28. Gross CP, Essien UR, Pasha S, Gross JR, Wang S. yi, Nunez-Smith M. Racial and ethnic disparities in population-level Covid-19 mortality. *J Gen Intern Med.* (2020) 35:3097–9. doi: 10.1007/s11606-020-06081-w

29. Parcha V, Malla G, Suri SS, Kalra R, Heindl B, Berra L, et al. Geographic variation in racial disparities in health and Coronavirus Disease-2019 (COVID-19) mortality. *Mayo Clin Proc Innov Qual Outcomes.* (2020) 4:703–16. doi: 10.1016/j.mayocpiqo.2020.09.005

30. Siegel M, Critchfield-Jain I, Boykin M, Owens A. Actual racial/ethnic disparities in COVID-19 mortality for the non-hispanic black compared to non-hispanic white population in 35 US states and their association with structural racism. *J Racial Ethn Health Disparities.* (2022) 9:886–98. doi: 10.1007/s40615-021-01028-1

31. Homan PA, Brown TH. Sick and tired of being excluded: structural racism in disenfranchisement as A threat to population health equity. *Health Aff (Millwood).* (2022) 41:219–27. doi: 10.1377/hlthaff.2021.01414

32. Lukachko A, Hatzenbuehler ML, Keyes KM. Structural racism and myocardial infarction in the United States. *Soc Sci Med 1982.* (2014) 103:42–50. doi: 10.1016/j.socscimed.2013.07.021

33. Mesic A, Franklin L, Cansever A, Potter F, Sharma A, Knopov A, et al. The relationship between structural racism and black-white disparities in fatal police shootings at the state level. *J Natl Med Assoc.* (2018) 110:106–16. doi: 10.1016/j.jnma.2017.12.002

34. Wallace M, Crear-Perry J, Richardson L, Tarver M, Theall K. Separate and unequal: structural racism and infant mortality in the US. *Health Place.* (2017) 45:140–4. doi: 10.1016/j.healthplace.2017.03.012

35. Brown T, Homan P. Structural racism and health stratification in the U.S: connecting theory to measurement. *SocArXiv.* (2022). doi: 10.31235/osf.io/3eacp

36. Chantarat T, Mentzer KM, Van Riper DC, Hardeman RR. Where are the labor markets?: examining the association between structural racism in labor markets and infant birth weight. *Health Place.* (2022) 74:102742. doi: 10.1016/j.healthplace.2022.102742

37. Adkins-Jackson PB, Chantarat T, Bailey ZD, Ponce NA. Measuring structural racism: a guide for epidemiologists and other health researchers. *Am J Epidemiol.* (2021) 191:539–47. doi: 10.1093/aje/kwab239

38. Krieger N, Testa C, Hanage WP, Chen JT. US racial and ethnic data for COVID-19 cases: still missing in action. *Lancet.* (2020) 396:e81. doi: 10.1016/S0140-6736(20)32220-0

39. Krieger N. Structural racism, health inequities, and the two-edged sword of data: structural problems require structural solutions. *Front Public Health.* (2021) 9:655447. doi: 10.3389/fpubh.2021.655447

40. Cowger TL, Davis BA, Etkins OS, Makofane K, Lawrence JA, Bassett MT, et al. Comparison of weighted and unweighted population data to assess inequities in Coronavirus Disease 2019 deaths by race/ethnicity reported by the US centers for disease control and prevention. *JAMA Netw Open.* (2020) 3:e2016933. doi: 10.1001/jamanetworkopen.2020.16933

41. CDC WONDER. Centers for Disease Control and Prevention, National Center for Health Statistics. National Vital Statistics System, Provisional Mortality on CDC WONDER Online Database. (2022). (Data are from the final Multiple Cause of Death Files, 2018-2020 and from provisional data for years 2021-2022, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.). Available online at: https://wonder.cdc.gov/wonder/help/mcd-provisional.html (accessed June 10, 2022).

42. Reskin B. The race discrimination system. *Annu Rev Sociol.* (2012) 38:17–35. doi: 10.1146/annurev-soc-071811-145508

43. Massey DS, Denton NA. *American Apartheid: Segregation and the Making of the Underclass. Later Printing edition*. Cambridge, Mass: Harvard University Press (1993). p. 312.

44. Muller C. Northward migration and the rise of racial disparity in American Incarceration, 1880–1950. *Am J Sociol.* (2012) 118:281–326. doi: 10.1086/666384

45. Uggen C, Larson, R, Shannon S, Pulido-Nava A. *Locked out 2020 Estimates of People Denied Voting Rights Due to A Felony Conviction. Sentencing Proj.* (2020). Available online at: https://www.sentencingproject.org/publications/locked-out-2020-estimates-of-people-denied-voting-rights-due-to-a-felony-conviction/ (accessed March 22, 2022).

46. Blalock HM. *Toward a Theory of Minority-group*. First Edition. New York, NY: Wiley. (1967). p. 227.

47. Du Bois WEB. *The Philadelphia Negro: A Social Study*. Philadelphia: University of Pennsylvania. (1899).

48. Laster Pirtle WN. Racial capitalism: a fundamental cause of novel Coronavirus (COVID-19) pandemic inequities in the United States. *Health Educ Behav.* (2020) 47:504–8. doi: 10.1177/1090198120922942

49. Gee GC, Hicken MT. Structural racism: the rules and relations of inequity. *Ethn Dis.* (2021) 31:293–300. doi: 10.18865/ed.31.S1.293

50. Sewell AA. The racism-race reification process: a mesolevel political economic framework for understanding racial health disparities. *Sociol Race Ethn.* (2016) 2:402–32. doi: 10.1177/2332649215626936

51. Solar O, Irwin A. *A Conceptual Framework for Action on the Social Determinants of Health*. Geneva: World Health Organization.

52. Jones C. Levels of racism: a theoretic framework and a gardener's tale. *Am J Public Health.* (2000) 90:1212–5. doi: 10.2105/AJPH.90.8.1212

53. Montez JK, Beckfield J, Cooney JK, Grumbach JM, Hayward MD, Koytak HZ, et al. US State policies, politics, and life expectancy. *Milbank Q.* (2020) 98:668–99. doi: 10.1111/1468-0009.12469

54. Montez JK, Hayward MD. Zajacova A. Educational disparities in adult health: US States as institutional actors on the association. *Socius.* (2019) 5:2378023119835345. doi: 10.1177/2378023119835345

55. Wolf DA, Monnat SM, Montez JK. Effects of US state preemption laws on infant mortality. *Prev Med.* (2021) 145:106417. doi: 10.1016/j.ypmed.2021.106417

56. Bruch SK, Rosenthal AJ, Soss J. Unequal positions: a relational approach to racial inequality trends in the US States, 1940–2010. *Soc Sci Hist.* (2019) 43:159–84. doi: 10.1017/ssh.2018.36

57. Alexander M. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York: The New Press (2012). p. 336.

58. Lucas SR. *Just Who Loses?* Philadelphia: Temple University Press. (2013). p. 350.

59. Fording RC, Soss J, Schram SF. Race and the local politics of punishment in the new world of welfare. *Am J Sociol.* (2011) 116:1610–57. doi: 10.1086/657525

60. Baker RS. The historical racial regime and racial inequality in poverty in the American South. *Am J Sociol.* (2022) 127:1721–81. doi: 10.1086/719653

61. Williams DT, Baker RS. Family structure, risks, and racial stratification in poverty. *Soc Probl.* (2021) 68:964–85. doi: 10.1093/socpro/spab018

62. Cogburn CD. Culture, race, and health: implications for racial inequities and population health. *Milbank Q.* (2019) 97:736–61. doi: 10.1111/1468-0009.12411

63. Ray R, Lantz P, Williams D. Upstream policy changes to improve population health and health equity. Urban Instititute (2023).

64. Kijakazi K, Brown S, Charleston D, Runes C. *Next50 Catalyst Brief: Structural Racism*. United States of America: Urban Institute. Available online at: https://policycommons.net/artifacts/630729/next50-catalyst-brief/1612019/

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 09 March 2023
DOI 10.3389/fpubh.2023.995529

OPEN ACCESS

EDITED BY
Michel Guillot,
Pennsylvania Hospital, United States

REVIEWED BY
Alexis Santos,
The Pennsylvania State University (PSU), United States
Irma Elo,
University of Pennsylvania, United States

*CORRESPONDENCE
Catherine García
✉ cgarci24@syr.edu

SPECIALTY SECTION
This article was submitted to Life-Course Epidemiology and Social Inequalities in Health, a section of the journal Frontiers in Public Health

RECEIVED 16 July 2022
ACCEPTED 14 February 2023
PUBLISHED 09 March 2023

CITATION
García C, Garcia MA, McEniry M and Crowe M (2023) The neighborhood context and all-cause mortality among older adults in Puerto Rico. *Front. Public Health* 11:995529. doi: 10.3389/fpubh.2023.995529

The neighborhood context and all-cause mortality among older adults in Puerto Rico

Catherine García[1]*, Marc A. Garcia[2], Mary McEniry[3] and Michael Crowe[4]

[1]Department of Human Development and Family Science, Aging Studies Institute, Center for Aging and Policy Studies, and the Lerner Center for Public Health Promotion and Population Health, Syracuse University, Syracuse, NY, United States, [2]Department of Sociology, Aging Studies Institute, Center for Aging and Policy Studies, and the Lerner Center for Public Health Promotion and Population Health, Syracuse University, Syracuse, NY, United States, [3]Center for Demography and Ecology, and Center for Demography of Health and Aging, University of Wisconsin-Madison, Madison, WI, United States, [4]University of Alabama at Birmingham, Birmingham, AL, United States

Background: Recent efforts have been made to collect data on neighborhood-level attributes and link them to longitudinal population-based surveys. These linked data have allowed researchers to assess the influence of neighborhood characteristics on the health of older adults in the US. However, these data exclude Puerto Rico. Because of significantly differing historical and political contexts, and widely ranging structural factors between the island and the mainland, it may not be appropriate to apply current knowledge on neighborhood health effects based on studies conducted in the US to Puerto Rico. Thus, we aim to (1) examine the types of neighborhood environments older Puerto Rican adults reside in and (2) explore the association between neighborhood environments and all-cause mortality.

Methods: We linked data from the 2000 US Census to the longitudinal Puerto Rican Elderly Health Conditions Project (PREHCO) with mortality follow-up through 2021 to examine the effects of the baseline neighborhood environment on all-cause mortality among 3,469 participants. Latent profile analysis, a model-based clustering technique, classified Puerto Rican neighborhoods based on 19 census block group indicators related to the neighborhood constructs of socioeconomic status, household composition, minority status, and housing and transportation. The associations between the latent classes and all-cause mortality were assessed using multilevel mixed-effects parametric survival models with a Weibull distribution.

Results: A five-class model was fit on 2,477 census block groups in Puerto Rico with varying patterns of social (dis)advantage. Our results show that older adults residing in neighborhoods classified as *Urban High Deprivation* and *Urban High-Moderate Deprivation* in Puerto Rico were at higher risk of death over the 19-year study period relative to the *Urban Low Deprivation* cluster, controlling for individual-level covariates.

Conclusions: Considering Puerto Rico's socio-structural reality, we recommend that policymakers, healthcare providers, and leaders across industries to (1) understand how individual health and mortality is embedded within larger social, cultural, structural, and historical contexts, and (2) make concerted efforts to reach out to residents living in disadvantaged community contexts to understand better what they need to successfully age in place in Puerto Rico.

KEYWORDS

neighborhood characteristics, mortality, Puerto Rican adults, latent variable analysis, multilevel survival analysis, PREHCO, older adults, social determinants of health (SDOH)

1. Introduction

The twenty-first century in the Commonwealth of Puerto Rico (hereafter, Puerto Rico)—an unincorporated United States (US) territory—is an era characterized by rapid population aging, reductions in social and economic resources, rampant disparities in access to adequate healthcare, and the ongoing reconstruction of the built environment post-Hurricanes Irma and María (1–6). The constellation of these factors infers that many older Puerto Rican adults may lack access to resources, services, and contexts considered necessary for promoting healthy aging.[1] In order to understand contemporary conditions in Puerto Rico, it is important to consider how historical contexts contribute to health inequities over time, particularly for older adults at increased risk for poor health and mortality.

Researchers have argued that social, political, and economic inequalities in Puerto Rico derive from the impacts of US colonialism—a structural and social determinant of health (7, 8). One significant impact of US colonialism was the transition of Puerto Rico from a rural agricultural society to an urban industrial society in the early twentieth century (9). This transition brought public health benefits, including improved sanitation practices and housing conditions, the creation of local health boards and hospitals, and increased access to primary education. However, urbanization in Puerto Rico also led to widening economic and racial disparities that resulted in unfavorable neighborhood and living conditions among socially marginalized individuals (e.g., poor and Black Puerto Ricans) (10).

For example, San Juan, the capital of Puerto Rico, has a long history of continuous urban growth and economic development. Under US control, San Juan experienced substantial modernization, including changes in land use efficiency and aggregation of local areas that connected land use with global-scale factors. Notably, new and growing opportunities in the San Juan wage labor market were a major driver for rural-dwelling Puerto Ricans to relocate in the early twentieth century; this rural-to-urban migration affected the subsequent development and preservation of several neighborhoods in the metropolitan area (11). Due to their lower socioeconomic position, Puerto Ricans from rural areas were forced to reside in poor and disadvantaged, communities in San Juan, such as La Perla (12). In addition to rural-urban migration patterns, rapid population growth and efforts to mirror the US model of suburbanization were additional factors that influenced variations in the investment of resources across neighborhoods in San Juan throughout the twentieth century that contributed to contemporary residential segregation patterns (13). For example, a study examining residential segregation in the San Juan-Bayamón metropolitan area, the most racially diverse metropolitan area in Puerto Rico, found that neighborhoods with a higher percentage of Black residents were associated with lower socioeconomic status (14).

In addition, a study focusing on the socioeconomic features of neighborhoods to assess health disparities in Puerto Rico found that municipalities (considered county-equivalents by the US Census) with a low socioeconomic position (SEP) were linked to higher cancer-related mortality rates (15). Importantly, the study showed that more deprived municipalities of Puerto Rico were in the island's central region.[2] In contrast, less deprived municipalities were concentrated in the San Juan metropolitan area. This suggests that residents living in municipalities with lower SEPs may lack access to healthcare services and health-promoting resources due to economic, environmental, and physical barriers that impact health and increase the risk of mortality. However, these findings are conditional based on the assumptions made regarding area-based socioeconomic status. Better inference of neighborhood effects would require a more nuanced approach on how specific constructs of the neighborhood environment are measured (e.g., census tract vs. census block group) and how they influence health and the risk of mortality (16).

Recent efforts have been made to collect data on neighborhood-level attributes and link them to longitudinal population-based surveys [e.g., the Health and Retirement Study Contextual Data Resource (HRS-CDR)] (17). These linked data have allowed researchers to assess the influence of neighborhood characteristics on the health of older adults in the US. However, these data only include the contiguous US and exclude Alaska, Hawai'i, and the five permanently inhabited US territories, including Puerto Rico. Because of significantly differing historical and political contexts, and widely ranging structural factors between Puerto Rico and the US mainland, it is not appropriate to apply current knowledge on neighborhood health effects based on studies conducted in the contiguous US to Puerto Rico. In addition, despite Puerto Rico's status as an unincorporated US territory, its social and economic contexts are more like Latin American and Hispanic-Caribbean countries than the US, which may lead to substantially different risk factors for poor health and mortality.

In this study, we aim to highlight multilevel perspectives and analyses of social determinants of health among older adults residing in Puerto Rico. We address a gap in the literature by using longitudinal data from the Puerto Rican Elderly Health Conditions Project (PREHCO) linked with 2000 US Census data to (1) examine the types of neighborhood environments older Puerto Rican adults reside in and, (2) explore the association between neighborhood environment and all-cause mortality.

2. Background

It is widely recognized that physical and social environments influence health behaviors, health outcomes, and mortality in the US. Although the neighborhood environment affects the health of people of all ages, the effects of the neighborhood environment may be accentuated among older adults as they are more likely than younger adults to have spent decades in the same community, have decreased physical mobility and cognitive functioning, and rely more on community resources for social integration and support (18). The combination of these factors may result in an early onset of age-related diseases (19), reduced life expectancy (20), and an increased risk of all-cause mortality (21–23). Notably, a vast array of

1 Healthy aging here is defined as good health, high mental and physical functioning, and active involvement in life.

2 The central region of Puerto Rico has many mountainous places and a higher proportion of individuals living in rural areas.

research has shown that individuals residing in neighborhoods with greater deprivation have poorer health behaviors (24), lack access to preventive health services (25), are exposed to chronic stress and pollutants (26), experience greater biological weathering (27), have worse health outcomes (28), and experience higher mortality rates (29). In many of these studies, neighborhood deprivation is based on socioeconomic contextual variables or indices related to income, education, employment, and housing, typically at the census tract level. Although these socioeconomic indicators have different meanings for older adults, it is noteworthy that the influence of socioeconomic deprivation persists in the oldest ages (30). Indeed, several studies suggest a cumulative effect of disadvantage across the lifespan that results in poor health and an increased risk of mortality (31). However, there is limited knowledge of how these multilevel processes influence population health and mortality in Puerto Rico due to the lack of data infrastructure to support these inquiries.

2.1. Neighborhood socioeconomic context

Various theoretical perspectives and conceptual frameworks have been put forth to explain why the neighborhood socioeconomic context (NSEC) plays a vital role in poor health outcomes and mortality risk. For instance, the ecological framework with a life course perspective would suggest that individuals living in disadvantaged NSECs are more likely to have a low socioeconomic position themselves due to constrained opportunity structures (22, 32, 33). Individuals who spend their early life in lower-income neighborhoods have less access to quality education than their peers residing in higher-income communities. This limits opportunities to obtain higher levels of education and marketable job skills and reduces lifetime earnings (34, 35). Thus, the importance of neighborhood context as a fundamental cause of mortality cannot be overlooked (36), particularly given the vast literature documenting how education shapes access to resources that promote better health and an individual's exposure to multiple health risks (37).

Another theoretical consideration is the systemic perspective, which infers that the NSEC affects the social, service, and physical environments of communities shared by residents. Namely, neighborhoods characterized by low socioeconomic levels are linked to underinvestment in health-promoting resources, such as lack of green and recreational spaces, adequate public transportation, affordable and high-quality grocery stores, and access to medical and social services (23). For example, individuals residing in high-poverty neighborhoods are less likely to have access to recreational opportunities to walk and exercise and are more likely to live in food swamps[3] (38, 39). Not being able to engage in healthy behaviors due to these structural challenges can increase the likelihood of early disease onset, reduce active life expectancy (e.g., physical mobility), and increase the risk of mortality. Overall, the emphasis of the NSEC on health is important from a public health perspective since resource-poor environments can be potentially addressed through community-level interventions, including investments in public education, transportation, expansion of door-to-door services (e.g., Meals on Wheels), and affordable and quality housing to name a few.

Although research has overwhelmingly demonstrated that the NSEC is a crucial determinant of health, other neighborhood-related factors interplay with the NSEC, such as a neighborhood's age structure, racial composition, residential stability, and family structure that shape opportunities and health-enhancing resources made available for residents across communities. We provide a summary of how each of these neighborhood-level determinants potentially influences health outcomes and the risk of mortality.

2.2. Neighborhood age structure

The age structure of a neighborhood may be particularly important to older adults who age in place as it may influence the provision of health services and facilities (including Medicaid reimbursements), perceptions of neighborhood safety, and opportunities for social engagement (40). Previous research has shown that neighborhoods with a high concentration of older adults are associated with better health among older adults, including those who are socioeconomically disadvantaged (41, 42). Evidence suggests that the presence of older adults in the community facilitates social integration and cultivates social ties, mutual support, social cohesion, and perceived safety (43), which is independently associated with various population-level health outcomes, including mortality (44). Several pathways have been hypothesized on how aspects of the social environment may influence health and mortality, including the impact of health behaviors and physiology (e.g., allostatic load) (44, 45). Specifically, individuals with positive social ties are less likely to engage in smoking and drinking and are more likely to receive preventive health screenings (e.g., cancer screenings). In contrast, socially isolated individuals are more likely to have weakened immune function, cardiovascular disease, and cognitive impairment. Older adults with chronic health conditions, disabilities, who live alone, and have reduced social networks are at an increased risk of social isolation, which has been shown to negatively impact health and mortality.

With the population of Puerto Rico is rapidly aging—due to a combination of outmigration among younger cohorts of adults, declining fertility, and increased longevity—these demographic changes will challenge the ability of Puerto Rico and local communities to meet the growing demands of older adults, including care and quality of life, that may further strain the collective (and scarce) resources available (1, 3, 5, 46). Specifically, increases in poverty and declining economic conditions across the archipelago, changes in the family structure, and the limited availability (and proximity) of individuals and/or services to provide long-term care for older adults in Puerto Rico (due in part to out migration of family and professionals) may result in poor health and an increased risk of mortality. Older Puerto Ricans, cognizant of these social realities have expressed concerns with loss of family cohesion and intergenerational support due to their

3 Food swamps describe the combination of food deserts and a high-density of stores and restaurants that offer high-calorie fast food and junk food.

children's search of economic opportunities outside of Puerto Rico (47). This suggests that places in Puerto Rico that have a larger concentration of older adults, particularly in rural areas, may not have the resources necessary for older adults to successfully age in place.

2.3. Neighborhood racial composition

Neighborhood racial composition has been shown to be associated with poor health and an increased risk of mortality among older adults due in part to exposure of institutionalized and systemic anti-Black racism across the life course (48–50). A large body of research shows that Black (including African American and Afro-Latino) individuals in the US overwhelmingly reside in residentially segregated neighborhoods that are characterized by concentrated economic disadvantage, which is often associated with disinvestment of municipal resources (e.g., high-quality medical care), poorly maintained infrastructures (e.g., sidewalks and green spaces), and densely populated and subpar housing quality (51–53). These conditions stem from racial capitalism and environmental racism that intentionally create the underdevelopment of non-White spaces (54). The purposeful underdevelopment of these communities results in unequal exposure to contextual health-related risks that over time exact wear and tear on the body, which contributes to a process of "weathering," leading to physiological dysregulation, the early onset of disease and disability, and ultimately mortality (55).

Although Puerto Rico appears to have a more flexible attitude toward race (i.e., the concept of "racial democracy") than the US, there is ample evidence documenting that racial minorities, immigrants (e.g., Dominican immigrants), and phenotypically dark-skinned individuals in Puerto Rico are stigmatized, discriminated against, and experience more socioeconomic disadvantage than their more socially advantaged counterparts (14, 56–59). Notably, Black communities in Puerto Rico[4] are largely located along the coastal regions of the Puerto Rican archipelago—a legacy of plantation slavery—and are regions that exhibit lower levels of education, lower median household income, lower median housing values, and higher rates of poverty and unemployment relative to predominantly White communities in Puerto Rico (60). Indeed, for Black Puerto Ricans, systemic and institutional racism across generations and across the lifespan have led to the inequitable access of social, educational, and material resources that have direct (e.g., access to health care) and indirect (e.g., stress and psychosocial resources) effects on health and mortality.

A community-based study of Puerto Rican adults aged 25–55 years in Guayama, Puerto Rico (a southeastern coastal town) found that respondents that are culturally defined as *negro* (Black) have higher systolic blood pressure (SBP) and diastolic blood pressure (DBP) than those who are classified as *blanco* (White) or *trigueño* (racially mixed)[5] (61). Additionally, Black Puerto Ricans who occupy higher socioeconomic status (SES) positions exhibit higher SBP and DBP relative to their Black counterparts in low SES contexts (61). The authors posit that Black Puerto Ricans' chronic exposure to institutional and interpersonal discrimination may be linked to their adverse cardiovascular responses (i.e., high blood pressure). Thus, deeply embedded, and multiple dimensions of racism in Puerto Rico are associated with the pronounced residential segregation of Black Puerto Ricans that results in constrained access to resources and opportunities which affect health and mortality.

2.4. Neighborhood residential stability

Living in residentially stable neighborhoods is theorized to promote the health and wellbeing of its residents as it facilitates the development of interpersonal bonds and ties (i.e., social cohesion) that individuals can draw on in times of need (i.e., social support) and may encourage healthy behaviors, and extend longevity. However, a study by Ross et al. (62) found that residential stability was only associated with enhanced psychological wellbeing among residents in affluent neighborhoods. In contrast, residential stability did not benefit the mental health of residents in impoverished communities. Ross et al. posit that living in a poor, stable neighborhood does not confer mental health advantages since residents of these environments do not have the instrumental and material resources needed to mitigate the high levels of disorder in their communities. For example, the chronic stress associated with living long-term in a neighborhood where the streets are dirty, noisy, and dangerous repeatedly activates the stress response, which can contribute to blood pressure and brain changes associated with mental and physical health outcomes (63). Thus, the effects of residential stability need to be considered in the context of a neighborhood's economic resources available.

Data from the U.S. Census and Puerto Rican Community Surveys show that Puerto Rican have high residential stability (64); however, no study, to our knowledge, has examined whether neighborhood-level variation in residential stability is beneficial or detrimental to the health of older adults in Puerto Rico. The scant research that does exist on island-born Puerto Ricans residing in the mainland U.S. has shown that living in ethnically dense, low NSECs reported worse physical health than island-born Puerto Ricans living in other types of NSECs (65). Individuals residing in ethnic enclaves tend to share common sociocultural characteristics (e.g., language and cultural background) and have strong social ties with community members, which have been found to be beneficial for health and mortality. However, enclaves that are formed involuntarily due to housing discrimination may not offer opportunities necessary for economic development at the individual and community levels. Given the high rates of poverty across the archipelago, we can infer that residential stability may not confer health benefits for Puerto Ricans who are living in disadvantaged NSECs.

4 Predominantly Black communities in Puerto Rico are located in the municipalities of Loíza, Yabucoa, and Patillas to name a few.

5 Trigueño in English literally means "wheat colored" and is used to (1) describe a light-skinned person with a slightly tanned complexion; or (2) describe someone with European, African, and Native American heritage. Being classified as trigueño may be socially advantageous since it does not carry the stigma of blackness.

2.5. Neighborhood family structure

Research on the association between neighborhood family structure and mortality is scarce; however, neighborhood family structure is related to the formation of social ties, which has been shown to have a robust association with extended longevity (66). For example, residents in neighborhoods with high family dissolution (e.g., single-parent households) have lower participation rates in formal voluntary organizations and local affairs. These forms of participation provide opportunities for individuals to integrate within the larger community—additionally, neighborhoods with a high percentage of individuals living alone present opportunities for crime. Individuals who live alone are more likely to go outside alone, which increases the likelihood of a targeted crime (e.g., robbery). These incidents are more likely to instill perceptions of neighborhood disorder that may contribute to the dissolution of social ties and an increased risk of mortality.

Traditionally, Puerto Ricans are very family oriented, embody familism,[6] and their families encompass extended and non-blood relatives (e.g., godparents and informally adopted children). The traditional structure of family dynamics in Puerto Rico has historically benefited older family members who often rely on family-based care. Recent research shows that intergenerational co-residence (e.g., children living with their older parents) is associated with increased functional and health support among older adults in Puerto Rico (67). However, the outmigration of younger Puerto Ricans to the US mainland, has led to a significant reduction in the number of family members available to provide care for older adult family members. Moreover, with increasing numbers of Puerto Ricans migrating in search of economic and educational opportunities, we can expect a higher risk of social isolation and lower social participation among older adults, which may be detrimental to mental and physical health (3). Thus, we can expect that communities in Puerto Rico with a high proportion of older adults that live alone and have a high proportion of single-parent households may be associated with worse health and an increased risk of mortality.

2.5.1. The present study

There is compelling theoretical and empirical evidence illustrating how various dimensions of the neighborhood environment co-occur and/or interact to influence the risk of mortality. Given the limited knowledge on the types of residential environments that older Puerto Ricans reside in, it is important to characterize the places where they live based on the factors discussed above. Previous research has shown that using latent class (or profile) models offers an efficient and statistically robust means of summarizing many indicators that constitute neighborhood risks and resources that are not captured by continuous scales or indices (68, 69). We intend to employ this method to classify how various neighborhood characteristics cluster together to create distinct neighborhood typologies that capture risk for all-cause mortality.

6 Familism emphasizes obligation and duties of families to one another.

3. Materials and methods

3.1. Data

3.1.1. Individual-level data

This study used data from the Puerto Rico Elderly Health Conditions Project (PREHCO), a representative longitudinal cohort study of community-dwelling Puerto Ricans aged 60 and older residing on the archipelago's main island that began in May 2002, with follow-up interviews completed in 2006–2007 and 2021–2022 (the data and documentation are not yet publicly available) (70). Response rates for the first two waves of PREHCO are high (>90.0%). The 4,291 respondents included in the PREHCO baseline sample were derived from a multistage, stratified sample of older adults, including oversampling in regions heavily populated by Afro-descendant individuals (e.g., residents in Loíza) and individuals over 80 years of age. Face-to-face interviews were conducted with each respondent in Spanish or with a proxy if a respondent had cognitive limitations. Additional information on the study and its design is provided elsewhere (71–73).

PREHCO obtained mortality information on respondents using a combination of the National Death Index (NDI) mortality data and PREHCO-identified deaths using reports by family members or the Puerto Rican death registry. Respondents were matched to the National Death Index (NDI) from their first PREHCO interview in 2002–2003 to December 2020, using the available matching variables in the PREHCO study, including social security number (SSN), name (first, middle, father's last name and/or mother's last name), birth date (month and year), and sex (female or male). We would like to note that many Puerto Ricans use two surnames, which adds to the difficulty in NDI matching. Thus, the investigators examined different combinations of respondents' last names to increase the likelihood of a positive match for those with two last names. Additional deaths were identified through November 2021 using family reports or the Puerto Rican death registry. The data file comprising the currently restricted PREHCO mortality database contains the PREHCO respondent's case identification number, the mortality status of the respondent (presumed dead or alive), year of death, month of death, day of death (for some), and cause of death (for most respondents). Two thousand eight hundred and thirty-two all-cause presumed deaths were identified from the cohort of 4,291 PREHCO respondents.

3.1.2. Neighborhood-level data

Data on baseline neighborhood characteristics were constructed from the 2000 Decennial US Census at the block-group level downloaded from Social Explorer and were linked with the PREHCO data (74). Census block groups typically include 600 to 3,000 people and is the smallest geographical unit for which the US Census Bureau publishes sample data. PREHCO respondents were linked to their affiliated census block group by linking their records in the public-use PREHCO to the restricted-use PREHCO geographic data file. These data were then merged with the 2000 US Census data using Federal Information Processing Standard (FIPS) codes to link the files. Out of the 2,477 unique census block group identifiers

for Puerto Rico in 2000,[7] we identified 233 unique census block groups in which PRECHO respondents resided at the time of the baseline interview, with 1–47 observations in each block group.

3.1.3. Sample selection

The baseline PREHCO cohort sample consisted of 4,291 unique respondents aged 60 and older. Given the design of the present study, we focused on individuals who were able to complete the full interview at baseline (n = 3,713). Respondents that needed proxies to do the interview were not asked health-related questions relevant to the present study (n = 578). Furthermore, we excluded respondents (n = 24) in neighborhoods with <5 individuals in any given block group to minimize statistical bias (75). Lastly, we excluded ~6% of participants (n = 220) due to missingness on baseline covariates. The variables with the highest prevalence of missing values were body mass index (BMI; 5%) and receipt of government-related income and services (1%). The final analytical sample included 3,469 participants.

Participants excluded from the analytical sample were more likely to be older (76.8 vs. 70.3 years), less likely to be married or partnered (38.8 vs. 53.2%), reported lower levels of education (6.1 vs. 8.3 years), and were more likely to receive government-related income and services (see Supplementary Table 1). Additionally, excluded participants were less likely to be obese, current smokers, and physically active. Excluded participants were also more likely to report chronic conditions and disability. We caution readers that the health profiles of excluded participants may be underestimated since proxy interviews were not asked all of the health-related questions. Thus, our analytical sample of community-dwelling older Puerto Ricans is relatively healthier than the general population of older adults in Puerto Rico. A detailed scheme showing the exclusion criteria and the analytic sample inclusion is provided in Supplementary Figure 1.

Additionally, given that measures included in our analysis are time varying, we briefly describe changes in sample characteristics for Wave 2 of PREHCO. From our analytical sample of 3,469 participants identified in Wave 1, 941 respondents (27%) did not have information reported in Wave 2 relevant to the analysis, including 226 proxy interviews, 27 respondents that became institutionalized, 319 that were lost at follow-up, and 369 respondents that were reported dead. To keep all respondents in our analysis, we conducted multiple imputation using chained equations (MICE) for missing data at Wave 2 using the mi suite of commands in Stata (76). We used the distribution of the observed data from Waves 1 and 2 to estimate a set of plausible values for the missing data in Wave 2. We then used Bodner's approach of generating the number of imputed data sets equivalent to the total percent missing and Rubin's rule for combining across data sets (77–82).

3.2. Measures

3.2.1. Mortality

The outcome of interest is all-cause mortality from May 2002 to November 2021. We calculated the time to censoring or death from the year of the interview to the year of death or censoring. For those who did not die in the interval, the censoring date was November 2021. We used years as the time metric.

3.2.2. Individual-level characteristics

Three groups of individual-level variables were considered as potential confounders in examining the role of neighborhood context and all-cause mortality—demographic, socioeconomic, and health characteristics.

3.2.2.1. Demographic variables

Age is measured in continuous years. We also included an age squared term, so we can model more accurately the effect of age rather than assuming the effect is linear for all ages. Sex was dichotomized as male or female. Marital status was dichotomized as married or partnered vs. never married, widowed, separated, or divorced. A dichotomous indicator for whether the respondent had moved from their main residence reported at baseline was also included.

3.2.2.2. Socioeconomic variables

Educational attainment is measured as continuous years of education completed. Given that PREHCO has limited variables for assessing individual income (e.g., not having combined household annual income or poverty thresholds) and wealth (e.g., not having a standardized measure of all assets and debt), we used indirect measures of income, including whether a respondent reports having difficulty paying for their daily necessities (categorized as never, sometimes, and often), whether they receive income from social welfare or the department of the family[8] (yes/no), whether they receive income from the nutritional assistance program (yes/no), and whether they have government-sponsored health insurance (excluding Medicare; yes/no). Given the strong association between individual-level socioeconomic position and mortality, it is crucial to adjust for individual socioeconomic measures to ensure the validity of neighborhood-level factors (83).

3.2.2.3. Health characteristics

We included indicators related to the morbidity process such as health behaviors, health conditions, and disability (84). Health behaviors included obesity, current smoking status, and physical activity. Dichotomous indicators were used to classify respondents as obese (i.e., body mass index of ≥30 kg/m^2), for whether the respondent reported being a current smoker at the time of the interview (yes/no), and whether they engaged in either moderate or vigorous physical activity at least three times per week (yes/no).

Health conditions included cardiometabolic diseases, other chronic conditions, and severe depression. Cardiometabolic diseases were a sum of whether the respondent self-reported heart

7 Every census block group has a unique 12-digit FIPS code.

8 The Puerto Rico department of the family specializes in individual and family social services.

problems (e.g., coronary heart disease, congestive heart failure, and heart attack), stroke, hypertension (including medication use) and diabetes (including medication use), ranging from 0 to 4. Other chronic conditions were a sum of self-reported cancer, lung disease, and arthritis, ranging from 0 to 3. We used the geriatric depression scale in its 15-item version (GSD-15) to classify respondents as having severe depression (85). Possible scores range from 0 (no depression) to 15 points (severe depression). Respondents were classified as having severe depression if they scored ≥10 points.

Disability was based on whether a respondent reported limitations in activities of daily living (ADLs) and instrumental activities of daily living (IADLs). ADLs are a continuous measure ranging from 0 to 6 and included difficulty with bathing, eating, dressing, walking across a room, getting in and out of bed, and using the toilet (86). IADLs are a continuous measure ranging from 0 to 7 and included difficulty with using the telephone, managing transportation, buying food or clothing, preparing meals, doing household tasks, taking medications, and managing finances.

3.2.3. Neighborhood-level characteristics

We included variables at the block group level that are theoretically related to and have been identified in previous studies as being associated with all-cause mortality. Neighborhood characteristics included 19 indicators related to the neighborhood constructs of socioeconomic status, household composition, minority status, and housing and transportation. These indicators included the proportion of the population living in a rural area,[9] Black residents, residents aged ≥65 years, older adults living alone, residents that lived in the same house past 5 years (residential stability), residents with <9 years of education, residents aged ≥16 years unemployed, residents aged ≥16 years employed in management, professional, and related occupations, households with income ≥$40,000, households with interest, dividend, or rental income, households with public assistance income, residents below 150% of the poverty threshold, single-parent households with children <18 years of age, renter-occupied housing units, residents living in crowded housing units, occupied housing units without complete plumbing, occupied housing units without a telephone, occupied housing units without a motor vehicle, and homes valued ≥150 k.

3.3. Statistical analysis

A latent profile analysis (LPA) was conducted using the gsem feature on Stata to characterize the types of neighborhood environments that older Puerto Ricans resided in at baseline (88). LPA is a semi-parametric finite mixture model that identifies homogenous subgroups based on common characteristics, creating mutually exclusive and exhaustive latent classes. LPA sorts data using posterior probabilities that calculate the probability of membership in each latent class given. Unlike other agglomerative approaches, such as cluster analysis and factor analysis, LPA is a non-parametric statistical technique that relaxes assumptions about normality and linearity in the variables used in the analyses, making LPA ideal for analyzing neighborhood-level characteristics since they do not have normal distributions. We selected the class solution that best represented the data using a combination of model fit statistics, the interpretability of the classes that emerged, and sample size per class once combined with the PREHCO data. When comparing class solutions based on model fit statistics, generally, lower values of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are preferred (89); and entropy with values approaching 1, indicating a clear delineation of classes, are preferred (90).

Next, we described the characteristics of the PREHCO analytic sample by each neighborhood cluster that emerged. Means and percentages were calculated using the xtsum and xttab features in Stata to account for the multilevel design and repeat observations.

Lastly, we estimated hazard ratios (HRs) and 95% confidence intervals (CIs) for all-cause mortality by applying a multilevel mixed-effects parametric survival model with a Weibull distribution and Berndt–Hall–Hall–Hausman (BHHH) optimization algorithm using the mestreg feature in Stata. We modeled our data with a three-level hierarchical structure: respondents (level 1) nested within each wave (level 2) and census block groups (level 3). Time-to-event was defined as the elapsed time, in years, from the baseline interview to the date of death or the end of the study follow-up, whichever came first. When we fitted a model, we included the neighborhood clusters and controlled for individual-level demographic variables: sex, age, age squared, and marital status (Model 1). Next, we proceeded to add individual-level socioeconomic indicators: education, income from social welfare, income from the nutritional assistance program, and government-sponsored health insurance (Model 2). Lastly, we added individual-level health characteristics: obesity, smoking, physical activity, cardiometabolic conditions, other chronic conditions, severe depression, and disability (i.e., ADLs and IADLs; Model 3).

All data wrangling, visualization, and analyses were conducted in Stata/MP version 17.0 (91). The data were weighted using PREHCO-provided sampling weights to ensure the representativeness of the PREHCO survey and to account for the sampling design to get reliable statistical estimates. The study protocol was deemed exempt by the Institutional Review Board at Syracuse University.

4. Results

4.1. Neighborhood clusters derived from the LPA

Latent profile models were fit based on 19 block-group level indicators using the 2,477 observations (i.e., unique block groups) available in the 2000 US Census for Puerto Rico, ranging from two to seven classes. Based on the model fit statistics, sample size, and accounting for interpretability, we chose the five-class model as

9 The US Census Bureau classifies urban/rural areas based on population thresholds, density, and land-use characteristics. Typically, rurality is designated as all population, housing, and territory not included within an urbanized area or urban cluster (≥2,500 people) (87).

10.3389/fpubh.2023.995529

TABLE 1 Summary of latent classes based on the year 2000 census block groups in Puerto Rico.

	All block groups	Urban low deprivation	Urban low-moderate deprivation	Rural moderate deprivation	Urban high-moderate deprivation	Urban high deprivation
Probability (class)		0.079	0.323	0.059	0.470	0.068
Probability of						
Rural	0.055	0.000	0.005	0.679	0.029	0.000
Black	0.082	0.040	0.083	0.040	0.086	0.129
Adults ≥65 years of age	0.124	0.172	0.135	0.102	0.118	0.081
Older adults living alone	0.328	0.323	0.302	0.299	0.338	0.417
Lived in same house past 5 years	0.727	0.666	0.714	0.770	0.746	0.688
<9 years of education	0.259	0.072	0.175	0.382	0.321	0.336
Unemployed	0.207	0.061	0.137	0.265	0.244	0.404
Employed in management and professional occupations	0.252	0.516	0.300	0.208	0.196	0.143
Households with ≥ $40,000 income	0.147	0.490	0.201	0.067	0.078	0.040
Households with interest, dividend, or rental income	0.048	0.188	0.053	0.024	0.029	0.020
Households with public assistance income	0.205	0.044	0.117	0.290	0.249	0.425
Population living below 150% of the poverty threshold	0.262	0.076	0.153	0.349	0.314	0.569
Single-parent households with children <18 years of age	0.190	0.117	0.168	0.131	0.189	0.436
Renter-occupied housing units	0.288	0.226	0.249	0.190	0.271	0.751
Living in crowded housing	0.194	0.095	0.156	0.259	0.217	0.269
Homes without complete plumbing	0.054	0.011	0.025	0.080	0.071	0.094
Homes without a telephone	0.241	0.047	0.137	0.346	0.305	0.422
Homes without a motor vehicle	0.302	0.158	0.215	0.294	0.339	0.635
Homes valued ≥ $150,000	0.124	0.556	0.108	0.081	0.072	0.099
Number of census block groups	2,477	236	783	143	1,149	166

Due to rounding, not all values add to 1.0.

having the best fit for further analysis (see Supplementary Tables 2–7). A summary of the five-class solution of neighborhood clusters is presented in Table 1.

We labeled the first cluster *Urban Low Deprivation* (Class 1), representing 7.9% of census block groups in Puerto Rico ($N = 236$). This cluster was characterized by block groups that were almost all urban, had the lowest proportion of Black individuals present, the highest proportion of older adults present, very favorable socioeconomic conditions, stable family structure, and favorable housing features relative to the other classes.

The second cluster was labeled *Urban Moderate-Low Deprivation* (Class 2) and represented 32.3% census block groups in Puerto Rico ($N = 783$). This cluster was characterized by block groups that were like the previous neighborhood cluster but notably had lower socioeconomic conditions, family structures that were somewhat less stable, and less favorable housing conditions compared to the first neighborhood cluster.

We labeled the third cluster *Rural Moderate Deprivation* (Class 3), representing 5.9% of census block groups in Puerto Rico ($N = 143$). This cluster was characterized by block groups that were predominantly rural, had a low proportion of Black individuals present, the lowest proportion of older adults living alone, unfavorable socioeconomic conditions, stable family structure, and unfavorable housing conditions relative to previous classes.

We labeled the fourth cluster *Urban Moderate-High Deprivation* (Class 4), representing 47.0% of census block groups in Puerto Rico ($N = 1{,}149$). This cluster was characterized by block groups that were predominantly urban, a higher proportion of older adults living alone, less favorable socioeconomic conditions, family structures that were less stable, and less favorable housing conditions relative to previous classes.

The final cluster represented 6.8% of census block groups in Puerto Rico ($N = 166$) and was labeled *Urban High Deprivation* (Class 5). This cluster was characterized by block groups that were urban, had the highest proportion of Black individuals present, the lowest proportion of older adults present yet the highest proportion of older adults living alone, very unfavorable socioeconomic conditions, unstable family structure, and unfavorable housing conditions relative to the other classes.

To better contextualize where these neighborhood clusters are geographically located in Puerto Rico, we provide a map of the neighborhood clusters identified in Puerto Rico by census block group (Figure 1). Neighborhoods that were classified as *Urban Low Deprivation* and *Urban High Deprivation* were mainly found in the municipalities of San Juan (the largest municipality), Ponce (the largest municipality outside the San Juan area), and Mayagüez (the largest municipality on the west side of the island). Neighborhoods characterized as *Urban Low-Moderate Deprivation* tended to be clustered outside larger municipalities (e.g., outside of San Juan). Neighborhoods characterized as *Urban High-Moderate Deprivation* and *Rural Moderate Deprivation* were distributed across the archipelago. Notably, neighborhoods in the *Rural Moderate Deprivation* cluster tended to be in the mountainous regions of the archipelago (i.e., the central part of Puerto Rico).

4.2. Characteristics of older Puerto Ricans by neighborhood cluster

The summary statistics of the PREHCO study sample by neighborhood cluster are presented in Table 2. We find that PREHCO respondents who resided in neighborhoods classified as *Urban Low Deprivation* (Class 1; $n = 224$), *Urban Low-Moderate Deprivation* (Class 2; $n = 1{,}153$), and *Rural Moderate Deprivation* (Class 3; $n = 153$) had a lower proportion of deaths over the study period relative to those residing in more disadvantaged neighborhood contexts. Older Puerto Ricans residing in the most advantaged neighborhood contexts included a higher proportion of female respondents, were older, less likely to move residences between waves, more educated, did not suffer from economic deprivation, and were less likely to report cardiometabolic conditions and disability. In contrast, respondents residing in the *Urban High Deprivation* (Class 5; $n = 281$) cluster had a higher proportion of individuals who died over the study period. The composition of this neighborhood cluster included a lower proportion of females, were younger, were the least likely to be married or partnered, more likely to move residences between waves, were less educated, suffered from economic deprivation, and were more likely to be classified with severe depression. Respondents in the *Rural Moderate Deprivation* (Class 3; $n = 153$) and *Urban High-Moderate Deprivation* (Class 4; $n = 1{,}658$) neighborhood clusters had similar demographic compositional profiles. However, respondents in the *Rural Moderate Deprivation* cluster had the lowest years of education attained relative to the other neighborhood clusters and had worse socioeconomic profiles relative to the *Urban High-Moderate Deprivation* cluster. Moreover, respondents in the *Rural Moderate Deprivation* cluster had relatively healthier behavioral profiles (e.g., more physically active, and lower proportion of obese individuals and current smokers) compared to the *Urban High-Moderate Deprivation* cluster.

4.3. Association of neighborhood clusters with all-cause mortality

The results of the fitted multilevel survival models are summarized in Table 3. Hazard ratios (HR) are presented with 95% confidence intervals (CI). Hazard ratios >1 indicate that the mortality hazard is increasing, whereas hazard ratios <1 indicate that the mortality hazard is decreasing. The results of Model 1 (our base model) show that neighborhood clusters are associated with an increased hazard in all-cause mortality among older Puerto Ricans. Older adults that resided in the *Urban Low-Moderate Deprivation* [HR: 2.94; 95% CI (1.33, 6.49)], *Rural Moderate Deprivation* [HR: 2.60; 95% CI (1.10, 6.13)], *Urban High-Moderate Deprivation* [HR: 3.55; 95% CI (1.58, 7.94)], and *Urban High Deprivation* [HR: 5.59; 95% CI (2.24, 13.96)] clusters at baseline had higher mortality rates over the study period relative to the *Urban Low Deprivation* cluster. We also observed that female and married or partnered respondents had lower mortality rates over the study period, and that increasing age was associated with higher mortality rates, which is consistent with results from studies in high- and middle-income countries.

Controlling for individual-level socioeconomic characteristics (Model 2) reduced the HR gradient of all the neighborhood clusters associated with all-cause mortality observed in Model 1. For example, adjusting for individual-level socioeconomic characteristics decreased the HR by ∼32–37% for the *Urban Low-Moderate Deprivation*, *Urban High-Moderate Deprivation*, and *Urban High Deprivation* clusters but they remained significantly associated with all-cause mortality. Conversely, adjusting for individual-level socioeconomic characteristics reduced the *Rural Moderate Deprivation* cluster to non-significance [HR: 1.83; 95% CI (0.78, 4.30)]. Furthermore, our results indicate that higher levels of education and receiving nutritional assistance was associated with lower mortality over the study period, whereas reporting government-sponsored health insurance was associated with higher mortality over the study period.

Additionally controlling for individual-level health characteristics (Model 3) further reduced (changed) the HR for all the neighborhood clusters. The *Urban High-Moderate Deprivation* and *Urban High Deprivation* clusters exhibited an ∼20–25% decrease (change) in the HR and were still significantly associated with all-cause mortality. For the *Urban Low-Moderate Deprivation* cluster, adjusting for individual-level health characteristics reduced the association to non-significance [HR: 1.86; 95% CI (0.92, 3.78)]. We also found that current smoker status, and reporting affirmative to individual items for cardiometabolic disease, other

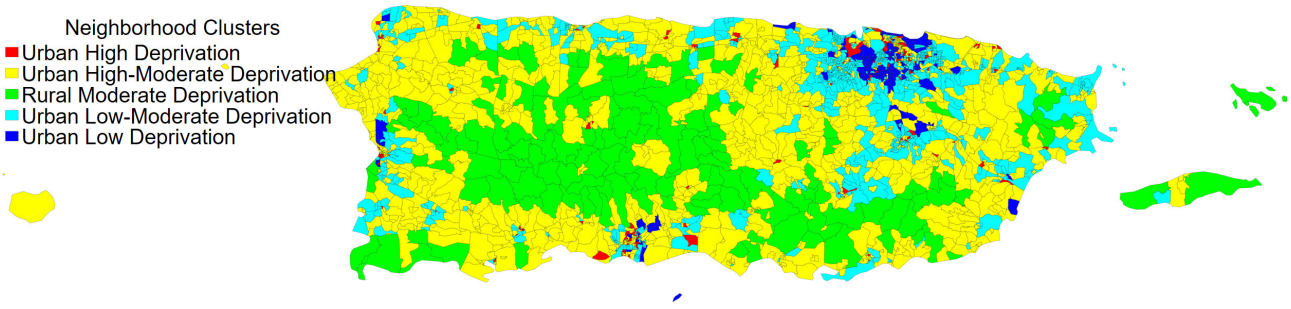


FIGURE 1
The distribution of neighborhood clusters by year 2000 census block groups in Puerto Rico. Data source: 2000 U.S. Decennial Census.

chronic conditions, and IADL limitations increased the hazard by 53, 19, 33, and 12%, respectively. In contrast, respondents that reported engaging in physical activity decreased the hazard by 28%.

Post-estimation tests of coefficients from the final model indicated that older Puerto Ricans residing in the *Urban High Deprivation* cluster were at the highest risk of death over the study period compared to all the other neighborhood clusters in Puerto Rico. Smoothed hazard estimates of the risk of mortality by neighborhood cluster demonstrating this are shown in Figure 2.

5. Discussion

Using a population-based sample of community-residing individuals aged 60 and older in Puerto Rico, this study builds on prior literature documenting the effect of neighborhood environments on all-cause mortality among older adults. Using latent profile analysis to classify neighborhoods based on indicators related to the constructs of socioeconomic status, household composition, minority status, and housing and transportation resulted in five neighborhood clusters with varying patterns of social (dis)advantage: *Urban Low Deprivation*, *Urban Low-Moderate Deprivation*, *Rural Moderate Deprivation*, *Urban High-Moderate Deprivation*, and *Urban High Deprivation*. Our results show that older Puerto Ricans residing in neighborhoods classified as *Urban High Deprivation* and *Urban High-Moderate Deprivation* in Puerto Rico (over half of our analytical sample) exhibited an increased risk of mortality over the 19-year study period after adjustment for individual-level covariates. This suggests that a high concentration of unsupportive contexts for healthy aging increases the risk of premature death. This finding is consistent with other studies in the US and Latin America that have found exposure to disadvantaged neighborhood contexts to be a robust predictor of poor health outcomes and increased risk of mortality (27, 28, 92).

In contrast, residing in neighborhoods classified as *Rural Moderate Deprivation* and *Urban Low-Moderate Deprivation* was associated with all-cause mortality among older adults, however the association was attenuated once individual-level socioeconomic factors and health-related characteristics were accounted for, respectively. Previous research has shown that individuals residing in rural communities in the US tend to be less educated, have higher rates of poverty, and are less likely to have health insurance than those residing in urban communities (93). In the case of older Puerto Ricans that reside in *Rural Moderate Deprivation* contexts, our results indicate that the socioeconomic composition of individuals residing within these communities is a more important risk factor for all-cause mortality than the deprivation that exists at the community level. Furthermore, we can infer that older adults with socioeconomic or material advantages living in these communities can alleviate some of the adverse effects and exposures associated with these environments, which may be a family-level social selection mechanism that is unaccounted for in this study (94). It is possible that individuals with economic advantages residing in rural areas in Puerto Rico have been there for generations and chose to stay for reasons related to social, cultural, human, and financial capital (95). For older adults in *Urban Low-Moderate Deprivation* neighborhood contexts, we can infer that these individuals may self-select into neighborhoods with access to a wealth of social and material resources, such as having access to preventive health care services, and access to medical care that allows for the management of age-related diseases, which can increase longevity.

With the combination of rapid aging and compounding disasters in Puerto Rico, it is imperative to document and account for multilevel determinants of mortality for older adults across later stages of the life course. From a risk environment perspective, there is a need to delineate the environmental factors associated with the risk of mortality, such as the types of environments (e.g., physical, social, economic, and policy) and level of environmental influence (micro and macro), because understanding the places in which harm is produced and reduced offers a broader vision for intervention (96). For instance, a recent review found that the long-term impacts of air pollution, heavy metals, chemicals, ambient temperature, noise, radiation, and urban residential surroundings are associated with increased mortality (97). Since aging is an active response to "weathering," we must consider how these environmental exposures are related to increases in inflammation, metabolic dysregulation, and genetic damage across the life span, increasing mortality risk. Specific to older adults, as their biological capacity declines with normal

10.3389/fpubh.2023.995529

TABLE 2 Characteristics of observations included in multilevel analysis by neighborhood cluster, PREHCO 2002–2007.

		Neighborhood cluster of residence				
		Urban low deprivation	Urban low-moderate deprivation	Rural moderate deprivation	Urban high-moderate deprivation	Urban high deprivation
	Full sample	Class 1	Class 2	Class 3	Class 4	Class 5
	% or mean ± SD					
Individual-level demographic variables						
Presumed dead	62.3	62.5	58.6	59.5	63.8	69.4
Female	59.7	70.5	61.9	51.0	57.1	61.9
Age (years)	73.4 ± 8.3	76.4 ± 9.1	73.1 ± 8.3	73.1 ± 8.3	73.1 ± 8.2	74.1 ± 8.3
Married or partnered	40.9	32.2	43.7	45.4	42.0	27.0
Moved from baseline residence	9.7	8.7	9.6	11.1	9.0	14.1
Individual-level socioeconomic variables						
Education (years)	8.1 ± 4.6	12.0 ± 3.8	9.4 ± 4.4	6.0 ± 3.9	7.0 ± 4.3	7.2 ± 4.4
Difficulty with daily needs being met						
Never	52.8	69.6	56.8	39.0	50.3	45.5
Sometimes	34.8	23.1	33.0	45.7	35.9	38.7
Often	12.4	7.3	10.2	15.3	13.8	15.9
Receives income from social welfare/department of the family	3.4	1.7	1.7	4.2	3.9	8.5
Receives income from the nutritional assistance program	29.0	10.6	19.6	48.4	33.8	43.6
Has government-sponsored health insurance	50.2	13.0	35.0	67.3	61.2	68.0
Individual-level health variables						
Obese (BMI ≥ 30 kg/m^2)	27.4	27.9	29.5	21.3	26.5	26.5
Current smoker	6.9	4.5	5.0	7.2	8.5	6.9
Physically active	57.4	63.0	60.1	61.5	54.7	55.6
Cardiometabolic diseases (0–4)	1.1 ± 0.9	0.9 ± 0.8	1.1 ± 0.9	1.2 ± 1.0	1.1 ± 0.9	1.1 ± 0.9
Other chronic conditions (0–3)	0.4 ± 0.6	0.5 ± 0.6	0.4 ± 0.6	0.4 ± 0.6	0.4 ± 0.6	0.4 ± 0.6
Severe depression (GDS ≥ 10)	7.6	6.2	6.5	6.4	8.1	11.1
Activities of daily living (0–5)	0.3 ± 0.9	0.2 ± 0.8	0.3 ± 0.9	0.3 ± 0.8	0.4 ± 0.9	0.4 ± 1.0
Instrumental activities of daily living (0–5)	0.7 ± 1.3	0.6 ± 1.2	0.6 ± 1.2	0.7 ± 1.3	0.7 ± 1.3	0.8 ± 1.3
N	3,469	224	1,153	153	1,658	281

Weighted percentages and means; N's unweighted.

aging, the effects of deleterious environmental exposures may be exacerbated among individuals who enter the later stage of the life course with pre-existing health conditions and disabilities (98). Indeed, the biophysiological mechanisms underlying the neighborhood-mortality association are just beginning to be elucidated. Nonetheless, evidence does show that there are links between social factors, physiological dysregulation, and adult mortality (99). Future data collection efforts of older adults in

TABLE 3 All-cause mortality estimated from multilevel survival models of older Puerto Rican adults (n = 3,469).

All-cause mortality	Model 1			Model 2			Model 3		
	HR		95% CI	HR		95% CI	HR		95% CI
Neighborhood-level variables									
Neighborhood clusters (ref = urban low deprivation)									
Urban low-moderate deprivation	2.94	**	[1.33, 6.49]	2.30	*	[1.07, 4.97]	1.86		[0.92, 3.78]
Rural moderate deprivation	2.60	*	[1.10, 6.13]	1.83		[0.78, 4.30]	1.76		[0.80, 3.88]
Urban high-moderate deprivation	3.55	**	[1.58, 7.94]	2.80	*	[1.24, 6.33]	2.16	*	[1.02, 4.56]
Urban high deprivation	5.59	***	[2.24, 13.96]	4.74	**	[1.82, 12.30]	3.45	**	[1.39, 8.54]
Individual-level demographic variables									
Female (ref = male)	0.53	***	[0.42, 0.66]	0.53	***	[0.42, 0.67]	0.51	***	[0.39, 0.67]
Age	1.37	***	[1.24, 1.52]	1.47	***	[1.29, 1.66]	1.47	***	[1.29, 1.68]
Age squared	1.00	***	[1.00, 1.00]	1.00	***	[1.00, 1.00]	1.00	***	[1.00, 1.00]
Married or partnered	0.75	**	[0.63, 0.90]	0.78	**	[0.67, 0.92]	0.80	**	[0.69, 0.93]
Moved from baseline residence	1.36		[0.84, 2.21]	1.29		[0.80, 2.07]	1.20		[0.80, 1.78]
Individual-level socioeconomic variables									
Education (years)				0.97	*	[0.94, 0.99]	0.98		[0.96, 1.01]
Difficulty with daily needs being met (ref = often)									
Sometimes				0.90		[0.68, 1.18]	1.01		[0.78, 1.30]
Never				1.20		[0.85, 1.69]	1.31		[0.95, 1.80]
Receives income from social welfare/department of the family				0.98		[0.63, 1.50]	1.10		[0.72, 1.68]
Receives income from the nutritional assistance program				0.78	*	[0.63, 0.97]	0.88		[0.73, 1.06]
Has government-sponsored health insurance				1.36	**	[1.12, 1.65]	1.19		[0.99, 1.44]
Individual-level health variables									
Obese (BMI $\geq$ 30 kg/m^2)							0.95		[0.80, 1.13]
Current smoker							1.53	***	[1.29, 1.81]
Physically active							0.72	**	[0.59, 0.89]
Cardiometabolic diseases (0–4)							1.19	***	[1.09, 1.29]
Other chronic conditions (0–3)							1.33	**	[1.09, 1.63]
Severe depression (GDS $\geq$ 10)							0.75		[0.53, 1.07]
Activities of daily living (0–5)							1.08		[0.96, 1.21]
Instrumental activities of daily living (0–5)							1.12	**	[1.03, 1.21]

*p < 0.05; **p < 0.01; ***p < 0.001.
HR, Hazard ratio; CI, Confidence interval.

10.3389/fpubh.2023.995529

Puerto Rico should include measures that represent multiple regulatory physiological systems (e.g., cardiovascular, metabolic, and immune) to comprehensively capture neighborhood influences on biology, and their contribution to health and mortality risks.

Considering Puerto Rico's socio-structural reality—including high levels of poverty, a deficient infrastructure, a fragile healthcare system, the dismantling of the public education system, and hazardous environmental exposures—a health disparities framework was established to reflect historical and sociocultural influences of the Puerto Rican population (100). We can draw on this framework to highlight how present disparities are rooted in historical, cultural, political, and economic factors that influence biology and behaviors and to illustrate the complex relationship between the neighborhood environment and mortality. For example, a recent study found that Puerto Rican adults residing in San Juan had multiple lifestyle risk factors and cardiometabolic conditions and recommended targeted efforts to improve the health care system and material resources among socially disadvantaged populations (101). While increasing material resources among older residents in the most disadvantaged neighborhood contexts may ease some of the challenges of aging in place, it does not get at the systemic causes of these challenges. For instance, the ports of Puerto Rico are controlled by mainland US agencies, leading to the high costs of (healthy) food on the archipelago (100). As a result, some older adults may forgo eating foods that may improve or better manage their health and decrease their mortality risk since they must make constrained choices on what to spend their limited incomes on. Thus, we recommend that policymakers, health care providers, and leaders across industries to (1) understand how individual health and mortality is embedded within larger social, cultural, structural, and historical contexts, and (2) make concerted efforts to reach out to residents living in disadvantaged community contexts to understand better what they need to successfully age in place in Puerto Rico. A study of residents in *La Perla* (an informal shantytown in San Juan with a high proportion of older adults) found that despite living in socially and economically disadvantaged residential environments, the residents reported high residential satisfaction because they built their neighborhood environment according to their community needs and have a network of support (102). This suggests that community engagement is essential to identify the health and social needs of Puerto Rican older adults and improve health in neighborhoods directly affected by inequities (103).

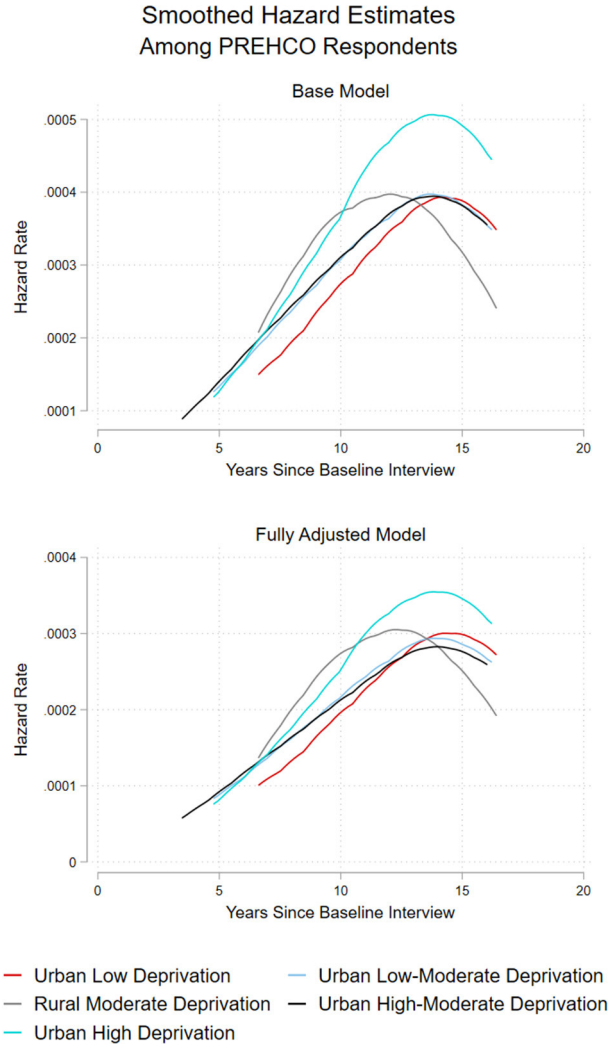


FIGURE 2
Smoothed hazard estimates of all-cause mortality by neighborhood cluster.

5.1. Limitations

Several limitations of this research should be acknowledged. First, we must recognize the physical resilience and robustness of Puerto Ricans who survived to older ages (i.e., aged ≥ 60 years) who were able to participate in the PREHCO study. Previous research has found that survival bias (or, selective survival) can attenuate associations between harmful exposures and age-related diseases, suggesting that the effects of harmful neighborhood environments may not be as pronounced among older adults and are likely underestimated (104).

Second, there are limitations associated with the operationalization of neighborhoods. We selected the smallest census unit for which we could obtain data—census block groups—to conceptualize neighborhoods in this study, an improvement from previous studies that have used census tracts as a neighborhood unit. However, recent research has emerged on the importance of activity spaces—defined as the places individuals encounter due to their day-to-day activities, which may not necessarily include their residential areas (105). Older adults may have activity spaces in more favorable or less advantageous environments relative to their residential settings that affect resources, exposures, benefits, and risks that have multifaceted effects on health and mortality. Future data collection efforts should consider capturing mobility and location information on older adults in Puerto Rico.

Third, using LPA to classify neighborhood clusters depends on the measures included to identify class types. Our findings

may be biased by the exclusion of neighborhood characteristics important for distinguishing underlying neighborhood clusters, such as the built environment (e.g., availability of green spaces), availability of health care (e.g., number of physicians and number of facilities), neighborhood crime (e.g., violent offenses), and air pollution (e.g., $PM_{2.5}$), which we lacked data on, to determine whether the identification of neighborhood clusters is improved. Nonetheless, we included multiple neighborhood variables across multiple neighborhood constructs that have been used in previous studies of all-cause mortality.

Fourth, as with any observational study, this study has unmeasured potential confounders that limit causal inference. For example, due to the limited measures related to income and wealth available in PREHCO, we could not examine if the influence of the neighborhood context differed by individual-level socioeconomic status (SES; e.g., low vs. moderate vs. high SES). Previous research has shown that death rates were higher among low SES individuals residing in high SES neighborhoods (92, 106). This suggests that there are potentially other subpopulations not captured in this study who are at higher risk for death.

Finally, we did not examine residential trajectories over time, which is especially relevant for Puerto Rico given the budget crisis, the great recession, the debt crisis, and Hurricanes Irma and María that may have resulted in increases in spatial inequality. PREHCO has publicly available data for two waves (2002–2003 and 2006–2007). The third wave of surviving respondents of PREHCO will be publicly available soon, and the fourth wave of data collection will begin later this year. These data will allow the creation of a longitudinal database to examine residential trajectories over time and their association with mortality.

Despite these limitations, our study makes several contributions on the role of neighborhoods on older adult mortality. First, we focus on older adults in Puerto Rico—a segment of the US and Latino population that is overlooked in US-based neighborhoods research and aging research more broadly. Second, we used latent profile analysis to summarize multiple indicators that constitute neighborhood risks and resources that are not captured by continuous scales or indices. Using this approach to identify neighborhood clusters associated with an increased risk of death in late life may help inform "upstream" points for structural interventions that can extend healthy years of life among older adults who have had adverse experiences throughout their life course. Third, we used longitudinal data to help establish causal inference. Using multilevel methods and longitudinal data, we assessed the temporal relationship of the association between the neighborhood context at baseline and 19-year all-cause mortality, controlling for possible confounders, allowing for more robust causal inferences. This investigation serves as a foundation to highlight a multilevel perspective of social determinants of health in Puerto Rico. Collectively, we must reframe the narrative on the aging process in Puerto Rico to understand the interplay that historical, environmental, social, behavioral, and biological factors have on health and mortality in this rapidly aging population. Through these efforts, we can identify opportunities to assess and improve the health and wellbeing of older Puerto Rican adults.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: https://doi.org/10.3886/ICPSR34596.v1 for PREHCO and https://www.socialexplorer.com/ for the 2000 US Decennial Census.

Ethics statement

The original two waves of the PREHCO study complied with all the IRB requirements at the University of Wisconsin-Madison and the University of Puerto Rico. The use of the NDI mortality data with the PREHCO study complied with IRB requirements at the University of Alabama-Birmingham. The use of the geographic data with the PREHCO study complied with IRB requirements at Syracuse University.

Author contributions

CG conceptualized and designed the study, organized and conducted the statistical analysis, interpreted the results and findings, prepared all data visualization, and wrote the manuscript. MG assisted with the interpretation and validation of results and revised the manuscript critically for important intellectual content. MM created the mortality data for PREHCO for use in this study and revised the manuscript. MC assisted with the conceptualization of the study, revised the manuscript, and acquired financial support for PREHCO. All authors approve the submitted version of this manuscript and agree to be accountable for the content of the work.

Funding

The authors gratefully acknowledge the funding of this research by the National Institute on Aging of the National Institutes of Health (R01AG064769, R01AG064769-03S1, and L30AG074407) and the services and support of the Center for Aging and Policy Studies, awarded to Syracuse University, in consortium with Cornell University and the University at Albany, by the National Institute on Aging of the National Institutes of Health (P30AG066583). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of

their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.995529/full#supplementary-material

References

1. Pérez C, Ailshire JA. Aging in Puerto Rico: A comparison of health status among Island Puerto Rican and Mainland US older adults. *J Aging Health.* (2017) 29:1056–78. doi: 10.1177/0898264317714144

2. García C, Rivera FI, Garcia MA, Burgos G, Aranda MP. Contextualizing the COVID-19 era in Puerto Rico: Compounding disasters and parallel pandemics. *J Gerontol Ser B.* (2021) 76:e263–7. doi: 10.1093/geronb/gbaa186

3. Matos-Moreno A, Verdery AM, Mendes de Leon CF, De Jesús-Monge VM, Santos-Lozada AR. Aging and the left behind: Puerto Rico and its unconventional rapid aging. *Gerontologist.* (2022) 2022:gnac082. doi: 10.1093/geront/gnac082

4. Keenan JM, Hauer ME. Resilience for whom? Demographic change and the redevelopment of the built environment in Puerto Rico. *Environ Res Lett.* (2020) 15:e074028. doi: 10.1088/1748-9326/ab92c2

5. Lozada ARS, Estrada ALV. *The Population Decline of Puerto Rico: An Application of Prospective Trends in Cohort-Component Projections.* (2015). Available online at: http://rgdoi.net/10.13140/RG.2.1.2895.5927 (accessed May 16, 2022).

6. Ramos JGP, Garriga-López A, Rodríguez-Díaz CE. How is colonialism a sociostructural determinant of health in Puerto Rico? *AMA J Ethics.* (2022) 24:305–12. doi: 10.1001/amajethics.2022.305

7. Valle AJ. Race and the empire-state: Puerto Ricans' unequal U.S. *Citizenship Sociol Race Ethn.* (2019) 5:26–40. doi: 10.1177/2332649218776031

8. Czyzewski K. Colonialism as a broader social determinant of health. *Int Indig Policy J.* (2011) 2:7337. doi: 10.18584/iipj.2011.2.1.5

9. Colón-López A, García C. 20th century Puerto Rico and later-life health: The association between multigenerational education and chronic conditions in Island-Dwelling older adults. *J Aging Health.* (2022) 2022:89826432210975. doi: 10.1177/08982643221097532

10. Zulawski A. Environment, urbanization, and public health: The bubonic plague epidemic of 1912 in San Juan, Puerto Rico. *Lat Am Res Rev.* (2022) 53:500–16. doi: 10.25222/larr.424

11. Grau HR, Aide TM, Zimmerman JK, Thomlinson JR, Helmer E, Zou X. The ecological consequences of socioeconomic and land-use changes in postagriculture Puerto Rico. *Bioscience.* (2003) 53:1159. doi: 10.1641/0006-3568(2003)053[1159:TECOSA]2.0.CO;2

12. Urban F. La Perla – 100 years of informal architecture in San Juan, Puerto Rico. *Plan Perspect.* (2015) 30:495–536. doi: 10.1080/02665433.2014.1003247

13. Azar D, Rain D. Identifying population vulnerable to hydrological hazards in San Juan, Puerto Rico. *GeoJournal.* (2007) 69:23–43. doi: 10.1007/s10708-007-9106-8

14. Denton NA, Villarrubia J. Residential segregation on the Island: The role of race and class in Puerto Rican neighborhoods. *Sociol Forum.* (2007) 22:51–76. doi: 10.1111/j.1573-7861.2006.00004.x

15. Torres-Cintrón M, Ortiz AP, Ortiz-Ortiz KJ, Figueroa-Vallés NR, Pérez-Irizarry J, Díaz-Medina G, et al. Using a socioeconomic position index to assess disparities in cancer incidence and mortality, Puerto Rico, 1995-2004. *Prev Chronic Dis.* (2012) 9:E15. doi: 10.5888/pcd9.100271

16. Diez Roux AV. Estimating neighborhood health effects: The challenges of causal inference in a complex world. *Soc Sci Med.* (2004) 58:1953–60. doi: 10.1016/S0277-9536(03)00414-3

17. Dick C. The health and retirement study: Contextual data augmentation. *For Health Econ Pol.* (2022) 2021:68. doi: 10.1515/fhep-2021-0068

18. Glass TA, Balfour JL. Neighborhoods, aging, and functional limitations. In: I Kawachi, LF Berkman, editors, *Neighborhoods and Health*. Oxford: Oxford University Press (2003). p. 303–34. Available online at: http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195138382.001.0001/acprof-9780195138382-chapter-14 (accessed April 17, 2020).

19. Clarke PJ, Weuve J, Barnes L, Evans DA, Mendes de Leon CF. Cognitive decline and the neighborhood environment. *Ann Epidemiol.* (2015) 25:849–54. doi: 10.1016/j.annepidem.2015.07.001

20. Gill TM, Zang EX, Murphy TE, Leo-Summers L, Gahbauer EA, Festa N, et al. Association between neighborhood disadvantage and functional well-being in community-living older persons. *J Am Med Assoc Intern Med.* (2021) 181:1297. doi: 10.1001/jamainternmed.2021.4260

21. Geronimus AT. The weathering hypothesis and the health of African-American women and infants: Evidence and speculations. *Ethn Dis.* (1992) 2:207–21.

22. Lawton MP, Nahemow L. Ecology and the aging process. In: C Eisdorfer, MP Lawton, editors, *The Psychology of Adult Development and Aging*. Washington, DC: American Psychological Association (1973). p. 619–74. Available online at: http://content.apa.org/books/10044-020 (accessed June 2, 2020).

23. Robert SA. Socioeconomic position and health: The independent contribution of community socioeconomic context. *Annu Rev Sociol.* (1999) 25:489–516. doi: 10.1146/annurev.soc.25.1.489

24. Stimpson JP, Ju H, Raji MA, Eschbach K. Neighborhood deprivation and health risk behaviors in NHANES III. *Am J Health Behav.* (2007) 31:215–22. doi: 10.5993/AJHB.31.2.10

25. Kirby JB, Kaneda T. Neighborhood socioeconomic disadvantage and access to health care. *J Health Soc Behav.* (2005) 46:15–31. doi: 10.1177/002214650504600103

26. Ailshire JA, García C. Unequal places: The impacts of socioeconomic and race/ethnic differences in neighborhoods. *Generations.* (2018) 42:20–7.

27. Lei MK, Beach SRH, Simons RL. Biological embedding of neighborhood disadvantage and collective efficacy: Influences on chronic illness via accelerated cardiometabolic age. *Dev Psychopathol.* (2018) 30:1797–815. doi: 10.1017/S0954579418000937

28. Powell WR, Buckingham WR, Larson JL, Vilen L, Yu M, Salamat MS, et al. Association of neighborhood-level disadvantage with Alzheimer disease neuropathology. *J Am Med Assoc Netw Open.* (2020) 3:e207559. doi: 10.1001/jamanetworkopen.2020.7559

29. Ward-Caviness CK, Pu S, Martin CL, Galea S, Uddin M, Wildman DE, et al. Epigenetic predictors of all-cause mortality are associated with objective measures of neighborhood disadvantage in an urban population. *Clin Epigenetics.* (2020) 12:44. doi: 10.1186/s13148-020-00830-8

30. Yen IH, Michael YL, Perdue L. Neighborhood environment in studies of health of older adults. *Am J Prev Med.* (2009) 37:455–63. doi: 10.1016/j.amepre.2009.06.022

31. Dannefer D. Cumulative advantage/disadvantage and the life course: Cross-fertilizing age and social science theory. *J Gerontol B Psychol Sci Soc Sci.* (2003) 58:S327–37. doi: 10.1093/geronb/58.6.S327

32. Moore KD. An ecological framework of place: Situating environmental gerontology within a life course perspective. *Int J Aging Hum Dev.* (2014) 79:183–209. doi: 10.2190/AG.79.3.a

33. Elder GH, Rockwell RC. The life-course and human development: An ecological perspective. *Int J Behav Dev.* (1979) 2:1–21. doi: 10.1177/016502547900200101

34. Cutler DM, Lleras-Muney A. Understanding differences in health behaviors by education. *J Health Econ.* (2010) 29:1–28. doi: 10.1016/j.jhealeco.2009.10.003

35. Frisvold D, Golberstein E. School quality and the education–health relationship: Evidence from Blacks in segregated schools. *J Health Econ.* (2011) 30:1232–45. doi: 10.1016/j.jhealeco.2011.08.003

36. Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav.* (1995) 35:80. doi: 10.2307/2626958

37. Mirowsky J, Ross CE. *Education, Social Status, and Health.* 1st ed. London: Routledge (2017). Available online at: https://www.taylorfrancis.com/books/9781351328074 (accessed July 15, 2022).

38. Cooksey-Stowers K, Schwartz M, Brownell K. Food swamps predict obesity rates better than food deserts in the United States. *Int J Environ Res Public Health.* (2017) 14:1366. doi: 10.3390/ijerph14111366

39. Perry A, Patlak M, Ramirez AG, Gallion KJ. *Making Healthy Food and Beverages the Affordable, Available, Desired Choices Among Latino Families.* San Antonio, TX: Salud America! (2015).

40. Cagney KA. Neighborhood age structure and its implications for health. *J Urban Health.* (2006) 83:827–34. doi: 10.1007/s11524-006-9092-z

41. Kubzansky LD, Subramanian SV, Kawachi I, Fay ME, Soobader MJ, Berkman LF. Neighborhood contextual influences on depressive symptoms in the elderly. *Am J Epidemiol.* (2005) 162:253–60. doi: 10.1093/aje/kwi185

42. Subramanian SV, Kubzansky L, Berkman L, Fay M, Kawachi I. Neighborhood effects on the self-rated health of elders: Uncovering the relative importance of structural and service-related neighborhood environments. *J Gerontol B Psychol Sci Soc Sci.* (2006) 61:S153–60. doi: 10.1093/geronb/61.3.S153

43. Cagney KA, Cornwell EY. Place, aging, and health. In: *Future Directions for the Demography of Aging: Proceedings of a Workshop*. National Academies Press (US) (2018). Available online at: https://www.ncbi.nlm.nih.gov/books/NBK513076/ (accessed June 17, 2022).

44. Seeman TE, Crimmins E. Social environment effects on health and aging: Integrating epidemiologic and demographic approaches and perspectives. *Ann N Y Acad Sci.* (2006) 954:88–117. doi: 10.1111/j.1749-6632.2001.tb02749.x

45. Berkman LF. Social integration, social networks, and health. In: *Encyclopedia of Health and Behavior*. Thousand Oaks, CA: SAGE Publications, Inc. (2004). Available online at: http://sk.sagepub.com/reference/behavior/n192.xml (accessed April 16, 2020).

46. Matos-Moreno A, Santos-Lozada AR, Mehta N, Mendes de Leon CF, Lê-Scherban F, De Lima Friche AA. Migration is the driving force of rapid aging in Puerto Rico: A research brief. *Popul Res Policy Rev.* (2022) 41:801–10. doi: 10.1007/s11113-021-09683-2

47. Zsembik BA, Bonilla Z. Eldercare and the changing family in Puerto Rico. *J Fam Issues.* (2000) 21:652–74. doi: 10.1177/019251300021005007

48. Charles CZ. The dynamics of racial residential segregation. *Annu Rev Sociol.* (2003) 29:167–207. doi: 10.1146/annurev.soc.29.010202.100002

49. Massey DS. American apartheid: Segregation and the making of the underclass. *Am J Sociol.* (1990) 96:329–57. doi: 10.1086/229532

50. Sewell AA. The racism-race reification process: A mesolevel political economic framework for understanding racial health disparities. *Sociol Race Ethn.* (2016) 2:402–32. doi: 10.1177/2332649215626936

51. Williams DR, Rucker TD. Understanding and addressing racial disparities in health care. *Health Care Financ Rev.* (2000) 21:75–90.

52. Williams DR, Collins C. Racial residential segregation: A fundamental cause of racial disparities in health. *Public Health Rep.* (2001) 116:404–16. doi: 10.1016/S0033-3549(04)50068-7

53. Williams DR, Lawrence JA, Davis BA. Racism and health: Evidence and needed research. *Annu Rev Public Health.* (2019) 40:105–25. doi: 10.1146/annurev-publhealth-040218-043750

54. Seamster L, Purifoy D. What is environmental racism for? Place-based harm and relational development. *Environ Sociol.* (2021) 7:110–21. doi: 10.1080/23251042.2020.1790331

55. Crimmins EM, Seeman TE. Integrating biology into the study of health disparities. *Popul Dev Rev.* (2004) 30:89–107.

56. Caraballo-Cueto J, Godreau IP. Colorism and health disparities in home countries: The case of Puerto Rico. *J Immigr Minor Health.* (2021) 23:926–35. doi: 10.1007/s10903-021-01222-7

57. Liu MM, Crowe M, Telles EE, Jiménez-Velázquez IZ, Dow WH. Color disparities in cognitive aging among Puerto Ricans on the archipelago. *SSM Popul Health.* (2022) 17:100998. doi: 10.1016/j.ssmph.2021.100998

58. Rodriguez-Díaz CE, Lewellen-Williams C. Race and racism as structural determinants for emergency and recovery response in the aftermath of Hurricanes Irma and Maria in Puerto Rico. *Health Equity.* (2020) 4:232–8. doi: 10.1089/heq.2019.0103

59. Lloréns H, García-Quijano C, Godreau IP. Racismo En Puerto Rico: Surveying perceptions of racism. *Cent J.* (2017) 29:154–83.

60. U.S. Census Bureau. *Quick Facts: Puerto Rico*. Washington, DC (2021). Available online at: https://www.census.gov/quickfacts/fact/table/PR/IPE120221 (accessed July 2022).

61. Gravlee CC, Dressler WW, Bernard HR. Skin color, social classification, and blood pressure in Southeastern Puerto Rico. *Am J Public Health.* (2005) 95:2191–7. doi: 10.2105/AJPH.2005.065615

62. Ross CE, Reynolds JR, Geis KJ. The contingent meaning of neighborhood stability for residents' psychological well-being. *Am Sociol Rev.* (2000) 65:581. doi: 10.2307/2657384

63. Yang L, Zhao Y, Wang Y, Liu L, Zhang X, Li B, et al. The effects of psychological stress on depression. *Curr Neuropharmacol.* (2015) 13:494–504. doi: 10.2174/1570159X1304150831150507

64. U.S. Census Bureau. *QuickFacts: Puerto Rico*. Washington, DC. (2020). Available online at: https://www.census.gov/quickfacts/fact/table/PR/AGE775218 (accessed July 2022).

65. Roy AL, Hughes D, Yoshikawa H. Intersections between nativity, ethnic density, and neighborhood SES: Using an ethnic enclave framework to explore variation in Puerto Ricans' Physical Health. *Am J Community Psychol.* (2013) 51:468–79. doi: 10.1007/s10464-012-9564-0

66. Sampson RJ. Neighborhood family structure and the risk of personal victimization. In: JM Byrne, RJ Sampson, editors, *The Social Ecology of Crime*. New York, NY: Springer New York (1986). p. 25–46. Available online at: http://link.springer.com/10.1007/978-1-4613-8606-3_2 (accessed July 15, 2022).

67. Quashie NT, Andrade FCD, Meltzer G, García C. Living arrangements and intergenerational support in Puerto Rico: Are fathers disadvantaged? *J Gerontol Ser B.* (2022) 2022:gbac044. doi: 10.1093/geronb/gbac044

68. Weden MM, Escarce JJ, Lurie N. *Technical Detail and Appendices for a Study of Neighborhood Archetypes for Population Health Research*. (*RAND Labor and Population working paper series). Report No.: WR-805*. Santa Monica, CA: RAND Corporation (2010). Available online at: https://www.rand.org/pubs/working_papers/WR805.html (accessed July 2022).

69. Humphrey J, Lindstrom M, Barton K, Shrestha P, Carlton E, Adgate J, et al. Social and environmental neighborhood typologies and lung function in a low-income, urban population. *Int J Environ Res Public Health.* (2019) 16:1133. doi: 10.3390/ijerph16071133

70. Palloni A, Luisa Davila A, Sanchez-Ayendez M. Puerto Rican elderly: Health Conditions (PREHCO) Project, 2002–2003, 2006–2007. *Inter-university Consort Polit Soc Res.* (2013) 2013:ICPSR34596. doi: 10.3886/ICPSR34596

71. McEniry M, Palloni A. Early life exposures and the occurrence and timing of heart disease among the older adult Puerto Rican population. *Demography.* (2010) 47:23–43. doi: 10.1353/dem.0.0093

72. Palloni A, McEniry M, Dávila AL, Gurucharri AG. The influence of early conditions on health status among elderly Puerto Ricans. *Soc Biol.* (2005) 52:132–63. doi: 10.1080/19485565.2005.9989106

73. McEniry M, Palloni A, Davila AL, Gurucharri AG. Early life exposure to poor nutrition and infectious diseases and its effects on the health of older Puerto Rican adults. *J Gerontol B Psychol Sci Soc Sci.* (2008) 63:S337–48. doi: 10.1093/geronb/63.6.S337

74. U.S. Census Bureau. *SF3 Prefixed Tables Summary File 3 (SF 3) - Sample Data*. Social Explorer (2000). Available online at: https://www.socialexplorer.com/reports/socialexplorer/en/report/183c339a-e43d-11ec-9428-d3638f2d61a2 (accessed July 2022).

75. Clarke P. Theory and methods: When can group level clustering be ignored? Multilevel models versus single-level models with sparse data. *J Epidemiol Community Health.* (2008) 62:752–8. doi: 10.1136/jech.2007.060798

76. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med.* (2011) 30:377–99. doi: 10.1002/sim.4067

77. Bodner TE. What improves with increased missing data imputations? *Struct Eq Model.* (2008) 15:651–75. doi: 10.1080/10705510802339072

78. Bennett DA. How can I deal with missing data in my study? *Aust N Z J Public Health.* (2001) 25:464–9. doi: 10.1111/j.1467-842X.2001.tb00659.x

79. Graham JW, Cumsille PE, Shevock AE. Methods for handling missing data. In: *Handbook of Psychology: Research Methods in Psychology*. Vol 2, 2nd ed. Hoboken, NJ: John Wiley & Sons Inc (2013). p. 109–41. doi: 10.1002/9781118133880.hop202004

80. Schafer JL. Multiple imputation: A primer. *Stat Methods Med Res.* (1999) 8:3–15. doi: 10.1191/096228099671525676

81. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York, NY: John Wiley & Sons (2004).

82. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc.* (1996) 91:473–89. doi: 10.1080/01621459.1996.10476908

83. Meijer M, Röhl J, Bloomfield K, Grittner U. Do neighborhoods affect individual mortality? A systematic review and meta-analysis of multilevel studies. *Soc Sci Med.* (2012) 74:1204–12. doi: 10.1016/j.socscimed.2011.11.034

84. Crimmins EM, Kim JK, Vasunilashorn S. Biodemography: New approaches to understanding trends and differences in population health and mortality. *Demography.* (2010) 47:S41–64. doi: 10.1353/dem.2010.0005

85. Yesavage JA, Sheikh JI. 9/Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clin Gerontol.* (1986) 5:165–73. doi: 10.1300/J018v05n01_09

86. Katz S. Assessing self-maintenance: Activities of daily living, mobility, and instrumental activities of daily living. *J Am Geriatr Soc.* (1983) 31:721–7. doi: 10.1111/j.1532-5415.1983.tb03391.x

87. Ratcliffe M, Burd C, Holder K, Fields A. *Defining Rural at the U.S. Census Bureau. American Community Survey and Geography Brief*. Washington, DC: U.S. Census Bureau (2016). Available online at: https://www.census.gov/content/dam/Census/library/publications/2016/acs/acsgeo-1.pdf (accessed July 2022).

88. Collins LM, Lanza ST. *Latent Class and Latent Transition Analysis. Wiley Series in Probability and Statistics*. Hoboken, NJ: John Wiley & Sons, Inc. (2009). Available online at: http://doi.wiley.com/10.1002/9780470567333 (accessed March 28, 2020).

89. Akaike H. Factor analysis and AIC. *Psychometrika.* (1987) 52:317–32. doi: 10.1007/BF02294359

90. Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. *J Classif.* (1996) 13:195–212. doi: 10.1007/BF01246098

91. StataCorp. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC (2021).

92. Yen IH, Kaplan GA. Neighborhood social environment and risk of death: Multilevel evidence from the Alameda county study. *Am J Epidemiol.* (1999) 149:898–907. doi: 10.1093/oxfordjournals.aje.a009733

93. Centers for Disease Control and Prevention. *About Rural Health*. Washington, DC (2017). Available online at: https://www.cdc.gov/ruralhealth/about.html (accessed June 2, 2020).

94. Jokela M. Are neighborhood health associations causal? A 10-year prospective cohort study with repeated measurements. *Am J Epidemiol.* (2014) 180:776–84. doi: 10.1093/aje/kwu233

95. Flora CB, Flora JL, Gasteyer SP. *Rural Communities: Legacy and Change*. 4th ed. London: Routledge (2018). Available online at: https://www.taylorfrancis.com/books/9780429963247 (accessed November 26, 2022).

96. Rhodes T. The "risk environment": A framework for understanding and reducing drug-related harm. *Int J Drug Policy.* (2002) 13:85–94. doi: 10.1016/S0955-3959(02)00007-5

97. 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) 18:704. doi: 10.3390/ijerph18020704

98. Geller AM, Zenick H. Aging and the environment: A research framework. *Environ Health Perspect.* (2005) 113:1257–62. doi: 10.1289/ehp.7569

99. Yang YC, McClintock MK, Kozloski M, Li T. Social isolation and adult mortality: The role of chronic inflammation and sex differences. *J Health Soc Behav.* (2013) 54:183–203. doi: 10.1177/0022146513485244

100. Lafarga Previdi I, Vélez Vega CM. Health disparities research framework adaptation to reflect Puerto Rico's socio-cultural context. *Int J Environ Res Public Health.* (2020) 17:8544. doi: 10.3390/ijerph17228544

101. Mattei J, Tamez M, Ríos-Bedoya CF, Xiao RS, Tucker KL, Rodríguez-Orengo JF. Health conditions and lifestyle risk factors of adults living in Puerto Rico: A cross-sectional study. *BMC Public Health.* (2018) 18:491. doi: 10.1186/s12889-018-5359-z

102. Caldieron J. Residential satisfaction in La Perla Informal Neighborhood, San Juan, Puerto Rico. *OIDA Int J Sustain Dev.* (2011) 2:77–84.

103. García AA, West Ohueri C, Garay R, Guzmán M, Hanson K, Vasquez M, et al. Community engagement as a foundation for improving neighborhood health. *Public Health Nurs.* (2021) 38:223–31. doi: 10.1111/phn.12870

104. McGuinness MB, Karahalios A, Kasza J, Guymer RH, Finger RP, Simpson JA. Survival bias when assessing risk factors for age-related macular degeneration: A tutorial with application to the exposure of smoking. *Ophthalmic Epidemiol.* (2017) 24:229–38. doi: 10.1080/09286586.2016.1276934

105. Cagney KA, York Cornwell E, Goldman AW, Cai L. Urban mobility and activity space. *Annu Rev Sociol.* (2020) 46:623–48. doi: 10.1146/annurev-soc-121919-054848

106. Winkleby MA, Cubbin C. Influence of individual and neighbourhood socioeconomic status on mortality among black, Mexican-American, and white women and men in the United States. *J Epidemiol Community Health.* (2003) 57:444–52. doi: 10.1136/jech.57.6.444

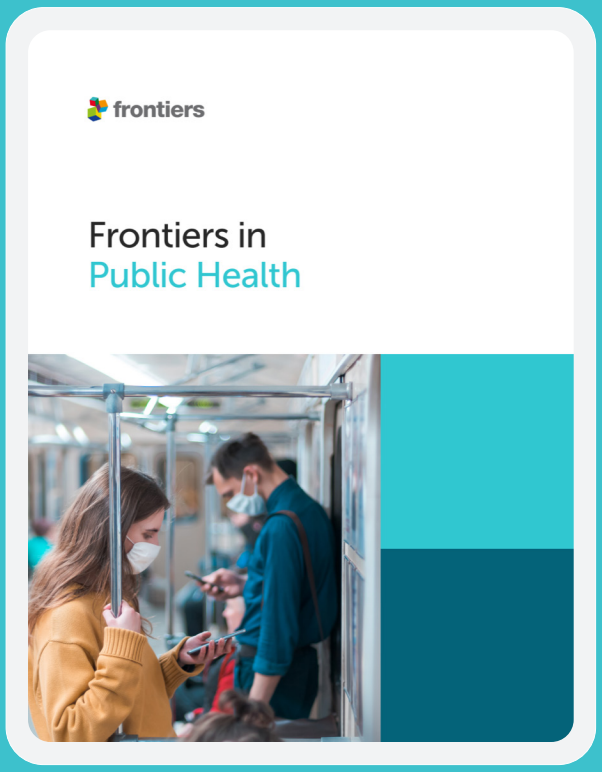
Frontiers in Public Health

Explores and addresses today's fast-moving healthcare challenges

One of the most cited journals in its field, which promotes discussion around inter-sectoral public health challenges spanning health promotion to climate change, transportation, environmental change and even species diversity.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact