# Factors Associated With Cancer Disparities Among Low-, Medium-, and High-Income US Counties 

Jeremy M. O'Connor, MD, MHS; Tannaz Sedghi, MPH; Meera Dhodapkar; Michael J. Kane, PhD; Cary P. Gross, MD


#### Abstract

BACKGROUND There are concerns about the degree to which county income level might underlie the stark disparities in cancer death rates among US counties; at the same time, there is uncertainty about the factors that may mediate the disparities.

OBJECTIVES To assess county-level cancer death rates and to identify possible mediators of the association between county-level median incomes and cancer death rates.

DESIGN, SETTING, AND PARTICIPANTS Cross-sectional study using death records from the National Center for Health Statistics in 2014, with data collected and analyzed between October 1, 2016, and July 31, 2017. All US counties and county equivalents were included.

EXPOSURES County-level median household income.

MAIN OUTCOMES AND MEASURES County-level age-standardized cancer death rate.

RESULTS In 3135 counties, median incomes ranged from $\$ 22126$ to $\$ 121250$ per year. Low-income counties (median income, $\$ 33445$ ) vs high-income counties (median income, $\$ 55780$ ) had higher proportions of residents who were non-Hispanic black, lived in rural areas, or reported poor or fair health. The mean (SD) cancer death rate was 185.9 (24.4) per 100000 person-years in high-income counties, compared with 204.9 (26.3) and 229.7 (32.9) per 100000 person-years in medium- and low-income counties, respectively. In mediation models, health risk behaviors (smoking, obesity, and physical inactivity); clinical care factors (unaffordable care and low-quality care); health environments (food insecurity); and health policies (state smoke-free laws and Medicaid payment rates) in aggregate accounted for more than $80 \%$ of the income-related disparity. The strongest possible mediators were food insecurity (explaining $19.1 \%$ [ $95 \% \mathrm{Cl}, 12.5 \%-26.5 \%$ ] of the association between county incomes and cancer deaths), low-quality care (17.9\%; 95\% CI, 14.0\%-21.8\%), smoking (12.7\%; 95\% CI, 9.4\%-15.6\%), and physical inactivity (12.2\%; 95\% CI, 9.4\%-15.6\%).

CONCLUSIONS AND RELEVANCE There are wide gaps in cancer death rates between low-, medium-, and high-income counties. Future studies are needed to assess whether targeting the possible mediators might ameliorate the substantial socioeconomic cancer disparities.


$\qquad$
JAMA Network Open. 2018;1(6):e183146. doi:10.1001/jamanetworkopen.2018.3146

## Introduction

Advances in cancer prevention, diagnosis, and treatment have led to rapid reductions in cancer mortality in the United States, with cancer death rates decreasing from 240 per 100000 personyears in 1980 to 192 per 100000 person-years in 2014. ${ }^{1-3}$ Reductions in cancer mortality, however, are not synonymous with reductions in cancer disparities-perhaps in part because of variation in

[^0]
## Key Points

Question How substantial are cancer disparities on the basis of county levels of income, and what are the factors that may mediate the disparities?

Findings In this cross-sectional study of 3135 US counties, cancer death rates varied significantly in counties of different income levels, with a mean cancer death rate per 100000 personyears of 185.9 in high-income counties, 204.9 in medium-income counties, and 229.7 in low-income counties. The strongest possible mediators were health risk behaviors, cost and quality of clinical care, and food insecurity.

Meaning There are multiple countylevel factors that may serve as mediators of cancer disparities and that may be targeted by future efforts to achieve equity in cancer outcomes.

## Supplemental content

Author affiliations and article information are listed at the end of this article.
access to advances in care. Indeed, cancer disparities remain substantial in the United States by geographic area and by socioeconomic status. ${ }^{4-6}$ The socioeconomic cancer disparities are notable in particular because they may worsen over time owing to the soaring costs of cancer diagnosis and treatment. Thus, there are major concerns about US socioeconomic disparities in cancer deaths.

A second major concern is growing evidence of wide disparities in cancer mortality at the county level, with cancer death rates that varied more than 7-fold across US counties in 2014. ${ }^{2}$ At the same time, there is a gap in knowledge about the degree to which socioeconomic factors might underlie the county disparities, or more specifically, the clustering of counties into area hot spots with a disproportionate burden of cancer deaths. Identifying area hot spots can be useful as a guide for public health programs to target the neediest clusters of counties. ${ }^{3}$ The success of such programs, however, depends on understanding the factors that underlie the disparities.

To reduce cancer disparities, it is imperative to understand the degree to which multiple environmental, clinical, and behavioral factors may serve as mediators of the association between county income and cancer mortality. ${ }^{3,7-10}$ There are gaps in knowledge, however, about these key factors. First, it is unclear which of the factors are serving as the strongest mediators of the association between socioeconomic status and cancer mortality. Second, it is unclear whether such factors adequately explain the large differences in cancer mortality between counties. This is important in part because a better understanding of the possible mediators of high death rates among low-income counties is needed to inform future efforts to lessen disparities.

In this context, we studied the association between county-level incomes and cancer death rates. Our aims were to assess the disparities in cancer death rates between low-, medium-, and highincome counties; to assess geographic variation in cancer death rates within and between income groups; to identify factors that may serve as mediators of county disparities; and finally, to compare geographic variations in associated factors with geographic variations in cancer deaths.

## Methods

## Study Design

We conducted a cross-sectional study to assess cancer disparities between counties on the basis of socioeconomic status and to identify clusters of counties with high cancer death rates. Then we used regression models to identify multiple factors associated with the disparities and to test our hypothesis that these factors may serve as mediators of disparities between counties. Finally, to assess geographic variation in possible mediators across counties, we used the sums of the values of the possible mediators to calculate a standardized risk score that we called the disparity risk index. The Yale institutional review board determined this study was exempt from review because we used publicly available data. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.

## Data Sources

We included all counties with cancer death rates available from the year 2014 in a database the Institute for Health Metrics and Evaluation created and published. ${ }^{11}$ To create the database, institute researchers used death record data from the National Center for Health Statistics to generate small area estimates and to eliminate garbage codes, which are implausible or nonspecific causes of death. ${ }^{12}$ As noted in prior studies, this approach is likely to improve the validity of county-level death records as key metrics for identifying specific causes of death. ${ }^{2,12}$

We linked the cancer death rate of each county to the median household income (MHI). We used income data from the 2012 US Census Bureau Small Area Income and Poverty Estimates, thus allowing for a 2-year lag prior to our primary outcome measure from 2014. ${ }^{13}$ We used MHI instead of using employment, education, or a combined socioeconomic indicator because MHI is a widely used, readily available marker of socioeconomic status, ${ }^{14-16}$ and because we expected it to have the most relevance to cancer death rates because of the high financial burden of cancer care.

## Outcomes, Exposures, and Possible Mediators

Our primary outcome was the age-adjusted cancer death rate per 100000 person-years. To identify the factors that may serve as mediators of the association between the exposure and the outcome at the county level, we used time-lagged variables from the Robert Wood Johnson Foundation County Health Rankings conceptual model (eTable 1 in the Supplement). This model systematically evaluates and ranks counties according to a series of health risk factors that are selected on the basis of their validity and their importance in public health. ${ }^{17,18}$ It then groups these health factors into domains: health risk behaviors, clinical care factors, socioeconomic factors, and physical environment factors. ${ }^{18}$ We added a fifth domain to include cancer-relevant health policies-for example, the number of state-level mandates for insurance coverage of cancer care. Last, we added other factors, such as the presence of a comprehensive cancer center nearby, that were not included in the County Health Rankings model but were expected to be relevant to cancer outcomes. ${ }^{19,20} \mathrm{We}$ obtained factors from multiple sources, including the Centers for Disease Control and Prevention, the American Lung Association, the National Cancer Institute, and the American Society of Clinical Oncology (eTable 1 in the Supplement). 21-26

## Statistical Analysis

We stratified counties into low- (lowest quartile), medium- (middle quartiles), and high-income (highest quartile) groups. We used descriptive statistics to determine the mean cancer death rates and the demographic characteristics.

To identify the factors that account for the county disparities, we fit a series of models of possible mediators (eTable 1 in the Supplement includes a list of variables). Mediators are helpful in modeling pathways between exposures and outcomes, as shown in Figure 1. ${ }^{27}$ Furthermore, models can simplify complex pathways by separating direct effects, which are independent of the mediator, from indirect effects, which are dependent on the mediator. In this way, we used the models to identify the intermediary factors (mediators) that help to explain the association between MHI (the exposure) and cancer mortality (the outcome).

To quantify the associations, we used a series of linear regression models. First, we used singlemediator models to assess for changes in parameter estimates after adding possible mediators to a base model of the outcome (cancer death rate, a continuous variable) and the exposure (MHI, a continuous variable that we log transformed). ${ }^{28}$ We selected variables having significant associations in the single mediator models ( 2 -sided a < .05) to be tested in the multiple mediator model. We used backward stepwise elimination to retain the variables that remained significant in the multivariable model. We adjusted for demographic factors including racial and ethnic distributions of county residents (eTable 1 in the Supplement) and we compared the stepwise approach with other approaches (eg, least absolute shrinkage and selection operator) in sensitivity analyses. After identifying the possible mediators, we estimated the indirect effects, which are given by the product of the coefficients of the association between (1) the exposure and the mediators and (2) the

Figure 1. Schema for Models of Mediation


We adapted schemas for single-mediator (A) and multiple-mediator (B) models from Preacher and Hayes. ${ }^{27}$ The subsequent models of mediation can be used to estimate the direct effects of $E$ on $O$ and the indirect effects of $E$ on $O$ that might be attributable to $M$. The subscripts in panel $B$ identify the separate pathways for the multiple mediators (ie, $M_{1}$ vs $M_{2}$ ).
mediators and the outcome. ${ }^{27}$ We used a method called seemingly unrelated regression, which is used to correct for the possibility of correlations between error terms in a series of similar models. ${ }^{29}$ We used the indirect effects to calculate the percentage mediated (ie, the percentage of the exposure-outcome association that can be explained by the possible mediator) and we used bootstrapped standard errors with 5000 repetitions because the errors for indirect effects are often skewed. ${ }^{27}$

We used multivariate normal regression to impute 20 sets of values for variables with greater than $5 \%$ but less than $20 \%$ missingness and excluded variables with greater than $20 \%$ missingness (eTable 2 in the Supplement). In sensitivity analyses, we compared models with or without outlier counties (for example, those with very large or small populations or with >40\% Native American ethnicity). In addition, because of concerns in prior studies about counties with much higher than expected rates of mortality (for example, rates being more than 3 times the interquartile range beyond the quartile values), ${ }^{30}$ we performed sensitivity analyses that excluded these counties.

## Calculating the Disparity Risk Index

To assess the geographic distribution of the possible mediators, we calculated a composite score that we called the disparity risk index (eFigure 1 in the Supplement). Each county's value for the risk index was the weighted sum of the variables' standardized values (or $z$ scores). Thus, we used the risk index to standardize the comparisons of variables across counties.

To assess the geographic variation in the outcome, we mapped the rates of cancer deaths using geocoded templates. We identified area hot spots using the Getis-Ord statistic for spatial autocorrelation with a significance threshold of 2-sided $P<.05 .^{31,32}$ We then mapped the possible mediators (as represented by the disparity risk index) and we identified and then mapped the hot spots for each of the possible mediators. We completed the analysis between October 1, 2016, and July 31, 2017, using Stata statistical software version 14.2 (StataCorp LLC).

## Results

In our total sample of 3135 US counties, the median incomes ranged from $\$ 22126$ to $\$ 121250$ per year. Compared with counties in the high-income group (median income, $\$ 55780$ ), those in the low-income group (median income, \$33 445) had smaller populations in addition to having higher proportions of residents who were non-Hispanic black, lived in rural areas, or reported poor or fair health (Table 1). Cancer death rates varied widely across counties (mean [range] rate, 206.4 [70.7503.5] deaths per 100000 person-years).

We found significant variation in cancer death rates across income groups, with a mean (SD) rate of 229.7 (32.9) deaths per 100000 person-years in low-income counties vs 204.9 (26.3) (difference, 24.8; 95\% CI, 22.4-27.4) and 185.9 (24.4) (difference, 43.8; 95\% Cl, 41.0-46.7) per 100000 person-years in medium- and high-income counties, respectively ( $P$ < 001 for all pairwise comparisons). We found geographic clusters, or hot spots, with the highest cancer death rates in the South, including the Mississippi River Delta, in addition to Appalachia ( $n=507$ counties in hot spots at a threshold of $P<.05$ ) (Figure 2; eFigure 2 in the Supplement). Many of these hot spots were constituted with low-income counties (Figure 2).

Of the 38 county-level factors we assessed in this study, 19 met criteria for inclusion in the multiple-mediator model, and 8 retained significance in the fully adjusted model, suggesting that 8 county-level factors may be serving as mediators of county-level socioeconomic cancer disparities (Table 2). These factors are important because they correlate with the exposure (county-level MHI) and the outcome (cancer death rate) and because they might fit plausibly into the relevant causal pathways. Three of these factors were health risk behaviors (rates of obesity, smoking, and physical inactivity); 2 were clinical care factors (indicators of unaffordable care and low-quality care); 2 were health policies (smoke-free laws and the state Medicaid fee index, which is a state-level ratio of provider payments from Medicaid vs Medicare); and 1 was a health environment factor (food
insecurity, defined as the percentage of the population that lacks a reliable source of food). In aggregate, these factors explained more than four-fifths ( $81.25 \%$ ) of the association between county-level median incomes and cancer death rates.

We found substantial variation in the degree to which each of the factors may be mediating the county-level association between incomes and death rates. The strongest possible mediators were food insecurity ( $\beta=-0.12 ; 95 \% \mathrm{Cl},-0.17$ to -0.08 ), low-quality care ( $\beta=-0.11 ; 95 \% \mathrm{Cl},-0.14$ to -0.09 ), smoking ( $\beta=-0.08 ; 95 \% \mathrm{Cl},-0.10$ to -0.06 ), physical inactivity ( $\beta=-0.08 ; 95 \% \mathrm{Cl},-0.10$ to -0.06 ), and obesity ( $\beta=-0.07 ; 95 \% \mathrm{Cl},-0.09$ to -0.05 ) (Table 2). Per the final model, the percentage mediated may be largest for food insecurity ( $19.1 \%$; $95 \% \mathrm{Cl}, 12.5 \%-26.5 \%$ ), low-quality care ( $17.9 \%$; $95 \% \mathrm{Cl}, 14.0 \%-21.8 \%$ ), smoking ( $12.7 \%$; $95 \% \mathrm{Cl}, 9.4 \%-15.6 \%$ ), and physical inactivity ( $12.2 \% ; 95 \% \mathrm{Cl}, 9.4 \%-15.6 \%$ ), and it may be smallest for smoke-free laws (1.1\%; 95\% CI, 0.5\%-1.6\%) and the Medicaid fee index ( $1.7 \% 95 \% \mathrm{Cl}, 1.6 \%-3.1 \%$ ) (Figure 3).

We summarized the geographic spread of the factors using the disparity risk index. We found stark differences in the index between low-income (mean [SD], 0.64 [0.57]), medium-income (mean [SD], 0.03 [0.48]), and high-income (mean [SD], -0.58 [0.50]) counties, with higher scores indicating higher concentrations of possible mediators. In models adjusted for county income, the disparity index was associated with cancer death rates overall ( $\beta=35.7 ; 95 \% \mathrm{CI}, 34.2-37.2 ; P<.001$ ) and in the subgroup of low-income counties ( $\beta=39.5$; 95\% CI, 36.2-42.8; $P<.001$ ). We used maps of the index to visualize whether groups of low-income counties having high concentrations of possible mediators matched with groups of low-income counties having high concentrations of cancer deaths (Figure 2). In addition, we identified hot spots of adjacent counties with similarly high values for each of the possible mediators, which included 391 counties in hot spots for smoking; 361, obesity; 469, physical inactivity; 391 , food insecurity; 255, the fee index; 468, smoke-free laws; 315 , unaffordable care; and 358 , low-quality care (all $P<.05$ ) (eFigure 2 in the Supplement).

In sensitivity analysis, we found similar estimates in models with or without counties that were outliers in demographic characteristics or cancer death rates. In addition, other approaches to variable selection (eg, least absolute shrinkage and selection operator) led to identical models. Thus, the final multivariable model seemed to provide an accurate and robust estimate of the degree to which the factors may explain the disparities.

| Characteristics | County-Level Income Groups, Mean (SD) |  |  |
| :---: | :---: | :---: | :---: |
|  | Low ( $\mathrm{n}=783$ ) | Medium ( $\mathrm{n}=1568$ ) | High ( $\mathrm{n}=784$ ) |
| Household income, median (IQR), \$ | 33445 (31 253-35 483) | 43010 (40 298-46 306) | 55780 (52 409-63 376) |
| Population, median (IQR), No. | 16996 (9474-28 543) | 26932 (10659-62 800) | 59227 (17 417-187 729) |
| Age $\geq 65 \mathrm{y}$, \% | 17.4 (4.0) | 17.5 (4.4) | 14.6 (3.8) |
| Female, \% | 50.0 (2.9) | 50.0 (2.1) | 50.1 (1.8) |
| Race, \% |  |  |  |
| Non-Hispanic white | 68.3 (24.8) | 81.9 (16.7) | 80.2 (16.5) |
| Non-Hispanic black | 18.0 (21.4) | 6.0 (9.9) | 5.3 (7.9) |
| Hispanic | 8.8 (17.4) | 8.1 (12.1) | 9.0 (10.5) |
| Asian | 0.5 (0.6) | 0.9 (1.3) | 2.7 (4.6) |
| Native American | 3.4 (11.8) | 1.9 (5.6) | 1.4 (4.4) |
| Rural, \% | 72.4 (25.7) | 59.2 (30.2) | 43.6 (32.5) |
| Poverty (income < 100\% federal poverty level), \% | 24.2 (5.9) | 16.2 (4.0) | 10.8 (3.3) |
| Premature deaths per 100000 person-years, No. ${ }^{\text {a }}$ | 489.3 (96.7) | 382.4 (72.7) | 310.6 (61.8) |
| Reported poor or fair health, \% | 23.3 (5.8) | 16.5 (4.8) | 12.8 (3.7) |
| Poor physical health, \% | 4.7 (1.2) | 3.7 (1.0) | 3.2 (0.7) |
| Cancer deaths per 100000 person-years, No. ${ }^{\text {a }}$ | 229.7 (32.9) | 204.9 (26.3) | 185.9 (24.4) |

Abbreviation: IQR, interquartile range.

[^1]
## Discussion

We found substantial socioeconomic disparities in cancer death rates across US counties and identified the factors that may account for more than $80 \%$ of the county-level disparity. By creating a disparity risk index, a composite measure of the factors that may mediate the association between county levels of income and cancer death rates, we identified heterogeneity in the geographic spread of the factors that aligned with the distribution of cancer death rates in low-income counties. Together, these findings suggest there are multiple factors in low-income areas that might be useful targets for actions to ameliorate cancer disparities between counties.

Figure 2. County Rates of Cancer Deaths and Potential Mediators of Disparities in Cancer Death Rates Between Low-, Medium-, and High-Income US Counties All Counties


Lowest-income quartile of counties

C Rates of cancer deaths


D Possible mediators of county disparities


A, Age-adjusted annual cancer death rate. B and D, Values for the disparity risk index, which we used to represent the geographic spread of the key factors that may act as
mediators of county income-related disparities in cancer death rates. C, Income-adjusted annual cancer death rate in each county for the lowest-income quartile of counties.

Our findings are consistent with prior studies of cancer disparities by area markers of socioeconomic status, ${ }^{33}$ such as income, ${ }^{34-36}$ education, ${ }^{36-38}$ and employment. ${ }^{39-42}$ Existing population-level studies are limited, however, by their focus on single risk factors or on single cancer types; their use of death records unadjusted for garbage codes; and their reliance on data registries

Table 2. Potential Mediators of the Association Between County Median Incomes and Cancer Death Rates Identified by Multiple Mediator Model

| Independent Variables ${ }^{\text {a }}$ | $\beta$ Coefficient (95\% CI) | Percentage <br> Mediated ${ }^{\text {b }}$ |
| :---: | :---: | :---: |
| Direct Associations-Possible Mediator Variables |  |  |
| Health behaviors |  |  |
| Obesity | 1.41 (1.15 to 1.66) | NA |
| Smoking | 0.80 (0.63 to 0.98) | NA |
| Physical inactivity | 0.90 (0.66 to 1.13) | NA |
| Clinical care |  |  |
| Unaffordable care | 0.34 (0.17 to 0.52) | NA |
| Low-quality care | 2.22 (1.90 to 2.55) | NA |
| Health policies |  |  |
| Smoke-free laws | -0.33 (-0.49 to -0.17) | NA |
| Medicaid-Medicare fee index | -0.15 (-0.20 to -0.10) | NA |
| Health environment |  |  |
| Food insecurity | 1.12 (0.76 to 1.47) | NA |
| Direct Association-Primary Exposure Variable |  |  |
| Median household income | -0.12 (-0.18 to -0.07) | NA |
| Indirect Associations-Possible Mediator Variables ${ }^{\text {c }}$ |  |  |
| Health behaviors |  |  |
| Obesity | -0.07 (-0.09 to -0.05) | 10.8 |
| Smoking | -0.08 (-0.10 to -0.06) | 12.7 |
| Physical inactivity | -0.08 (-0.10 to -0.06) | 12.2 |
| Clinical care |  |  |
| Unaffordable care | -0.03 (-0.05 to -0.02) | 5.2 |
| Low-quality Care | -0.11 (-0.14 to -0.09) | 17.9 |
| Health policies |  |  |
| Smoke-free laws | -0.01 (-0.01 to -0.003) | 1.1 |
| Medicaid-Medicare fee index | -0.01 (-0.02 to -0.01) | 1.7 |
| Health environment |  |  |
| Food insecurity | -0.12 (-0.17 to -0.08) | 19.1 |
| Total indirect (mediated) | -0.52 (-0.58 to -0.46) | $81.3^{\text {d }}$ |

Figure 3. Percentage Mediated by County-Level Factors in a Multivariable Model of the Association Between County-Level Median Household Incomes and Cancer Death Rates


Abbreviation: NA, not applicable.
${ }^{\text {a }}$ We included the following county demographic factors as independent variables in the model because we expected them to be potential confounders: the racial/ethnic distributions of residents, the percentage of the county that is rural, the percentage female residents, and the percentage non-English-speaking residents. All $P$ values are $<.001$ except for $P=.002$ for the indirect association with smoke-free laws.
${ }^{\mathrm{b}}$ The percentage mediated is the percentage of the total exposure-outcome association that is attributed to the possible mediator. The numerator is the coefficient for the indirect (mediator-outcome) association, and the denominator is the sum of the coefficients (direct and total indirect associations) for the exposure-outcome association.
${ }^{\text {c }}$ Indirect associations are the subsets of the total exposure-outcome association attributed to each mediator. In contrast, direct associations are mediator-outcome associations that are independent from the exposure.
${ }^{d}$ Total values may differ slightly from the sum of the reported values because of rounding.

For the definitions and the sources of the data for the factors that may serve as mediators, see eTable 1 in the Supplement. To calculate the percentage mediated, we used the $\beta$ coefficients from the multiple mediator model reported in Table 2. Error bars represent 95\% Cls.
that omit certain geographic regions. In contrast, we used a death record data set that applied novel small area estimates and garbage code redistribution techniques to a full sample of US counties. We also used a series of models to explain the county disparities by identifying the most important possible mediators.

Health risk behaviors of smoking, obesity, and physical inactivity were the strongest possible mediators of the cancer disparities. This finding is consistent with prior work showing that health risk behaviors are the factors that correlate most strongly with health outcomes. ${ }^{43}$ With regard to cancer outcomes, there is strong evidence in particular for causal pathways that include smoking, ${ }^{7}$ along with evidence that rates of obesity and physical inactivity are strongly associated with cancer death rates. ${ }^{8-10,44}$ Thus, there is strong evidence to support community intervention targeting health risk behaviors. At the same time, there is a need to disrupt external factors that contribute to health risk behaviors in low- and medium-income areas. For example, efforts to increase rates of physical activity at the population level are likely to be inadequate in the absence of efforts to improve the built environments in certain communities. ${ }^{45}$

A second major finding was that food insecurity and the quality of clinical care may be the strongest individual mediators. There are multiple ways in which these 2 factors may account for disparities in cancer deaths. For example, low-quality clinical care may lead to delays in the diagnosis and treatment of cancer, and food insecurity may increase the incidence of certain cancers in populations due to poor nutrition, even if obesity rates are similar. Efforts to target nonbehavioral mediators might be useful in light of evidence that addressing health risk behaviors is necessary but not sufficient if the ultimate goal is to eliminate health disparities. ${ }^{46,47}$ In addition, the issue of regular access to healthy foods, or food security, might warrant further consideration in future studies of cancer disparities. This is because studies suggest that food insecurity is correlated with poor health, ${ }^{48}$ high costs, ${ }^{49}$ and obesity-a key risk factor for cancer. ${ }^{10}$ In addition, because obesity is a risk factor for diabetes and cardiovascular disease, efforts to address it might lead to substantial gains in population-level health outcomes.

To inform local health policies, it is important to generate rigorous evidence about the factors associated with cancer deaths. This is because there is a clear need for efficacy and efficiency in health programs that often are funded locally with limited budgets. Locoregional programs might be particularly helpful in addressing the clustering of cancer deaths in the South and in Appalachia. Because we identified multiple possible mediators, our findings may be useful for local policymakers to develop and evaluate future actions to eliminate cancer disparities.

Finally, to address disparities, it may be critical to maintain policies that are associated with better outcomes in low-income communities. For example, efforts to limit the expansion of Medicaid may undermine efforts to lessen socioeconomic cancer disparities, in part because the states with vs without Medicaid expansions have had larger improvements in screenings for and early detections of cancer. ${ }^{50,51}$ Our study adds to these findings by suggesting that limited access to affordable care is a mediator of cancer disparities, at the same time as many of the low-income counties with the highest mortality rates are in the states that eschewed the expansion. We also found that a disproportionate number of residents were non-Hispanic black Americans in low-income counties with high rates of cancer death. This highlights the need to dismantle structural racism, which contributes to inequalities in social and economic power and to the segregation of black Americans into area hot spots of counties having disproportionately more cancer deaths.

## Limitations

Our study has limitations. First, population studies are prone to interpretive errors, including ecological fallacies. For this reason, the findings from this population-level study should be used to inform population-level, but not individual-level, health programs. Second, because we studied cancer death rates in aggregate, we did not assess for differences in key factors across cancer types, which can be the focus of future work. Third, we were not able to assess for causality or for unmeasured confounding because of the cross-sectional nature of the study. Fourth, although
mixed-effects models are used in the literature to assess the effects of single mediators, ${ }^{52}$ we were unable to find a validated approach to using a mixed-effect model in a study of multiple mediators. In part for this reason, we used a fixed-effects model (eFigure 3 in the Supplement), which means that we cannot exclude the possibility that random effects at the state level might alter the model. Fifth, we did not assess city- or county-level regulations, which limited our assessment of local health policies. Sixth, mediator models may oversimplify associations because of assumptions about directionality and interactions. For this reason, they should be used to generate rather than to confirm hypotheses for causal pathways.

## Conclusions

We found substantial gaps in cancer death rates among low-, medium-, and high-income counties. We explained much of the disparity with health behaviors, clinical care factors, health policies, and health environments. Because of the unacceptable gaps in cancer outcomes across counties that persist despite major advances in cancer care, there is an urgent need for actions to determine whether targeting these factors might ameliorate disparities.

## ARTICLE INFORMATION

Published: October 5, 2018. doi:10.1001/jamanetworkopen.2018.3146
Open Access: This is an open access article distributed under the terms of the CC-BY License. © 2018 O'Connor JM et al. JAMA Network Open.

Corresponding Author: Cary P. Gross, MD, Section of General Internal Medicine, Department of Internal Medicine, Yale University School of Medicine, 333 Cedar St, New Haven, CT 06510 (cary.gross@yale.edu).

Author Affiliations: Yale University School of Medicine, New Haven, Connecticut (O'Connor, Sedghi, Gross); National Clinician Scholars Program, New Haven, Connecticut (O'Connor, Gross); Cancer Outcomes, Public Policy and Effectiveness Research Center, Yale University School of Medicine, New Haven, Connecticut (Sedghi, Gross); The University of Chicago, Chicago, Illinois (Dhodapkar); Yale University School of Public Health, New Haven, Connecticut (Kane).

Author Contributions: Drs O'Connor and Gross had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: O'Connor, Kane, Gross.
Acquisition, analysis, or interpretation of data: O'Connor, Sedghi, Dhodapkar.
Drafting of the manuscript: O'Connor, Sedghi.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: O'Connor, Sedghi, Kane.
Administrative, technical, or material support: Dhodapkar, Gross.
Supervision: Gross.
Conflict of Interest Disclosures: Dr Gross has received support for research from 21st Century Oncology, Johnson \& Johnson, Medtronic, and Pfizer. No other disclosures were reported.
Funding/Support: This study was made possible by the National Center for Advancing Translational Science, a component of the National Institutes of Health (grant TL1 TRO01864).
Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.
Disclaimer: The contents are solely the responsibility of the authors and do not necessarily represent the official view of the National Institutes of Health.

## REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016;66(1):7-30. doi:10.3322/caac. 21332
2. Mokdad AH, Dwyer-Lindgren L, Fitzmaurice C, et al. Trends and patterns of disparities in cancer mortality among US counties, 1980-2014. JAMA. 2017;317(4):388-406. doi:10.1001/jama.2016.20324
3. Wheeler SB, Basch E. Translating cancer surveillance data into effective public health interventions. JAMA. 2017;317(4):365-367. doi:10.1001/jama.2016.20326
4. Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, et al. Inequalities in life expectancy among US counties, 1980 to 2014: temporal trends and key drivers. JAMA Intern Med. 2017;177(7):1003-1011. doi:10.1001/jamainternmed. 2017.0918
5. Hood CM, Gennuso KP, Swain GR, Catlin BB. County health rankings: relationships between determinant factors and health outcomes. Am J Prev Med. 2016;50(2):129-135. doi:10.1016/j.amepre.2015.08.024
6. McGovern L, Miller G, Hughes-Cromwick P; Project HOPE. The relative contribution of multiple determinants to health outcomes. Health Affairs health policy brief. Published August 21, 2014. Accessed September 30, 2017. doi:10.1377/hpb20140821.404487
7. Lortet-Tieulent J, Goding Sauer A, Siegel RL, et al. State-level cancer mortality attributable to cigarette smoking in the United States. JAMA Intern Med. 2016;176(12):1792-1798. doi:10.1001/jamainternmed.2016.6530
8. Massetti GM, Dietz WH, Richardson LC. Excessive weight gain, obesity, and cancer: opportunities for clinical intervention. JAMA. 2017;318(20):1975-1976. doi:10.1001/jama.2017.15519
9. Doubeni CA, Major JM, Laiyemo AO, et al. Contribution of behavioral risk factors and obesity to socioeconomic differences in colorectal cancer incidence. J Nat/ Cancer Inst. 2012;104(18):1353-1362. doi:10.1093/jnci/djs346
10. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med. 2003;348(17):1625-1638. doi:10.1056/NEJMoaO21423
11. Institute for Health Metrics and Evaluation. United States mortality rates by county 1980-2014. http://ghdx. healthdata.org/record/united-states-mortality-rates-county-1980-2014. Updated September 8, 2017. Accessed June 1, 2016.
12. Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, et al. US county-level trends in mortality rates for major causes of death, 1980-2014. JAMA. 2016;316(22):2385-2401. doi:10.1001/jama.2016.13645
13. US Census Bureau. Small Area Income and Poverty Estimates (SAIPE) Program. https://www.census.gov/ programs-surveys/saipe.html. Accessed November 30, 2017.
14. Spatz ES, Beckman AL, Wang Y, Desai NR, Krumholz HM. Geographic variation in trends and disparities in acute myocardial infarction hospitalization and mortality by income levels, 1999-2013. JAMA Cardiol. 2016;1(3):255-265. doi:10.1001/jamacardio.2016.0382
15. Anderson RT, Sorlie P, Backlund E, Johnson N, Kaplan GA. Mortality effects of community socioeconomic status. Epidemiology. 1997;8(1):42-47. doi:10.1097/00001648-199701000-00007
16. Bucholz EM, Ma S, Normand S-LT, Krumholz HM. Race, socioeconomic status, and life expectancy after acute myocardial infarction. Circulation. 2015;132(14):1338-1346. doi:10.1161/CIRCULATIONAHA.115.017009
17. Remington PL, Catlin BB, Gennuso KP. The county health rankings: rationale and methods. Popul Health Metr. 2015;13(1):11. doi:10.1186/s12963-015-0044-2
18. University of Wisconsin Population Health Institute. County health rankings \& roadmaps. http://www. countyhealthrankings.org/explore-health-rankings/rankings-data-documentation. Published 2016. Accessed February 4, 2017.
19. Paulson EC, Mitra N, Sonnad S, et al. National Cancer Institute designation predicts improved outcomes in colorectal cancer surgery. Ann Surg. 2008;248(4):675-686. doi:10.1097/SLA.Ob013e318187a757
20. Onega T, Duell EJ, Shi X, Demidenko E, Goodman D. Determinants of NCI Cancer Center attendance in Medicare patients with lung, breast, colorectal, or prostate cancer. J Gen Intern Med. 2009;24(2):205-210. doi:10. 1007/s11606-008-0863-y
21. National Cancer Institute. State cancer profiles. https://statecancerprofiles.cancer.gov/index.html. Published 2016. Accessed February 4, 2017.
22. Kaiser Family Foundation. Medicaid-to-Medicare fee index. https://www.kff.org/medicaid/state-indicator/ medicaid-to-medicare-fee-index/. Published 2016. Accessed February 7, 2017.
23. American Lung Association. State of Tobacco Control, 2017. http://www.lung.org/our-initiatives/tobacco/ reports-resources/sotc/state-grades/state-rankings/. Accessed February 6, 2017.
24. Hanson K, Bondurant E; National Conference of State Legislatures. Cancer insurance mandates and exceptions. http://www.ncsl.org/portals/1/documents/health/CancerMandatesExcept09.pdf. Published 2009. Accessed August 14, 2018.
25. National Cancer Institute. Find a cancer center. https://cancercenters.cancer.gov/Center/CClist. Published 2017. Accessed February 7, 2017.
26. Kirkwood MK, Bruinooge SS, Goldstein MA, Bajorin DF, Kosty MP. Enhancing the American Society of Clinical Oncology workforce information system with geographic distribution of oncologists and comparison of data sources for the number of practicing oncologists. J Oncol Pract. 2014;10(1):32-38. doi:10.1200/JOP.2013.001311
27. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Methods. 2008;40(3):879-891. doi:10.3758/BRM.40.3.879
28. VanderWeele TJ. Mediation analysis: a practitioner's guide. Annu Rev Public Health. 2016;37:17-32. doi:10. 1146/annurev-publhealth-032315-021402
29. Zellner A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J Am Stat Assoc. 1962;57(298):348-368. doi:10.1080/01621459.1962.10480664
30. Hall JM, Shenkman EA, Fu CS. Outlier in analysis of cancer mortality by US county. JAMA. 2017;317(18): 1910-1911. doi:10.1001/jama.2017.4128
31. Kondo K. Hot and cold spot analysis using Stata. Stata J. 2016;16(3):613-631.
32. Getis A, Ord JK. The analysis of spatial association by use of distance statistics. Geogr Anal. 1992;24(3):

189-206. doi:10.1111/j.1538-4632.1992.tb00261.x
33. Singh GKMB, Hankey BF, Edwards BK. Area Socioeconomic Variations in US Cancer Incidence, Mortality, Stage, Treatment, and Survival, 1975-1999. Bethesda, MD: National Cancer Institute; 2003.
34. Al-Qurayshi Z, Randolph GW, Srivastav S, Kandil E. Outcomes in endocrine cancer surgery are affected by racial, economic, and healthcare system demographics. Laryngoscope. 2016;126(3):775-781. doi:10.1002/ lary. 25606
35. Artinyan A, Mailey B, Sanchez-Luege N, et al. Race, ethnicity, and socioeconomic status influence the survival of patients with hepatocellular carcinoma in the United States. Cancer. 2010;116(5):1367-1377. doi:10.1002/ cncr. 24817
36. Byers TE, Wolf HJ, Bauer KR, et al; Patterns of Care Study Group. The impact of socioeconomic status on survival after cancer in the United States: findings from the National Program of Cancer Registries Patterns of Care Study. Cancer. 2008;113(3):582-591. doi:10.1002/cncr. 23567
37. Singh GK, Miller BA, Hankey BF, Edwards BK. Persistent area socioeconomic disparities in U.S. incidence of cervical cancer, mortality, stage, and survival, 1975-2000. Cancer. 2004;101(5):1051-1057. doi:10.1002/ cncr. 20467
38. Albano JD, Ward E, Jemal A, et al. Cancer mortality in the United States by education level and race. J Natl Cancer Inst. 2007;99(18):1384-1394. doi:10.1093/jnci/djm127
39. Brewer KC, Peterson CE, Davis FG, Hoskins K, Pauls H, Joslin CE. The influence of neighborhood socioeconomic status and race on survival from ovarian cancer: a population-based analysis of Cook County, Illinois. Ann Epidemiol. 2015;25(8):556-563. doi:10.1016/j.annepidem.2015.03.021
40. Breen N, Lewis DR, Gibson JT, Yu M, Harper S. Assessing disparities in colorectal cancer mortality by socioeconomic status using new tools: health disparities calculator and socioeconomic quintiles. Cancer Causes Control. 2017;28(2):117-125. doi:10.1007/s10552-016-0842-2
41. Enewold L. Horner MJ, Shriver CD, Zhu K. Socioeconomic disparities in colorectal cancer mortality in the United States, 1990-2007. J Community Health. 2014;39(4):760-766. doi:10.1007/s10900-014-9824-z
42. Smith ND, Prasad SM, Patel AR, et al. Bladder cancer mortality in the United States: a geographic and temporal analysis of socioeconomic and environmental factors. J Urol. 2016;195(2):290-296. doi:10.1016/j.juro.2015. 07.091
43. Schroeder SA. Shattuck Lecture. We can do better-improving the health of the American people. NEnglJ Med. 2007;357(12):1221-1228. doi:10.1056/NEJMsa073350
44. Song M, Giovannucci E. Preventable incidence and mortality of carcinoma associated with lifestyle factors among white adults in the United States. JAMA Oncol. 2016;2(9):1154-1161. doi:10.1001/jamaoncol.2016.0843
45. Lovasi GS, Hutson MA, Guerra M, Neckerman KM. Built environments and obesity in disadvantaged populations. Epidemiol Rev. 2009;31(1):7-20. doi:10.1093/epirev/mxp005
46. Lantz PM, House JS, Lepkowski JM, Williams DR, Mero RP, Chen J. Socioeconomic factors, health behaviors, and mortality: results from a nationally representative prospective study of US adults. JAMA. 1998;279(21): 1703-1708. doi:10.1001/jama.279.21.1703
47. Nandi A, Glymour MM, Subramanian SV. Association among socioeconomic status, health behaviors, and all-cause mortality in the United States. Epidemiology. 2014;25(2):170-177. doi:10.1097/EDE.
0000000000000038
48. Gundersen C, Ziliak JP. Food insecurity and health outcomes. Health Aff (Millwood). 2015;34(11):1830-1839. doi:10.1377/hlthaff.2015.0645
49. Berkowitz SA, Seligman HK, Rigdon J, Meigs JB, Basu S. Supplemental Nutrition Assistance Program (SNAP) participation and health care expenditures among low-income adults. JAMA Intern Med. 2017;177(11):
1642-1649. doi:10.1001/jamainternmed.2017.4841
50. Wright BJ, Conlin AK, Allen HL, Tsui J, Carlson MJ, Li HF. What does Medicaid expansion mean for cancer screening and prevention? results from a randomized trial on the impacts of acquiring Medicaid coverage. Cancer. 2016;122(5):791-797. doi:10.1002/cncr. 29802
51. Jemal A, Lin CC, Davidoff AJ, Han X. Changes in insurance coverage and stage at diagnosis among nonelderly patients with cancer after the Affordable Care Act. J Clin Oncol. 2017;35(35):3906-3915. doi:10.1200/JCO.2017. 73.7817
52. Krull JL, MacKinnon DP. Multilevel modeling of individual and group level mediated effects. Multivariate Behav Res. 2001;36(2):249-277. doi:10.1207/S15327906MBR3602_06

SUPPLEMENT.
eTable 1. Variables We Evaluated for Testing in Models of Mediation
eFigure 1. Calculating the Disparity Risk Index
eTable 2. Missing Values
eFigure 2. Geographic Distributions (Maps) of "Hot Spots" Using Spatial Autocorrelation Analysis of Data for the Exposure (Median Income), Outcome (Cancer Death Rate), and Possible Mediating Factors
eFigure 3. Fixed Effects vs Mixed Effects Models of the Possible Mediators


[^0]:    色 Open Access.

[^1]:    ${ }^{\text {a }}$ Age adjusted.

