


RESEARCH ARTICLE

Open Access



The association of county-level socioeconomic factors with individual tobacco and alcohol use: a longitudinal study of U.S. adults

Rita Hamad^{1*} , Daniel M. Brown² and Sanjay Basu³

Abstract

Background: Place-based factors have been implicated as root causes of socioeconomic disparities in risky health behaviors such as tobacco and alcohol use. Yet few studies examine the effects of county-level socioeconomic characteristics, despite the fact that social and public health policies are often implemented at the county level. In this study, we tested the hypothesis that county-level socioeconomic disadvantage was associated with individual tobacco and alcohol use.

Methods: The sample included a panel of participants from the National Longitudinal Survey of Youth ($N = 9302$). The primary predictors were three time-varying measures of socioeconomic disadvantage in an individual's county of residence: educational attainment, percent unemployment, and per capita income. We first conducted traditional ordinary least squares (OLS) models, both unadjusted and adjusted for individual-level covariates. We then conducted fixed effects (FE) models to adjust for confounding by unmeasured time-invariant individual-level factors.

Results: OLS and FE models yielded contrasting results: higher county-level per capita income was associated with decreased drinking in OLS models and increased drinking in FE models, while decreased county-level educational attainment was associated with decreased smoking in OLS models and more cigarettes per day in FE models. The findings from FE models suggest that OLS models were confounded by unobserved time-invariant characteristics. Notably, the point estimates for the county-level measures were small, and in many cases they may not represent a clinically meaningful effect except at the population level.

Conclusions: These results suggest that county-level socioeconomic characteristics may modestly influence tobacco and alcohol use. Future work should examine the effects of specific county policies that might explain these findings.

Keywords: Alcohol, Tobacco, Area effects, County effects, Fixed effects, Health disparities, Social epidemiology

Background

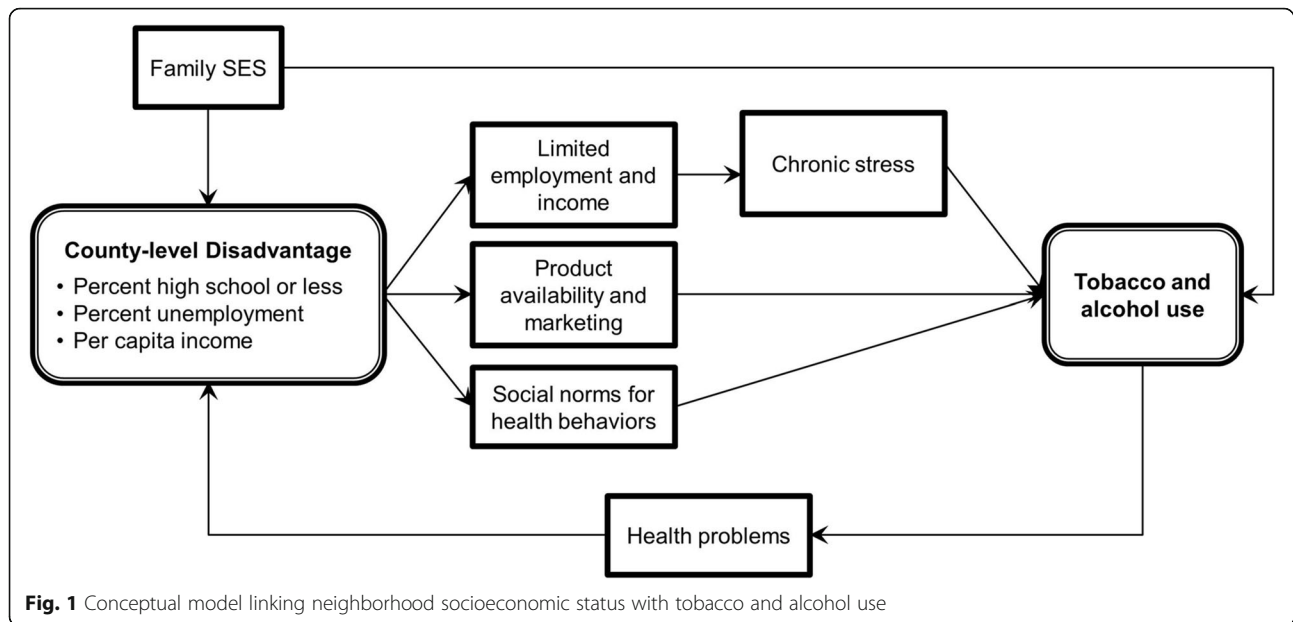
Place-based characteristics have been implicated as determinants of socioeconomic disparities in risky health behaviors, over and above the effects of individual-level socioeconomic status. For example, numerous studies have demonstrated associations between area-level disadvantage—including

measures of poverty, education, employment, or an aggregate index of all three—and tobacco and alcohol use [1–5]. Hypothesized mechanisms linking area-level disadvantage with healthy risk behaviors include limited employment and income, leading to stress and increased substance use [1, 6], the availability of harmful substances, for example through the increased marketing of tobacco products in low-income areas [7, 8], or differences in social norms [9–11] (Fig. 1). Area-level policies—such as taxation or smoking restrictions—may also drive differences in the prevalence of substance use [12–15].

* Correspondence: rita.hamad@ucsf.edu

¹Department of Family and Community Medicine, Philip R. Lee Institute for Health Policy Studies, University of California San Francisco, 995 Potrero Avenue, Building 80, Ward 83, San Francisco, CA 94110, USA
Full list of author information is available at the end of the article





While numerous prior studies have examined the health effects of small areas (e.g., neighborhoods or U.S. census tracts) and large areas (e.g., states, countries) [16–21], few have examined the effects of U.S. county-level characteristics on risky health behaviors [20, 22]. There are roughly 3000 counties in the U.S., and they represent administrative areas that are larger than neighborhoods but smaller than states. Analyses at the county level are important because relevant policies that influence health and the social determinants of health are often implemented by county governments [23]. Examples include county-level “smoke-free” policies that restrict smoking in certain places [24–26], those that limit the sale of specific tobacco products [27, 28], and others that limit the sale of alcohol [29]. Beyond just health policy, counties are also involved in policies that affect the *social and economic determinants* of health behaviors. For example, income is strongly correlated with tobacco and alcohol use [30, 31], and counties are often involved in policies that affect labor markets or other economic factors, like setting a local minimum wage [32].

A recent review concluded that the evidence on the associations between area-level characteristics and individual health behaviors remains inconclusive [33]. In part, prior work may demonstrate inconsistent results because of the use of different measures of disadvantage [33]. More problematically, some prior analyses may suffer from confounding or reverse causation, in that unhealthy individuals may be more likely to move into disadvantaged areas [34, 35]. Simple adjustment for observed individual-level covariates is unlikely to adequately control for this confounding, such that more rigorous study designs may be needed. Previous studies using more sophisticated statistical methods—including fixed effects (FE) and marginal structural models—find persistent

associations between area-level deprivation and tobacco and alcohol use [36, 37]. Yet these studies were not conducted using population-level U.S. data, which limits their generalizability, and they employed limited measures of area-level disadvantage.

In this study, we estimated the association between county-level characteristics and health behaviors using FE models, which more rigorously adjust for confounding relative to standard statistical techniques used in the prior literature. We employed a large nationally representative U.S. sample to test the hypothesis that greater county-level socioeconomic disadvantage is associated with increased risky health behaviors, even after adjusting for individual-level socioeconomic status. In addition, we examined multiple indices of area-level disadvantage. Determining the contributions of county-level socioeconomic characteristics to disparities in risky health behaviors has important implications for directing policy budgets towards effective interventions.

Methods

Data set

We used data from the 1979 National Longitudinal Survey of Youth (NLSY). The NLSY is a nationally representative longitudinal panel study of 12,686 men and women in the United States enrolled when they were 14–22 years old in 1979. It was conducted annually during 1979–1994 and biennially thereafter, via in-person interviews. Questions regarding the health outcomes of interest were included in surveys beginning in 1992 for smoking, and in 1994 for alcohol use. We restricted the sample to individuals who answered questions related to the health outcomes of interest in at least the first time period, and who

lived in counties for which county-level socioeconomic data were available. This resulted in a sample of 9302 individuals in 2117 counties. Additional details on the NLSY are provided elsewhere [38].

Individual-level covariates

Time-invariant characteristics included gender and race. Time-varying covariates included educational attainment, marital status, number of children in the household, annual total household income in inflation-adjusted U.S. dollars, and the number of weeks of unemployment in the last year. For the latter two variables, the natural logarithm was taken because of right-skewness. All models also included fixed effects (i.e., indicator variables) for year to account for secular trends.

County-level disadvantage

We constructed three variables to capture the level of disadvantage in each individual's county of residence in a given year: (1) educational attainment, i.e., the percent of people in a county with a high school education or less, (2) percent unemployment, and (3) inflation-adjusted per capita personal income. Each of these has been previously associated with substance use in correlational studies [2, 37, 39, 40]. These measures were obtained from online national public data sources [41–44]. The three time-varying exposure variables were then linked to NLSY respondents based on their county of residence during each survey wave. Figure 2 shows the variation in county-level disadvantage in 1992, the beginning of the study period.

Health behavior outcomes

Two alcohol-related outcomes were constructed using NLSY survey questions (Table 1), including the number of alcoholic drinks consumed in a typical day in the last month, and whether an individual consumed at least six drinks in a single day in the last month. We refer to the latter as binge drinking for conciseness, as it roughly corresponds to the term established by the U.S. National Institute on Alcohol Abuse and Alcoholism: four or more drinks per day for women and five or more drinks for men [45]. The NLSY does not contain a question that captures this standard definition of binge drinking. We also constructed two smoking-related variables, including the number of cigarettes smoked per day in the last month, and whether an individual was a current smoker in the last month.

Multiple imputation

We conducted multiple imputation using chained equations to impute missing predictor variables from the NLSY. The percentage of missing values ranged from roughly 2% for weeks of unemployment to about 30% for household income. We assumed that values were “missing at random,” rather than “missing completely at random” [46]. This imputation method does not assume that variables are normally distributed, and can therefore be employed for categorical and binary variables. The data were imputed in wide form, to allow for correlations between observations of the same individual in different years. All variables used in the analytic models were included in the imputation models, including

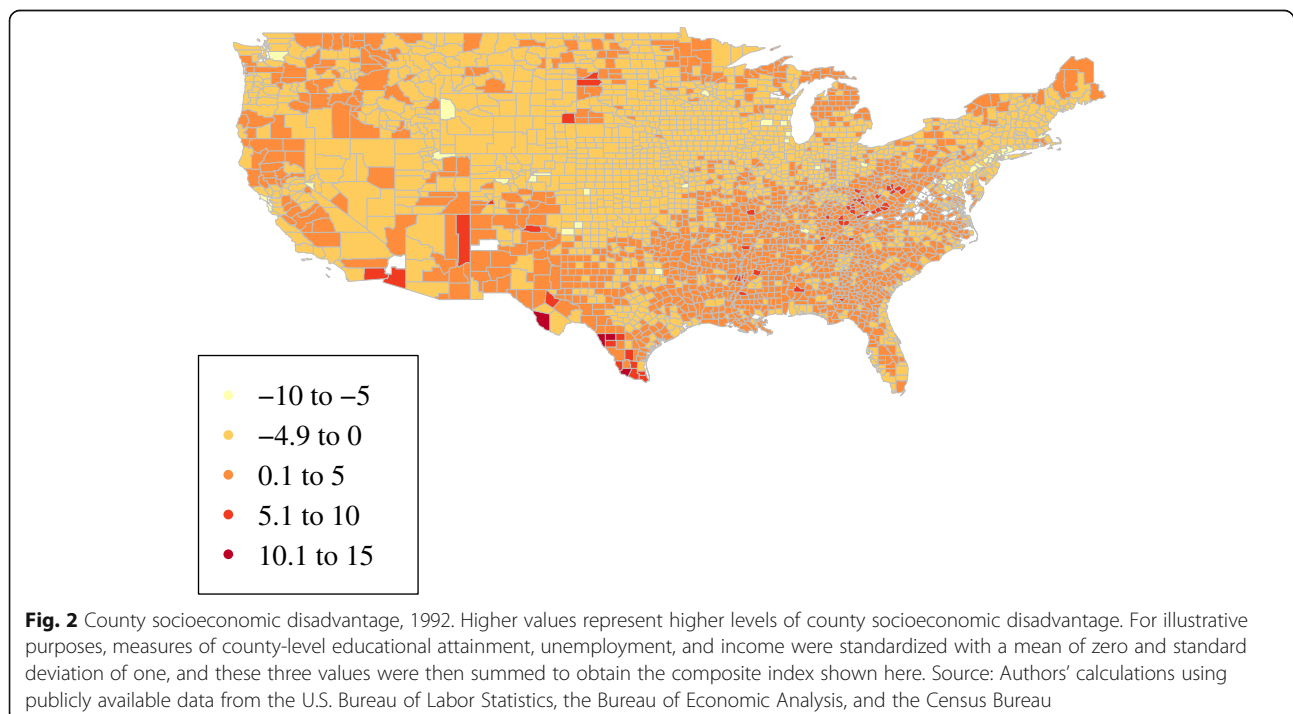


Table 1 Health Outcomes of Interest, U.S. National Longitudinal Study of Youth, 1992–2012

Variable	Type	Survey Years Including Outcome
Binge drinking in last month	Binary	1994, 2002, 2006, 2008, 2010, 2012
Typical number of drinks per day (zero for non-drinkers)	Continuous	1994, 2002, 2006, 2008, 2010, 2012
Smoked in last month	Binary	1992, 1994, 1998, 2008, 2010, 2012
Typical number of cigarettes per day (zero for non-smokers)	Continuous	1992, 1994, 1998, 2008, 2010, 2012

outcome variables, in order to improve the prediction of missing covariates. We did not use imputed values of the outcome variables in the analyses, however, as this is likely to add noise to subsequent estimates [47]. This resulted in differing numbers of observations for analyses examining each of the different health outcomes. We produced 20 imputations per observation, which is sufficient to ensure reproducibility between successive analytic runs [48].

As a sensitivity analysis, we also conducted the models described below using the complete cases, i.e., excluding observations with missing values.

Data analysis

We employed two types of models in this study. First, we conducted standard ordinary least squares (OLS) models to examine the association between health behaviors and county-level disadvantage. We then carried out individual-level fixed effects (FE) models, which adjust for time-invariant confounding and therefore captured the effects of “within-person” changes in county-level disadvantage. FE models represent an improvement over OLS models in that they compare each individual with herself at different time points, rather than comparing different individuals to one another. This amounts to adding a separate intercept for each individual, thereby controlling for any unobserved characteristics that are constant over time [49]. The main drawback of FE methods is that they rely on multiple observations per person; studies that only include a single measurement for a given individual cannot leverage this technique. In this study, we employed both techniques to investigate whether methodological differences may explain heterogeneity in the prior literature.

Logistic regressions with FE were not feasible due to the sheer number of parameters and the failure of these models to converge. We therefore report the results of linear regressions for continuous outcomes as well as binary outcomes (i.e., linear probability models). As a sensitivity analysis, we carried out logistic regressions for binary outcomes in the OLS models, and these resulted in findings that were similar in magnitude and statistical significance to our primary findings (results available upon request).

Ordinary least squares models

We first conducted multivariable regressions to examine the association between each of the four outcome variables and the three measures of county-level disadvantage. We fit two sets of models: the first included only the three measures of disadvantage (unadjusted), while the second also included the time-variant and time-invariant individual-level covariates listed above (adjusted).

Because standard errors between observations may be correlated over time, we employed Huber-White robust standard errors clustered at the individual level to account for potential heteroscedasticity [50], analogous to generalized estimating equations. Multi-level models (also known as hierarchical models) are primarily useful when the question of interest is decomposition of the variance at multiple levels of analysis [51], which was not our research question of interest.

Fixed effects models

We next conducted multivariable linear regressions, now with the inclusion of FE at the individual level. This accounted for confounding by unmeasured time-invariant characteristics of the individual and their contemporaneous county of residence. We carried out two sets of models, with and without adjusting for the time-varying individual-level covariates listed above. Robust standard errors were again clustered at the individual level to account for correlated observations.

Secondary analysis

Because of potential lagged effects of county-level socioeconomic characteristics on health behaviors, we also carried out an analysis in which the primary exposures were unemployment rates, per capita income, and educational attainment in an individual's county-of-residence during the *prior* survey wave. We conducted these analyses using OLS and FE models. These analyses were otherwise similar to our primary models, including adjustment for covariates and clustering of standard errors.

Results

Sample characteristics

The sample was diverse with respect to gender, educational attainment, and race (Table 2). The sociodemographic characteristics of those living in socioeconomically

Table 2 Sample Characteristics by County Disadvantage Level, U.S. National Longitudinal Study of Youth, 1992–2012 (N = 9302)

	County Not Socioeconomically Disadvantaged	County Socioeconomically Disadvantaged	Test of proportions or t-test (<i>p</i> -value)
Sociodemographic Characteristics			
Female (%)	51.2	51.9	<0.01
Education (%)			
Less than high school	14.00	18.8	<0.01
High school	36.3	40.4	<0.01
Some college	24.7	25.0	<0.01
College or more	24.9	15.7	<0.01
Race (%)			
White/other	52.8	41.5	<0.01
Black	30.9	32.4	<0.01
Hispanic	16.3	26.1	<0.01
Married (%)	66.5	64.4	<0.01
Children in household (mean (SD))	1.2 (1.2)	1.2 (1.3)	0.20
Annual household income (USD, mean (SD))	59,956 (61,500)	49,119 (53,005)	<0.01
Weeks unemployed in last year (mean (SD))	2.3 (8.2)	3.0 (9.7)	<0.01
Health Behaviors			
Tobacco-related outcomes			
Smoker (%)	30.0	30.2	0.16
Cigarettes per day (mean (SD))	4.2 (8.6)	4.1 (8.7)	<0.01
Alcohol-related outcomes			
Binge drinker (%)	18.4	17.8	<0.01
Drinks per day (mean (SD))	1.8 (2.3)	1.7 (2.5)	<0.01

Note: Values for sociodemographic characteristics are based on imputed data. Income was inflation-adjusted. For illustrative purposes to summarize sample characteristics, measures of county-level disadvantage (i.e., educational attainment, unemployment, and income) were each standardized with a mean of zero and standard deviation of one, these three values were summed to obtain a composite index, and this index was then split at the median

disadvantaged counties were statistically significantly different from those living in non-disadvantaged counties. Those in disadvantaged counties were more likely to be female, non-white, and unmarried, and were more likely to have lower educational attainment, lower household income, and more weeks unemployed in the last year. In terms of health behaviors, those living in disadvantaged counties smoked fewer cigarettes on average. They were less likely to be binge drinkers, and consumed fewer drinks per day.

Ordinary least squares models

For unadjusted OLS models (Table 3), increased county-level unemployment was associated with decreased smoking, fewer cigarettes per day, and more drinks per day. Increased county-level per capita income was associated with decreased smoking, fewer cigarettes per day, and less binge drinking. Lower county-level educational attainment was associated with less smoking. Results were largely similar in adjusted OLS models (Table 3),

although unemployment was no longer statistically significantly associated with drinks per day.

Analyses using complete cases yielded results similar to those obtained with imputed data (results available upon request).

Fixed effects models

In unadjusted FE models (Table 4), increased county-level unemployment was associated with decreased smoking, fewer cigarettes per day, and more drinks per day. Increased county-level per capita income was associated with higher rates of binge drinking and more drinks per day (both contradictory to OLS findings). Results were similar in adjusted FE models for unemployment, and additionally, lower county-level educational attainment was associated with more cigarettes per day (again contradictory to OLS findings).

Analyses using complete case data yielded results similar to those obtained with imputed data (results available upon request).

Table 3 Ordinary Least Squares Analysis of the Association between County-Level Characteristics and Individual Health Behaviors, U.S. National Longitudinal Study of Youth, 1992–2012

	β Coefficient [95% CI]							
	Smoker Unadjusted	Smoker Adjusted	Daily Cigarettes Unadjusted	Daily Cigarettes Adjusted	Binge Drinking Unadjusted	Binge Drinking Adjusted	Daily Drinks Unadjusted	Daily Drinks Adjusted
County characteristics								
Unemployment rate	-0.0055*	-0.0053*	-0.14*	-0.085*	-0.00064	-0.0011	0.021*	0.012
	[-0.0080, -0.0031]	[-0.0077, -0.0029]	[-0.19, -0.094]	[-0.13, -0.039]	[-0.0031, 0.0018]	[-0.0036, 0.0014]	[0.0037, 0.037]	[-0.0055, 0.030]
Income (per \$1000)	-0.0020*	-0.00080*	-0.055*	-0.030*	-0.00088*	-0.00070*	-0.0025	-0.0019
	[-0.0028, -0.0013]	[-0.0015, -0.00009]	[-0.067, -0.042]	[-0.042, -0.018]	[-0.0015, -0.00031]	[-0.0013, -0.00014]	[-0.0055, 0.00054]	[-0.0049, 0.0010]
% < high school	-0.00090*	-0.00099*	-0.0085	-0.0014	0.00025	0.00024	-0.00017	0.000077
	[-0.0016, -0.00023]	[-0.0017, -0.00028]	[-0.019, 0.0019]	[-0.012, 0.0098]	[-0.00032, 0.00081]	[-0.00036, 0.00083]	[-0.0033, 0.0029]	[-0.0032, 0.0033]
Female		-0.021*		-0.67*		-0.15*		-0.83*
		[-0.037, -0.0051]		[-0.97, -0.38]		[-0.17, -0.14]		[-0.90, -0.77]
Race (ref white/other)								
Black		-0.048*		-2.76*		-0.066*		-0.52*
		[-0.067, -0.029]		[-3.12, -2.40]		[-0.079, -0.053]		[-0.59, -0.44]
Hispanic		-0.12*		-3.64*		-0.015		0.015
		[-0.14, -0.096]		[-4.03, -3.25]		[-0.032, 0.0018]		[-0.086, 0.12]
Ln(Household income)		-0.0087*		-0.074*		0.0063*		0.055*
		[-0.012, -0.0057]		[-0.12, -0.025]		[0.0045, 0.0080]		[0.044, 0.066]
Education (ref < HS)								
High school		-0.15*		-2.84*		-0.013		-0.16*
		[-0.18, -0.13]		[-3.41, -2.28]		[-0.031, 0.0046]		[-0.28, -0.030]
Some college		-0.22*		-4.02*		-0.054*		-0.36*
		[-0.24, -0.19]		[-4.58, -3.47]		[-0.073, -0.034]		[-0.49, -0.24]
College or more		-0.36*		-6.57*		-0.092*		-0.54*
		[-0.39, -0.33]		[-7.13, -6.00]		[-0.11, -0.072]		[-0.67, -0.41]
Married		-0.074*		-0.96*		-0.053*		-0.37*
		[-0.094, -0.055]		[-1.30, -0.62]		[-0.068, -0.039]		[-0.45, -0.29]
Ln(Weeks unemployed)		0.036*		0.62*		0.011*		0.090*
		[0.030, 0.043]		[0.49, 0.74]		[0.0063, 0.016]		[0.058, 0.12]
Number of children		-0.016*		-0.26*		-0.0097*		-0.035*
		[-0.022, -0.011]		[-0.37, -0.16]		[-0.014, -0.0052]		[-0.062, -0.0092]

Table 3 Ordinary Least Squares Analysis of the Association between County-Level Characteristics and Individual Health Behaviors, U.S. National Longitudinal Study of Youth, 1992–2012 (Continued)

	β Coefficient [95% CI]							
	Smoker Unadjusted	Smoker Adjusted	Daily Cigarettes Unadjusted	Daily Cigarettes Adjusted	Binge Drinking Unadjusted	Binge Drinking Adjusted	Daily Drinks Unadjusted	Daily Drinks Adjusted
No. Observations	42,996	42,996	40,735	40,735	39,402	39,402	44,272	44,272
No. Individuals	9246	9246	9206	9206	8943	8943	9064	9064

* $p < 0.05$. HS = high school. County-level characteristics included annual unemployment rate, inflation-adjusted per capita personal income, and percent with less than a high school education. Analyses were conducted using multivariable linear regressions with imputed data. Linear probability models were used for binary outcomes. Additional controls included fixed effects for year

Secondary analyses

For adjusted OLS models using lagged exposures (Additional file 1: Table S1), increased county-level unemployment was associated with decreased smoking and fewer cigarettes per day, as in our primary models. Increased county-level per capita income was associated with decreased smoking, fewer cigarettes per day, and less binge drinking, as in our primary models, as well as fewer drinks per day. Lower county-level educational attainment was associated with increased binge drinking and drinks per day, neither of which was statistically significant in our primary models.

In adjusted FE models using lagged exposures (Additional file 2: Table S2), increased county-level unemployment was associated with decreased smoking and fewer cigarettes per day, as in our primary models, although drinks per day was no longer statistically significant. Increased county-level per capita income was associated with higher rates of binge drinking as in our primary models, and drinks per day was no longer statistically significant. There was no association between county-level educational attainment and health behaviors.

Discussion

In this study, we investigated how three measures of county-level socioeconomic disadvantage were associated with individual tobacco and alcohol use, using a large longitudinal nationally representative U.S. data set. In both OLS and FE models, higher unemployment rates were associated with less smoking and more drinks per day. Yet OLS and FE models gave contrasting results for the other county-level socioeconomic measures: higher county-level per capita income was associated with decreased drinking in OLS models and increased drinking in FE models, while decreased area-level educational attainment was associated with decreased smoking in OLS models and more cigarettes per day in FE models. Results for lagged models were similar, which may be because socioeconomic characteristics in a given county are correlated over time. The findings from the FE models suggest that OLS models are confounded by unobserved time-invariant individual-level characteristics. Of note, the point estimates for each of our analyses

were very small, and in many cases may not represent a meaningful effect except at the population level.

These findings suggest that interventions to address the social and economic determinants of health at the population level may influence levels of tobacco and alcohol use, thereby improving population health. Prior work has shown that policies at the state level in the U.S. are associated with improvements in child health and chronic disease [17, 52–54], although research on county-level policies is limited [32]. Future studies should specifically examine the impacts of newly implemented county policies that may affect the socioeconomic determinants of health behaviors, to determine whether the associations that we observed in this study may represent causal effects. For example, a recent systematic review of studies across international settings suggested that increased minimum wage policies reduce smoking [55]; additional work is needed to examine whether these results extend to recent county-level minimum wage increases or other similar policies in the U.S.

Our study suggests that the choice of methodology may be driving some of the inconsistencies in the existing literature in this field. The prior literature has relied primarily on statistical methods similar to our OLS models. These studies have been inconsistent, such that increased area-level disadvantage has been associated with both increased and decreased smoking and alcohol use, while others find no association [2, 6, 39, 56–58]. At the same time, prior studies using FE and marginal structural models have found persistent associations of area-level poverty with smoking and alcohol use [36, 37]. Randomized studies in this field are challenging due to logistical and ethical difficulties, although a handful exist. One randomized study found that poor individuals assigned to low-poverty neighborhoods had lower rates of short-term alcohol abuse [59], while another found no long-term impacts on risky healthy behaviors among youth whose families were randomly assigned housing vouchers [60]. Unsurprisingly, a recent systematic review found that the research on place-based effects on health behaviors is inconclusive [33]. Our findings suggest that future meta-analyses should pay special attention to the methods of included studies as a way of explaining contradictory findings.

Table 4 Fixed Effects Analysis of the Association between County-Level Characteristics and Individual Health Behaviors, U.S. National Longitudinal Study of Youth, 1992–2012

	β Coefficient [95% CI]							
	Smoker Unadjusted	Smoker Adjusted	Daily Cigarettes Unadjusted	Daily Cigarettes Adjusted	Binge Drinking Unadjusted	Binge Drinking Adjusted	Daily Drinks Unadjusted	Daily Drinks Adjusted
County characteristics								
Unemployment rate	-0.0036*	-0.0034*	-0.056*	-0.053*	-0.000045	0.00020	0.032*	0.026*
	[-0.0060, -0.0012]	[-0.0058, -0.00092]	[-0.098, -0.013]	[-0.098, -0.0084]	[-0.0030, 0.0029]	[-0.0029, 0.0033]	[0.0082, 0.055]	[0.0025, 0.050]
Income (per \$1000)	-0.00052	-0.00056	-0.0049	-0.0069	0.00095*	0.00077	0.0062*	0.0042
	[-0.0012, 0.00018]	[-0.0013, 0.00017]	[-0.017, 0.0074]	[-0.020, 0.0063]	[0.00008, 0.0018]	[-0.00013, 0.0017]	[0.00079, 0.012]	[-0.0013, 0.0097]
% < high school	-0.00020	-0.00015	0.0068	0.0086*	-0.000033	-0.000060	-0.0020	-0.0015
	[-0.00072, 0.00032]	[-0.00070, 0.00040]	[-0.0010, 0.015]	[0.00013, 0.017]	[-0.00059, 0.00052]	[-0.00065, 0.00053]	[-0.0053, 0.0012]	[-0.0049, 0.0019]
Ln(Household income)		-0.00023		0.048*		0.0059*		0.050*
		[-0.0024, 0.0019]		[0.0057, 0.091]		[0.0039, 0.0079]		[0.037, 0.064]
Education (ref < HS)								
High school		-0.0015		-0.10		-0.0063		0.045
		[-0.025, 0.022]		[-0.57, 0.36]		[-0.029, 0.017]		[-0.16, 0.25]
Some college		0.000067		-0.14		-0.015		0.051
		[-0.027, 0.027]		[-0.68, 0.39]		[-0.042, 0.012]		[-0.17, 0.28]
College or more		0.0013		0.34		-0.0042		0.14
		[-0.035, 0.038]		[-0.34, 1.02]		[-0.043, 0.034]		[-0.14, 0.43]
Married		-0.031*		-0.49		-0.044*		-0.21*
		[-0.047, -0.015]		[-0.80, -0.18]		[-0.065, -0.022]		[-0.34, -0.085]
Ln(Weeks unemployed)		0.0060*		0.12*		0.0017		0.032*
		[0.0021, 0.0099]		[0.046, 0.20]		[-0.0030, 0.0064]		[0.00084, 0.064]
Number of children		-0.0016		0.066		0.0044		0.0094
		[-0.0062, 0.0030]		[-0.018, 0.15]		[-0.0013, 0.010]		[-0.026, 0.045]
No. Observations	42,996	42,996	40,735	40,735	39,402	39,402	44,272	44,272
No. Individuals	9246	9246	9206	9206	8943	8943	9064	9064

* $p < 0.05$. HS = high school. County-level characteristics included annual unemployment rate, inflation-adjusted per capita personal income, and percent with less than a high school education. Analyses were conducted using multivariable linear regressions with imputed data, including fixed effects at the individual level to adjust for time-invariant individual characteristics. Linear probability models were used for binary outcomes. Additional controls included fixed effects for year

Our study has several strengths. We use more rigorous longitudinal statistical techniques—i.e., fixed effects models—to overcome the confounding and reverse causation present in prior work in this field. Our use of a nationally representative data set also means that our results are more generalizable than prior studies that examined limited

geographic areas. We also provide evidence on the effects of county-level measures of disadvantage, which are less frequently examined relative to place-based studies of smaller or larger geographic areas (e.g., U.S. census tracts or states). Relatedly, public health research departments and foundations have begun to support initiatives like

the County Health Rankings to create metrics of county-level differences in health disparities [61], recognizing the importance of county-level determinants of population health.

Our study has several limitations. First, there may be measurement error in self-reported individual characteristics, as well as reporting biases related to frequency of substance use. Second, while both OLS and FE models adjust for time-varying confounding on observed characteristics, there may be confounding on unobserved factors; these might include time-varying aspects of individual or family socioeconomic status not captured by existing variables, or time-varying county and state characteristics that might influence both county disadvantage and individual health (e.g., minimum wage policies or alcohol prices). Consequently, we would not interpret these findings as causal estimates. Nevertheless, FE models represent an improvement over standard OLS modeling techniques, which fail to consider time-invariant confounding and which have dominated the area effects literature [62]. Also, county-level socioeconomic measures beyond the three we examined here are generally not available during this time period for the entire country; however, future studies could seek to compile a richer set of county-level predictors. Finally, one can imagine many interventions to improve health behaviors by addressing individual- and county-level disadvantage, representing a violation of the consistency assumption in causal inference. Absent an exogenous intervention or natural experiment, observational studies can only obliquely inform such strategies [63]. Nevertheless, this avenue of research should be considered one component of a pluralistic approach to triangulate the effects of place-based factors on health [64].

Conclusions

Our findings highlight the challenge in disentangling the effects of county-level socioeconomic disadvantage on risky health behaviors, suggesting that methodological differences may explain some of the inconsistencies in the existing literature in this field. Few studies have implemented multiple statistical methods to disentangle these complex relationships. It is rare that place-based exposures can be randomized, and consequently, there is sparse inconsistent evidence that policymakers and advocates might use to design interventions to address the contextual determinants of risky health behaviors. While some have called for greater reliance on experimental studies [65], these are typically expensive and logistically or ethically unfeasible. Alternatives include increased attention to the use of more rigorous statistical methods and the identification of natural experiments, some of which suggest that area-level socioeconomic disadvantage influences health outcomes [66]. With the increasing availability of longitudinal and linked data, we are hopeful that our study contributes to a greater understanding of these pathways to guide future interventions.

Additional files

Additional file 1: Table S1. Ordinary Least Squares Analysis of the Association between Lagged County-Level Characteristics and Individual Health Behaviors, U.S. National Longitudinal Study of Youth, 1992–2012. (DOCX 23 kb)

Additional file 2: Table S2. Fixed Effects Analysis of the Association between Lagged County-Level Characteristics and Individual Health Behaviors, U.S. National Longitudinal Study of Youth, 1992–2012. (DOCX 23 kb)

Abbreviations

FE: Fixed effects; NLSY: National Longitudinal Survey of Youth; OLS: ordinary least squares

Acknowledgements

Not applicable.

Funding

Dr. Hamad is supported by the National Heart Lung and Blood Institute (K08 HL132106). Dr. Basu is supported by the National Institute on Minority Health and Health Disparities (DP2 MD010478). Publication made possible in part by support from the UCSF Open Access Publishing Fund. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Bureau of Labor Statistics. The funders had no role in study design, data collection, data analysis or interpretation, decision to publish, or preparation of the manuscript for publication.

Availability of data and materials

This research was conducted with restricted access to Bureau of Labor Statistics data. Restricted NLSY data can be obtained by applying at the study site of the U.S. Bureau of Labor Statistics: <https://www.bls.gov/nls>. The measures of county socioeconomic disadvantage are publicly available from the U.S. Bureau of Labor Statistics (<https://www.bls.gov/data>), the Bureau of Economic Analysis (<https://www.bea.gov/data>), and the Census Bureau (<https://www.census.gov/data.html>) [41–44].

Authors' contributions

RH contributed the conception and design of the study and the acquisition of data, conducted the analysis, interpreted the data, and prepared drafts of the manuscript for publication. DMB contributed to the analytic plan, interpreted the data, and revised the manuscript critically for intellectual content. SB contributed to the conception and design of the study, interpreted the data, and revised the manuscript critically for intellectual content. All authors gave approval of the final version of the manuscript and agree to be accountable for all aspects of the work.

Ethics approval and consent to participate

Ethics approval was provided by the Stanford Institutional Review Board (protocol #31603). Ethics approval for the NLSY was provided by the institutional review boards of Ohio State University and the National Opinion Research Center at the University of Chicago, and by the U.S. Office of Management and Budget. Written informed consent was obtained from study subjects for their participation in the NLSY.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Family and Community Medicine, Philip R. Lee Institute for Health Policy Studies, University of California San Francisco, 995 Potrero Avenue, Building 80, Ward 83, San Francisco, CA 94110, USA. ²Environmental

Health Sciences Division, University of California Berkeley, Berkeley, California, USA. ³Department of Medicine, Stanford University, Palo Alto, California, USA.

Received: 4 July 2018 Accepted: 25 March 2019

Published online: 11 April 2019

References

- Boardman JD, Finch BK, Ellison CG, Williams DR, Jackson JS. Neighborhood disadvantage, stress, and drug use among adults. *J Health Soc Behav.* 2001; 42(2):151–65.
- Galea S, Ahern J, Tracy M, Vlahov D. Neighborhood income and income distribution and the use of cigarettes, alcohol, and marijuana. *Am J Prev Med.* 2007;32(6, Supplement):S195–202.
- Cubbin C, Sundquist K, Ahlén H, Johansson S-E, Winkleby MA, Sundquist J. Neighborhood deprivation and cardiovascular disease risk factors: protective and harmful effects. *Scand J Soc Med.* 2006;34(3):228–37.
- Hastert TA, Ruterbusch JJ, Beresford SAA, Sheppard L, White E. Contribution of health behaviors to the association between area-level socioeconomic status and cancer mortality. *Soc Sci Med.* 2016;148:52–8.
- Shareck M, Kestens Y, Frohlich KL. Moving beyond the residential neighborhood to explore social inequalities in exposure to area-level disadvantage: results from the interdisciplinary study on inequalities in smoking. *Soc Sci Med.* 2014;108(0):106–14.
- Steptoe A, Feldman PJ. Neighborhood problems as sources of chronic stress: development of a measure of neighborhood problems, and associations with socioeconomic status and health. *Ann Behav Med.* 2001;23(3):177–85.
- Bluthenthal RN, Cohen DA, Farley TA, Scribner R, Beighley C, Schonlau M, Robinson PL. Alcohol availability and neighborhood characteristics in Los Angeles, California and southern Louisiana. *J Urban Health.* 2008;85(2):191–205.
- Bryden A, Roberts B, McKee M, Petticrew M. A systematic review of the influence on alcohol use of community level availability and marketing of alcohol. *Health Place.* 2012;18(2):349–57.
- Galea S, Nandi A, Vlahov D. The social epidemiology of substance use. *Epidemiol Rev.* 2004;26(1):36–52.
- Ahern J, Galea S, Hubbard A, Syme SL. Neighborhood smoking norms modify the relation between collective efficacy and smoking behavior. *Drug Alcohol Depend.* 2009;100(1–2):138–45.
- Musick K, Seltzer JA, Schwartz CR. Neighborhood norms and substance use among teens. *Soc Sci Res.* 2008;37(1):138–55.
- Wakefield MA, Durkin S, Spittal MJ, Siahpush M, Scollo M, Simpson JA, Chapman S, White V, Hill D. Impact of tobacco control policies and mass media campaigns on monthly adult smoking prevalence. *Am J Public Health.* 2008;98(8):1443–50.
- Wilson LM, Avila Tang E, Chandler G, Hutton HE, Odelola OA, Elf JL, Heckman-Stoddard BM, Bass EB, Little EA, Haberl EB. Impact of tobacco control interventions on smoking initiation, cessation, and prevalence: a systematic review. *J Environ Public Health* 2012, Article ID 961724.
- 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(6):352–65.
- Wagenaar AC, Salois MJ, Komro KA. Effects of beverage alcohol price and tax levels on drinking: a meta-analysis of 1003 estimates from 112 studies. *Addiction.* 2009;104(2):179–90.
- 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(6):444–52.
- Hamad R, Rehkopf DH. Poverty and child development: a longitudinal study of the impact of the earned income tax credit. *Am J Epidemiol.* 2016;183(9):775–84.
- Stuckler D, Basu S, Suhrcke M, Coutts A, McKee M. The public health effect of economic crises and alternative policy responses in Europe: an empirical analysis. *Lancet.* 2009;374:315–23.
- Clark CR, Williams DR. Understanding county-level, cause-specific mortality: the great value—and limitations—of small area data. *JAMA.* 2016;316(22):2363–5.
- Kim D, Subramanian SV, Gortmaker SL, Kawachi I. US state- and county-level social capital in relation to obesity and physical inactivity: a multilevel, multivariable analysis. *Soc Sci Med.* 2006;63(4):1045–59.
- Hamad R, Rehkopf D, Kuan K, Cullen M. Predicting later life health status and mortality using state-level socioeconomic characteristics in early life. *Social Science & Medicine - Population Health.* 2016;2:269–76.
- Shin ME, McCarthy WJ. The association between county political inclination and obesity: results from the 2012 presidential election in the United States. *Prev Med.* 2013;57(5):721–4.
- Remington PL, Booske BC. Measuring the health of communities—how and why? *J Public Health Manag Pract.* 2011;17:397–400.
- Hahn EJ, Rayens MK, Butler KM, Zhang M, Durbin E, Steinke D. Smoke-free laws and adult smoking prevalence. *Prev Med.* 2008;47(2):206–9.
- Hurt RD, Weston SA, Ebbert JO, McNallan SM, Croghan IT, Schroeder DR, Roger VL. Myocardial infarction and sudden cardiac death in Olmsted County, Minnesota, before and after smoke-free workplace LawsMI and cardiac death with smoke-free workplace law. *Arch Intern Med.* 2012; 172(21):1635–41.
- Weber MD, Bagwell DAS, Fielding JE, Glantz SA. Long term compliance with California's smoke-free workplace law among bars and restaurants in Los Angeles County. *Tob Control.* 2003;12(3):269–73.
- Ellis GA, Hobart RL, Reed DF. Overcoming a powerful tobacco lobby in enacting local smoking ordinances: the contra Costa County experience. *J Public Health Policy.* 1996;17(1):28–46.
- Kroon LA, Corelli RL, Roth AP, Hudmon KS. Public perceptions of the ban on tobacco sales in San Francisco pharmacies. *Tob Control.* 2012.
- Gerber AS, Gruber J, Hungerman DM. Does church attendance cause people to vote? Using blue laws' repeal to estimate the effect of religiosity on voter turnout. *Br J Polit Sci.* 2016;46(3):481–500.
- Centers for Disease Control: Burden of Tobacco Use in the U.S. In. Atlanta, Georgia; 2018.
- Cerdá M, Johnson-Lawrence VD, Galea S. Lifetime income patterns and alcohol consumption: investigating the association between long- and short-term income trajectories and drinking. *Soc Sci Med.* 2011;73(8):1178–85.
- Reich M, Montaloux C, Allegretto S, Jacobs K, Bernhardt A, Thomson S: The effects of a \$15 minimum wage by 2019 in San Jose and Santa Clara County. Institute for Research on labor and employment, June 2016.
- Bryden A, Roberts B, Petticrew M, McKee M. A systematic review of the influence of community level social factors on alcohol use. *Health Place.* 2013;21:70–85.
- Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med.* 2004;58(10):1929–52.
- Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing “neighborhood effects”: social processes and new directions in research. *Annu Rev Sociol.* 2002;28:443–78.
- Ivory VC, Blakely T, Richardson K, Thomson G, Carter K. Do changes in neighborhood and household levels of smoking and deprivation result in changes in individual smoking behavior? A large-scale longitudinal study of New Zealand adults. *Am J Epidemiol.* 2015;182(5):431–40.
- Cerdá M, Diez-Roux AV, Tchetgen ET, Gordon-Larsen P, Kiefe C. The relationship between neighborhood poverty and alcohol use: estimation by marginal structural models. *Epidemiology.* 2010;21(4):482.
- Center for Human Resource Research: A Guide to the 1979–2000 National Longitudinal Survey of Youth Data. In. Columbus, Ohio; 2001.
- Galea S, Ahern J, Tracy M, Rudenstine S, Vlahov D. Education inequality and use of cigarettes, alcohol, and marijuana. *Drug Alcohol Depend.* 2007; 90(Supplement 1):S4–S15.
- Miles R. Neighborhood disorder and smoking: findings of a European urban survey. *Soc Sci Med.* 2006;63(9):2464–75.
- Bode E. Annual educational attainment estimates for U.S. counties 1990–2005. *Leit Spat Resour Sci.* 2011;4(2):117–27.
- U.S. Census Bureau American Community Survey [<https://www.census.gov/programs-surveys/acs/>] Accessed 18 Mar 2019.
- Bureau of Labor Statistics. Local Area Unemployment Statistics [<http://www.bls.gov/lau/>] Accessed 18 Mar 2019.
- Bureau of Economic Analysis. Regional Economic Accounts [<http://www.bea.gov/regional/>] Accessed 18 Mar 2019.
- National Institute on Alcohol Abuse and Alcoholism: Drinking Levels Defined. In. Washington, D.C.; 2019.
- Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ.* 2009;338:b2393.
- Von Hippel PT. Regression with missing Ys: an improved strategy for analyzing multiply imputed data. *Sociol Methodol.* 2007;37(1):83–117.
- White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med.* 2011;30(4): 377–99.

49. Gardiner JC, Luo Z, Roman LA. Fixed effects, random effects and GEE: what are the differences? *Stat Med.* 2009;28(2):221–39.
50. Wooldridge JM: Chapter 8. Heteroskedasticity. In: *Introductory econometrics: A modern approach.* edn. Boston, Massachusetts: Cengage Learning; 2016.
51. Arceneaux K, Nickerson DW. Modeling certainty with clustered data: a comparison of methods. *Polit Anal.* 2009;17(2):177–90.
52. Hamad R, Rehkopf DH. Poverty, pregnancy, and birth outcomes: a study of the earned income tax credit. *Paediatr Perinat Epidemiol.* 2015;29(5):444–52.
53. Strully KW, Rehkopf DH, Xuan Z. Effects of prenatal poverty on infant health state earned income tax credits and birth weight. *Am Sociol Rev.* 2010;75(4):534–62.
54. Hamad R, Modrek S, White JS. Paid family leave effects on breastfeeding: a quasi-experimental study of U.S. policies. *Am J Public Health.* 2019;109(1):164–6.
55. Leigh JP, Leigh WA, Du J. Minimum wages and public health: a literature review. *Prev Med.* 2019;118:122–34.
56. Buu A, Wang W, Wang J, Puttler LI, Fitzgerald HE, Zucker RA. Changes in women's alcoholic, antisocial, and depressive symptomatology over 12 years: a multilevel network of individual, familial, and neighborhood influences. *Dev Psychopathol.* 2011;23(01):325–37.
57. Diez-Roux AV, Nieto FJ, Muntaner C, Tyroler HA, Comstock GW, Shahar E, Cooper LS, Watson RL, Szklo M. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol.* 1997;146(1):48–63.
58. Pollack CE, Cubbin C, Ahn D, Winkleby M. Neighbourhood deprivation and alcohol consumption: does the availability of alcohol play a role? *Int J Epidemiol.* 2005;34(4):772–80.
59. Fauth RC, Leventhal T, Brooks-Gunn J. Short-term effects of moving from public housing in poor to middle-class neighborhoods on low-income, minority adults' outcomes. *Soc Sci Med.* 2004;59(11):2271–84.
60. Ludwig J, Duncan GJ, Genetian LA, Katz LF, Kessler RC, Kling JR, Sanbonmatsu L. Long-term neighborhood effects on low-income families: evidence from moving to opportunity. *Am Econ Rev Pap Proc.* 2013;103(3):226–31.
61. University of Wisconsin Population Health Institute: County Health Rankings. In. Madison, Wisconsin; 2016.
62. Oakes JM, Andrade KE, Biyoow IM, Cowan LT. Twenty years of neighborhood effect research: an assessment. *Current Epidemiology Reports.* 2015;2(1):80–7.
63. Glass TA, Goodman SN, Hernán MA, Samet JM. Causal inference in public health. *Annu Rev Public Health.* 2013;34:61–75.
64. Vandembroucke JP, Broadbent A, Pearce N: Causality and causal inference in epidemiology: the need for a pluralistic approach. *Int J Epidemiol* 2016, Published online ahead of print, 22 Jan 2016.
65. Oakes JM, Fuchs EL, Tate AD, Galos DL, Biyoow IM. How should we improve neighborhood health? Evaluating evidence from a social determinant perspective. *Current Epidemiology Reports.* 2016;3(1):106–12.
66. White JS, Hamad R, Li X, Basu S, Ohlsson H, Sundquist J, Sundquist K. Long-term effects of neighbourhood deprivation on diabetes risk: quasi-experimental evidence from a refugee dispersal policy in Sweden. *The Lancet Diabetes & Endocrinology.* 2016;4(6):517–24.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

