Mapping and Measuring Social Disparities in Premature Mortality: The Impact of Census Tract Poverty within and across Boston Neighborhoods, 1999–2001

Jarvis T. Chen, David H. Rehkopf, Pamela D. Waterman,
S. V. Subramanian, Brent A. Coull, Bruce Cohen,
Mary Ostrem, and Nancy Krieger

ABSTRACT *The identification and documentation of health disparities are important functions of public health surveillance. These disparities, typically falling along lines defined by gender, race/ethnicity, and social class, are often made visible in urban settings as geographic disparities in health between neighborhoods. Recognizing that premature mortality is a powerful indicator of disparities in both health status and access to health care that can readily be monitored using routinely available public health surveillance data, we undertook a systematic analysis of spatial variation in premature mortality in Boston (1999–2001) across neighborhoods and sub-neighborhoods in relation to census tract (CT) poverty. Using a multilevel model based framework, we estimated that the incidence of premature mortality was 1.39 times higher (95% credible interval 1.09– 1.78) among persons living in the most economically deprived CTs (≥20% below poverty) compared to those in the least impoverished tracts (<5% below poverty). We present maps of model-based standardized mortality ratios that show substantial within-neighborhood variation in premature mortality and a sizeable decrease in spatial variation after adjustment for CT poverty. Additionally, we present maps of model-based direct standardized rates that can more readily be compared to externally published rates and targets, as well as maps of the population attributable fraction that show that in some of Boston's poorest neighborhoods, the proportion of excess deaths associated with CT poverty reaches 25–30%. We recommend that these methods be incorporated into routine analyses of public health surveillance data to highlight continuing social disparities in premature mortality.*

KEYWORDS *Age standardization, Area-based socioeconomic measures, Geocoding, Mapping, Multilevel models, Poverty, Premature mortality.*

INTRODUCTION

An important function of public health surveillance and monitoring is the identification of systematic disparities in who experiences the burden of disease and

Chen, Rehkopf, Waterman, Subramanian, and Krieger are with the Department of Society, Human Development, and Health, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA; Coull is with the Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA; Cohen is with the Massachusetts Department of Public Health, 250 Washington Street, Boston, MA, USA; Ostrem is with the Boston Public Health Commission, 1010 Massachusetts Avenue, Boston, MA, USA.

Correspondence: Jarvis T. Chen, ScD, Department of Society, Human Development, and Health, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. (E-mail: jarvis@hsph. harvard.edu)

death in the population. These health disparities, falling along lines of gender, race/ethnicity, and class, often reflect underlying social inequalities.[1] Efforts to identify, document, investigate, reduce, and ultimately eliminate these health disparities are thus of paramount interest to a just and democratic society. Additionally, public health practitioners must be concerned not only with *who* is at risk, but also *where* they live and how the communities in which they live do or do not support the health of their residents. To date, public health reports on health inequities have tended to present statistical summaries devoid of spatial analysis or depiction. To address this gap, we use a multilevel modeling approach in this paper to quantify and map social disparities in premature mortality in Boston, using techniques that yield age-standardized rates that can be readily depicted and compared to externally published rates and targets.

Given patterns of racial/ethnic and socioeconomic residential segregation, it should be no surprise that social disparities are often reflected in geographic variation in health. Over 175 years ago, the French physician Louis René Villermé published a pathbreaking paper examining patterns of mortality across different neighborhoods of Paris.[2,3] By linking newly available administrative census data with health records, he was able to demonstrate, for the first time, that the strongest predictor of neighborhood mortality was neighborhood poverty, which he measured using the number of households in each Parisian arrondissement exempt from rent tax due to low income. His study is not just remarkable for explicitly linking health to social inequities, but also for demonstrating how documentation of the geographic inequalities in health provided the initial impetus to investigate the social inequalities. To put it another way, geography does not just tell us where something happened, but also tells us who lived there and to what social circumstances they were exposed.

To quantify, depict, and routinely monitor the social geography of health, three conditions must be met. First, we require a routinely collected indicator of population health that is geographically locatable and sensitive to social disparities. Second, we require a valid measure of socioeconomic position that can be linked to both the public health data and relevant denominator data. And third, we require a statistical methodology that can handle estimation of small area rates and allow comparison of the variability in rates before and after adjustment for socioeconomic position.

Premature mortality lends itself to being monitored as a salient indicator of social disparities in health in the United States (U.S.), since all deaths must be reported to state health departments, and death certificates include the decedent's address and age. Additionally, a growing body of research documents that premature mortality is a powerful indicator of disparities in both health status and access to health care, with risk of premature death directly increasing with economic deprivation.[4–8] For example, in Massachusetts in 2001, the premature mortality rate (deaths before age 75, age-standardized to the year 2000 standard million) was two times higher in the worst-off cities and towns (with rates in excess of 450 per 100,000) compared with the best-off areas (rates ~220/100,000).[9]

Regarding socioeconomic measures, major obstacles have been the lack of routinely collected socioeconomic data in U.S. public health surveillance systems and the inability to link to population denominators stratified by the same measures. Though death certificates include information on decedent's education, these data may be misclassified.[10] Moreover, until recently the U.S. census public release files did not provide the necessary denominator data by age, race/ethnicity, and education.[11] To overcome these obstacles, in our Public Health Disparities

Geocoding Project we have validated the use of geocoding and linkage of public health records and population denominators to area-based socioeconomic measures at the census tract level.[12–14] In particular, the census tract (CT) level measure "percent of persons below the U.S. poverty line" consistently detects expected socioeconomic gradients in health across a wide range of health outcomes, among both the total population and diverse racial/ethnic-gender groups; yields maximal geocoding and linkage; and is readily interpretable by a wide range of audiences. Using this methodology, we found that overall premature mortality rates in Massachusetts in 1989–1991 were 2.2 times higher (95% confidence interval (CI) 2.0–2.3) for persons living in CTs with ≥20% of the population below poverty (the federal definition of a poverty area[15]) compared to <5% of the population below poverty.[14]

With respect to statistical considerations, multilevel modeling using generalized linear mixed models has become an increasingly popular method for dealing with small area rate instability by "smoothing" over unstable rates.[16–20] This approach decomposes overall variation at multiple levels of spatial aggregation and models the contribution of "place" using random effects that are assumed to be drawn from a common underlying distribution. This effectively permits one to "borrow strength" across neighboring areas in the multilevel hierarchy to yield more stable place level estimates. These models, which provide a framework for validly combining data from different spatial scales (e.g., neighborhood and CT) and also permit adjustment for multiple individual and area-level covariates, have been used in the spatial analysis of relationships between social environmental exposures and a wide variety of health outcomes.[21–27] However, to date these types of models have rarely been used by U.S. state and local health departments to analyze social disparities in routinely collected public health surveillance data.[28]

Instead, in order to communicate associations between social variables and health outcomes, the cartographic approach used by some health departments and health researchers is to present two maps side by side, one showing area-variation in health outcomes, the other area variation in socioeconomic conditions.[28–30] The net effect is to ask the map viewer to "eyeball" the two maps to gauge whether worse health occurs in poorer areas. While maps are an appealing way to display spatially varying data, this visual comparison of two adjacent maps, however, is not the same as a rigorous statistical analysis of the contribution of area socioeconomic conditions to area health profiles. State and local health departments rightly place a priority on being able to communicate health disparities clearly and succinctly to community groups and policy makers. Thus, it would be desirable to be able to translate the results of a multilevel model-based analysis of social disparities in health into easily interpretable epidemiologic quantities that can be effectively represented on a map.

Accordingly, to advance efforts to routinely monitor and map health disparities, in this study we examine the social geography of premature mortality at both the neighborhood and CT levels. We include both area-levels because (1) neighborhoods are a key level for the work of health departments, community health clinics, and community groups, and (2) our prior research has suggested the importance of addressing CT variation of economic conditions and health status within neighborhoods. The specific aims of our study were (1) to characterize CT and neighborhood variation in premature mortality in the city of Boston (1999–2001); (2) to estimate the magnitude of the socioeconomic gradient in premature mortality based on CT poverty and the contribution of CT poverty to census tract and neighborhood variation; and (3) to map the population impact of CT poverty on premature mortality.

MATERIALS AND METHODS

Study Population

We obtained records for all reported deaths in the city of Boston in 1999–2001 ($n = 13,361$) from the Massachusetts Department of Public Health (MDPH) and geocoded all records using the address on the death certificate. According to MDPH, reporting of deaths for this time period was effectively 100% complete.[31] Employing a commercial geocoding firm with known high accuracy,[32] we were able to geocode 97% of the records with certainty to the CT level. We also assigned deaths to 1 of 16 Boston neighborhoods defined by the Boston Public Health Commission based on the identified CT of residence. Premature mortality was defined as death before age 75 ($n = 5,945$). Age of the decedent was categorized as follows: <1, 1–4, 5–14, 15–24, 25–34, 35–44, 45–54, 55–64, and 65–74. We extracted the corresponding population counts by CT and age category from Summary File 1 of the U.S. 2000 Census.[33] Person-time denominators were calculated by multiplying these population counts by 3 years, such that estimated premature mortality rates from our analysis are interpretable as average annual rates. Following the methodology of the Public Health Disparities Geocoding Project,[12–14] we calculated the CT percent of persons below poverty measure based on Table P87 of the U.S. 2000 Census Summary File 3.[34] The CT poverty measure could be determined for all of the 156 CTs in Boston in 2000. The population distribution by CT poverty and neighborhood and corresponding age-standardized premature mortality rates (standardized using the direct method to the year 2000 standard million[35]) are presented in Table 1.

Modeling

The numerator, denominator, and CT poverty data were organized in a multilevel structure of age cells nested within CTs nested within neighborhoods. We modeled premature mortality using Poisson mixed effects models as follows. Let y_{ijk} represent the count of cases observed in age stratum i, census tract j, and neighborhood k, and let n_{ijk} represent the corresponding person-time denominator. Then a simple age-adjusted model can be fit as:

$$y_{ijk} \sim \mathrm{Poisson}\left(\lambda_{ijk}\right) \tag{1}$$
$$\log\left(\lambda_{ijk}\right) = \log\left(n_{ijk}\right) + \beta_0 + \beta_i^{age} + u_{jk} + \nu_k$$

where u_{jk} and ν_k are CT and neighborhood specific random effects, respectively, and $\exp(u_{jk} + \nu_k)$ and $\exp(\nu_k)$ are interpretable as CT and neighborhood specific standardized premature mortality ratios. For these analyses, we treated age as a categorical covariate so that β_i^{age} is a coefficient associated with age category i, with $\beta_1^{age} = 0$. To model the additional contribution of CT poverty, we fit:

$$\log\left(\lambda_{ijk}\right) = \log\left(n_{ijk}\right) + \beta_0 + \beta_i^{age} + \beta_p^{CTPOV} + u_{jk} + \nu_k \tag{2}$$

where p corresponds to categories of CT poverty (0–4.9%, 5–9.9%, 10–19.9%, 20–100%). We treated the random u_{jk} and ν_k terms as independent normal random effects with variances σ_u^2 and σ_v^2, respectively.

We took a Bayesian approach to model fitting, using a Markov Chain Monte Carlo (MCMC) algorithm that yields simulated samples from the posterior distribution of the parameters given the data. We present the posterior means and 95% credible intervals corresponding to these posterior samples when reporting

TABLE 1. Premature deaths (<75 years old), person-years, average annual age-standardized premature mortality rate and 95% confidence interval by CT poverty and Boston neighborhood, 1999–2001
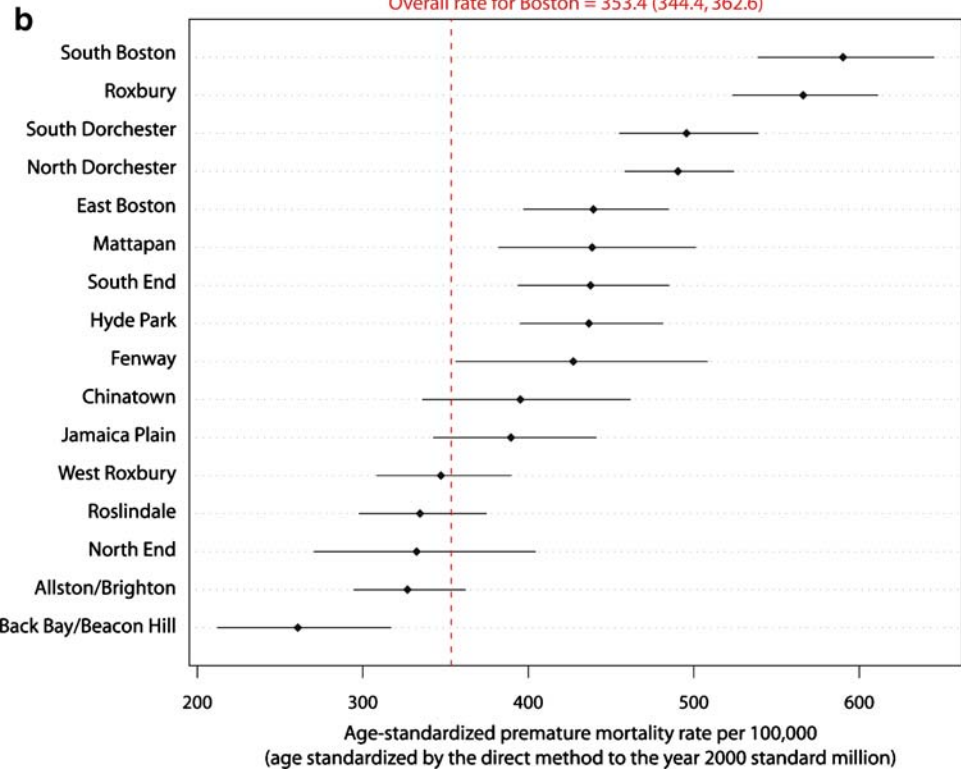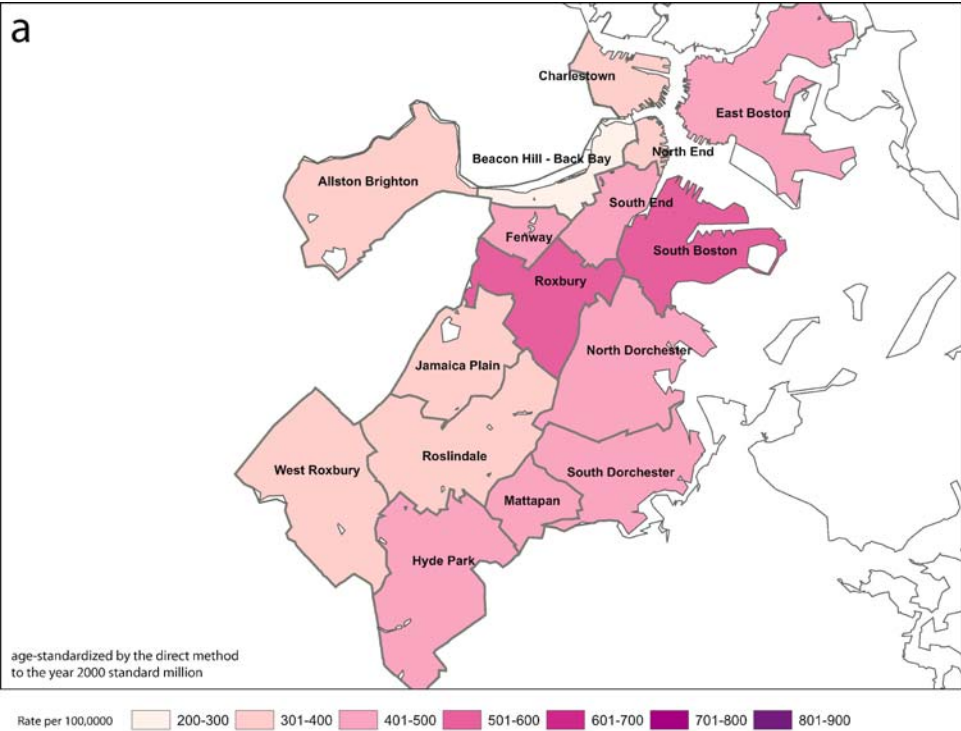
	Deaths	Person-years	Age-standardized[1]	Rate (95% CI)
CT poverty				
0–4.9%	360	91,155	323.3	(271.03,339.41)
5–9.9%	996	271,830	382.8	(335.83,384.77)
10–19.9%	2239	598,674	429.1	(385.39,421.57)
20–100%	2350	713,334	474.4	(426.40,465.79)
Neighborhood				
Back Bay/Beacon Hill	195	104,577	260.8	(212.44,316.87)
Allston/Brighton	404	198,486	327.0	(294.75,361.88)
North End	103	33,441	332.5	(270.73,404.26)
Roslindale	305	97,059	334.6	(297.86,374.63)
West Roxbury	298	69,897	347.2	(308.44,389.51)
Jamaica Plain	265	84,660	389.6	(342.77,440.94)
Chinatown	162	43,494	395.2	(336.18,461.65)
Fenway	139	87,387	427.2	(356.28,508.14)
Hyde Park	411	96,537	436.7	(395.28,481.19)
South End	396	95,022	437.7	(393.92,484.91)
Mattapan	221	57,369	438.6	(382.19,500.98)
East Boston	410	107,946	439.4	(397.36,484.58)
North Dorchester	908	241,773	490.5	(458.52,524.02)
South Dorchester	562	129,420	495.6	(455.18,538.73)
Roxbury	671	143,979	566.1	(523.72,611.04)
South Boston	495	83,946	590.2	(538.92,645.00)

[1] Age standardized using the direct method to the year 2000 standard million.

results of model fitting. Non-informative but proper priors were specified for all regression coefficients and variance components. We specified Normal(0,10000) priors for the β regression coefficients, and Uniform(0,50) priors for each of the random effect standard deviations.[36] Models were fit by calling the WinBUGS package[37] from within R, using functions from the R2WinBUGS library[38] and specifying 70,000 iterations with 10,000 burn-in and thinning the chain by 30. We evaluated model convergence by examining traceplots, the autocorrelation function for the model parameters, and Gelman–Rubin statistics. All these methods supported convergence, with all Gelman–Rubin statistics within 0.001 of 1.0. Model fit was evaluated using the deviance information criterion (DIC),[39] which weighs overall model fit against a penalty term for decreased model parsimony. We further evaluated the sensitivity of our models to choice of prior distributions for the variance components by re-running models using the half-Cauchy distribution recommended by Gelman[36] and found negligible differences.

Mapping

We obtained Census tract TIGER shape files for the year 2000 from the U.S. Bureau of the Census website.[40] We constructed neighborhood level shape files as aggregations of CTs based on neighborhood definitions from the Boston Public
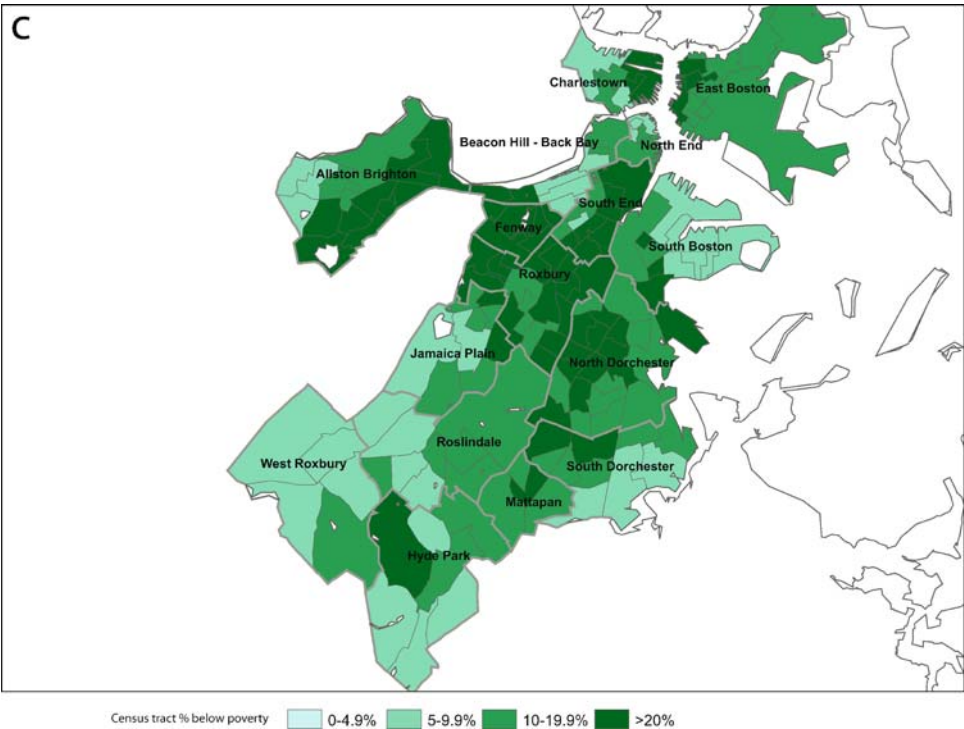
a

age-standardized by the direct method
to the year 2000 standard million

Rate per 100,0000   200-300   301-400   401-500   501-600   601-700   701-800   801-900

b

Overall rate for Boston = 353.4 (344.4, 362.6)

Age-standardized premature mortality rate per 100,000
(age standardized by the direct method to the year 2000 standard million)

FIGURE 1. (**a**) Average annual age-standardized premature mortality rates (deaths before age 75) by Boston neighborhood, 1999–2001 (*not smoothed*) (**b**) Dotchart of average annual age-standardized premature mortality rates and 95% confidence intervals by Boston neighborhood, 1999–2001 (*not smoothed*) (**c**) Percent persons living below poverty in Boston census tracts, 2000 (U.S. Census).

Health Commission. We constructed all maps using ArcGIS[41] (map colors based on http://www.ColorBrewer.org).[42]

In order to generate maps that clearly communicate patterns of premature mortality, area-level poverty, and the relationship between the two, we calculated several model-based quantities. First, we mapped CT- and neighborhood-specific standardized premature mortality ratios (SMR) *before* and *after* adjustment for CT poverty, where $SMR_{neigh} = \exp(v_k)$ and $SMR_{ct} = \exp(u_{jk}+v_k)$. Secondly, we calculated a corresponding model-based directly age-standardized rate, based on the expected count of cases that would be observed in an area if the total population of that area were assigned to age categories based proportionately on the U.S. 2000 standard million. Finally, we also estimated a model-based population attributable fraction, defined as the proportion of cases that would not have been observed if the population of the area experienced the premature mortality burden of the least impoverished tracts (i.e., if residents in each CT had the same average age-specific premature mortality rates as those experienced in the least impoverished CTs). We calculated this quantity by predicting the expected number of cases that would have arisen if CT poverty = 0–4.9%, based on the model, and by expressing the difference from the observed number of cases as a percent. Note that because they are model-based, both the model-based direct standardized rate and the population attributable fraction incorporate the CT- and neighborhood-specific random effects.

RESULTS

In 2000, the population of Boston included 589,141 persons, of whom 19.5% were living below the poverty line. For the time period of our study, 1999–2001, the overall average annual premature mortality rate in Boston, age standardized to the year 2000 standard million using the direct method, was 353.4 per 100,000 (95% CI 344.4–362.6). This overall premature mortality rate represents a composite of the premature mortality rates experienced in the city's different neighborhoods and the CTs within them, which in turn are shaped in part by their poverty levels, as shown by our analysis.

Overall, we found significant variation in premature mortality rates across CTs and neighborhoods in Boston, with notable variation at the CT level *within* the neighborhood units recognized by the Boston Public Health Commission. Importantly, neighborhoods with relatively low premature mortality rates can nevertheless contain CTs with high rates and vice versa. Inclusion of CT poverty in the model considerably reduced observed CT and neighborhood level variation, although spatial variation in premature mortality persisted.

Figure 1a presents a map of the neighborhood level direct age-standardized premature mortality rates tabulated in Table 1, with a corresponding dotchart (Fig. 1b) showing associated 95% credible limits. This neighborhood level map shows higher premature mortality rates in South Boston, Roxbury, and North and South Dorchester. The analogous CT level map is not presented, due to the sparseness of population in some of the age cells in many CTs and the resulting rate instability. However, in Fig. 1c, we do present the CT poverty rates in CTs in Boston, with neighborhood boundaries also shown. This is the type of socioeconomic map that might typically be paired with the health outcome map to invite visual inspection of whether areas with high poverty also have high rates of premature mortality. While it is tempting to draw conclusions about the overlap of neighborhoods with high premature mortality and census tracts with high poverty, one can do so only casually because of the different levels of aggregation in the two maps (neighborhood vs. CT).[43]

When CT poverty was added in Model 2, we observed increasing rate ratios for premature mortality across categories of CT poverty (Table 2). Rates were 39% higher in the most compared to least impoverished CTs (rate ratio of 1.39, 95% CI 1.09–1.78, comparing CTs meeting the federal criterion for "poverty areas,"[15] i.e., $\geq$ 20% poverty, to those with < 5% poverty). This estimate was similar to that presented in Table 1, based on aggregation of deaths and population at risk across census tracts and neighborhoods into strata defined by CT poverty. This suggests that, in the case of Boston, estimation of the overall impact of CT poverty on premature mortality was not biased by aggregation of counts across CTs with similar CT poverty. Comparison of the estimates of σ_u^2 and σ_v^2 from Model 2 to Model 1 shows that inclusion of CT poverty in the model reduced CT variation by 8.3% and neighborhood variation by 13.2%.

To visualize the geographic variation in premature mortality before and after adjusting for CT poverty, we generated maps of the standardized premature mortality ratio, derived from the posterior estimates of the model-based neighborhood and CT random effects. These ratios are interpretable as deviations from the mean rate for Boston overall. Figure 2a presents these quantities at the neighborhood level before adjustment for CT poverty. Though not directly comparable to the direct standardized rates in Fig. 1, we can see a similar pattern of geographic

TABLE 2.   Model estimates of fixed and random effects from Models 1 (BEFORE adjustment for CT poverty) and 2 (AFTER adjustment for CT poverty) for premature mortality in Boston, MA, 1999–2001

Fixed effects	Model 1a		Model 1b		
	IRR* (95% CI)		IRR (95% CI)		
CT poverty					
0–4.9%			1.00 (reference)		
5–9.9%			1.11	(0.87,1.42)	
10–19.9%			1.27	(1.00,1.63)	
20–100%			1.39	(1.09,1.78)	
Random effects	Estimate (standard error)		Estimate (standard error)		Percent change σ^2
CT level σ_u^2	0.048	(0.010)	0.044	(0.009)	8.3%
Neighborhood level σ_v^2	0.068	(0.035)	0.059	(0.029)	13.2%

* Incidence Rate Ratio (IRR).

variation in premature mortality rates, with the lowest rates seen in the neighborhoods of Back Bay/Beacon Hill and Allston/Brighton and the highest rates in South Boston and Roxbury.

Showing, however, that there can be substantial variation of premature mortality rates within the conventional neighborhood demarcations, in Fig. 2b we present the analogous map for CT level standardized premature mortality ratios. This map also depicts the boundaries of Boston's 16 neighborhoods. Results indicated that neighborhoods with overall relatively low premature mortality rates nevertheless can contain CTs with high premature mortality rates. Moreover, even among neighborhoods with an overall high premature mortality rate, some CTs fared far worse than others. For example, although overall the South End had a lower premature mortality rate than the adjacent Roxbury and South Boston neighborhoods, which together had the highest neighborhood premature mortality rates in Boston, the CT with the highest premature mortality rate nevertheless was in the South End. Additionally, within both Roxbury and South Boston, while the premature mortality rates in every CT were above the mean premature mortality rate for Boston, the ratios for excess premature mortality among their CT ranged from relatively low (1.05–1.25) to substantial (1.75–2.00).

Figure 3a and b shows the impact of additionally adjusting for CT poverty on the neighborhood and CT level standardized premature mortality ratios. As predicted by comparison of σ_u^2 and σ_v^2 from Models 1 and 2, this adjustment did on average reduce the overall spatial variation, although it did not alter the ordering of elevated rates in South Boston, Roxbury, and North and South Dorchester. Importantly, however, not all neighborhood and CT specific standardized premature mortality ratios were uniformly adjusted towards the null. For example, while standardized premature mortality ratios in Roxbury and North Dorchester were reduced after adjustment (8 and 6%, respectively), the standardized premature mortality ratio for South Boston actually increased by 4%. That is, South Boston had an unusually high premature mortality rate in spite of its relatively lower CT poverty composition. Conversely, adjustment for CT poverty reduced the already low standardized premature mortality ratio for Allston/Brighton even further, by 3%.
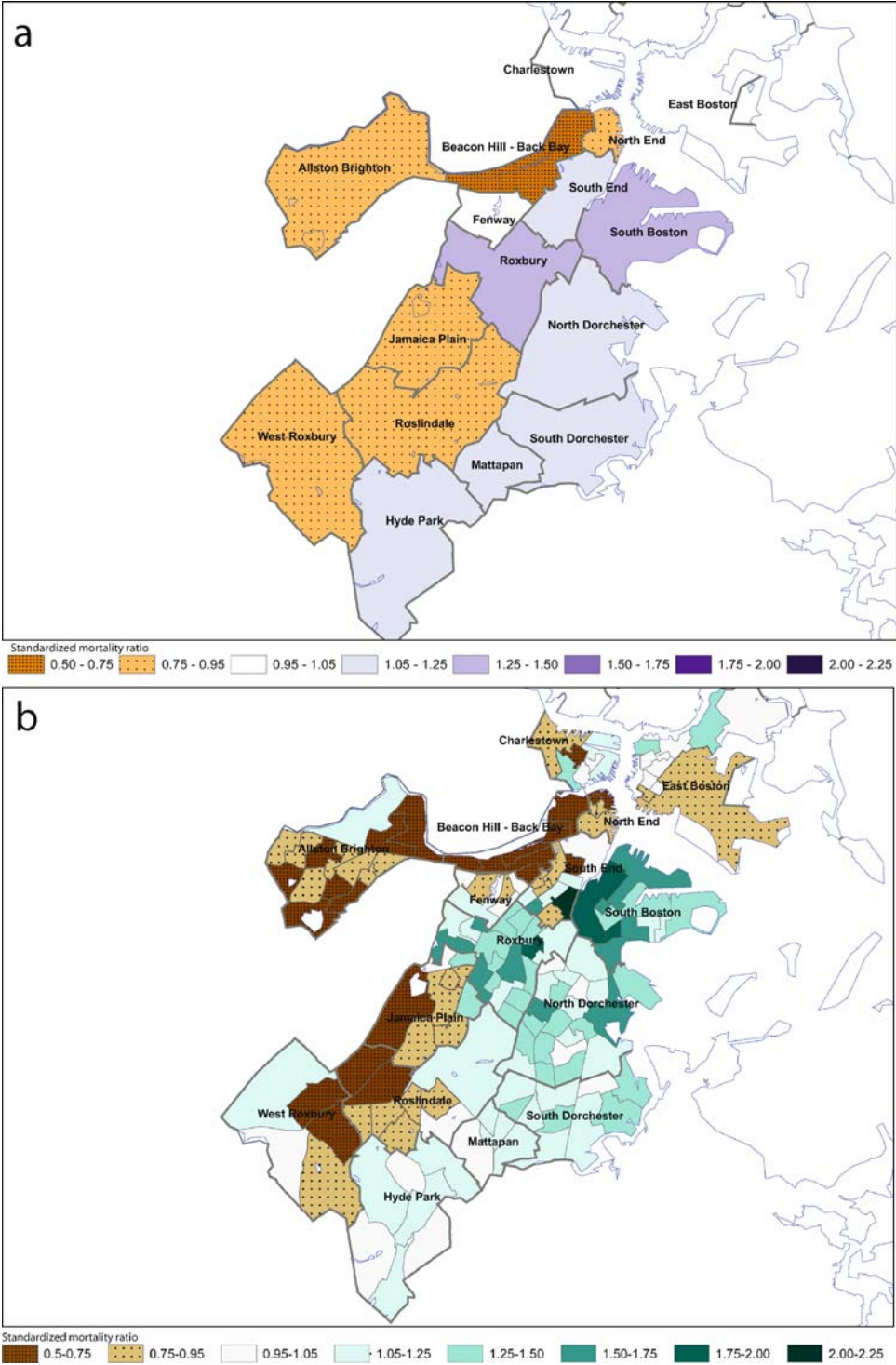
FIGURE 2. (**a**) Neighborhood standardized premature mortality ratios BEFORE adjustment for CT poverty, Boston, MA, 1999–2001. (**b**) Census tract standardized premature mortality ratios BEFORE adjustment for CT poverty, Boston, MA, 1999–2001.
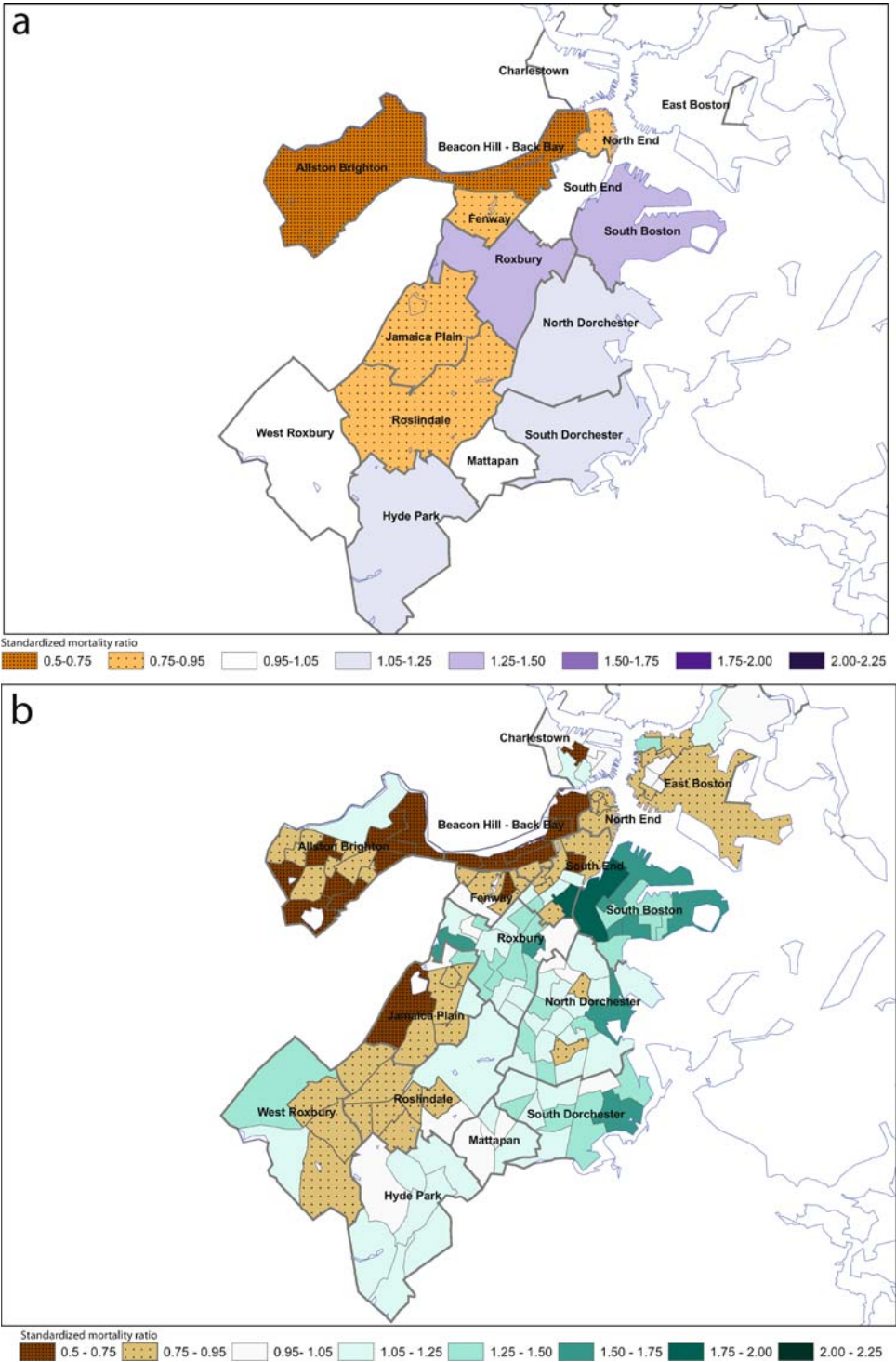
FIGURE 3. (**a**) Neighborhood standardized premature mortality ratios AFTER adjustment for CT poverty, Boston, MA, 1999–2001. (**b**) Census tract standardized premature mortality ratios AFTER adjustment for CT poverty, Boston, MA, 1999–2001.
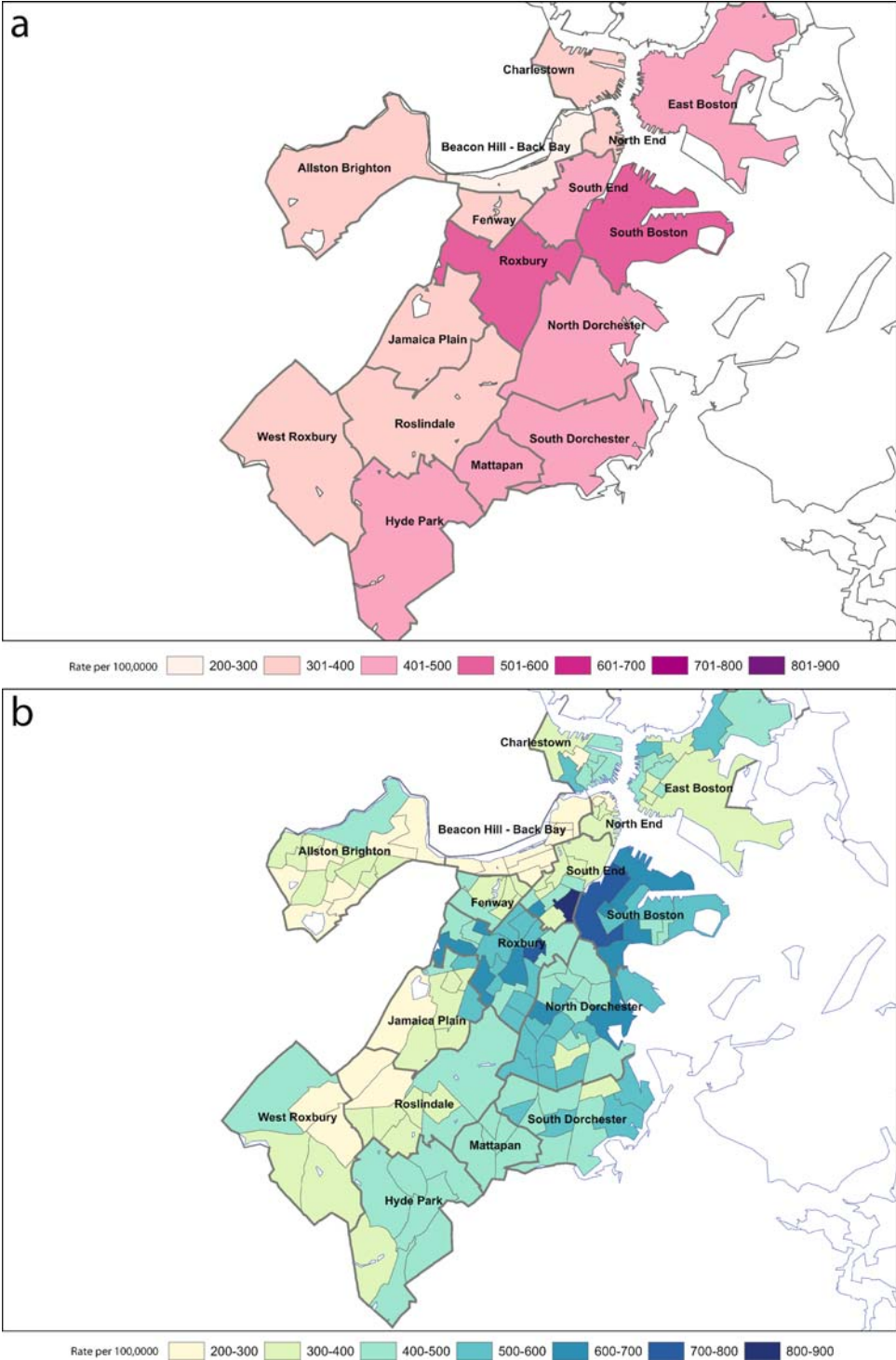
Rate per 100,0000 □ 200-300 □ 301-400 □ 401-500 ■ 501-600 ■ 601-700 ■ 701-800 ■ 801-900

Rate per 100,0000 □ 200-300 □ 300-400 □ 400-500 ■ 500-600 ■ 600-700 ■ 700-800 ■ 800-900

FIGURE 4. (**a**) Neighborhood model-based direct standardized rates BEFORE adjustment for CT poverty, Boston, MA, 1999–2001. (**b**) Census tract model-based direct standardized rates BEFORE adjustment for CT poverty, Boston, MA, 1999–2001.
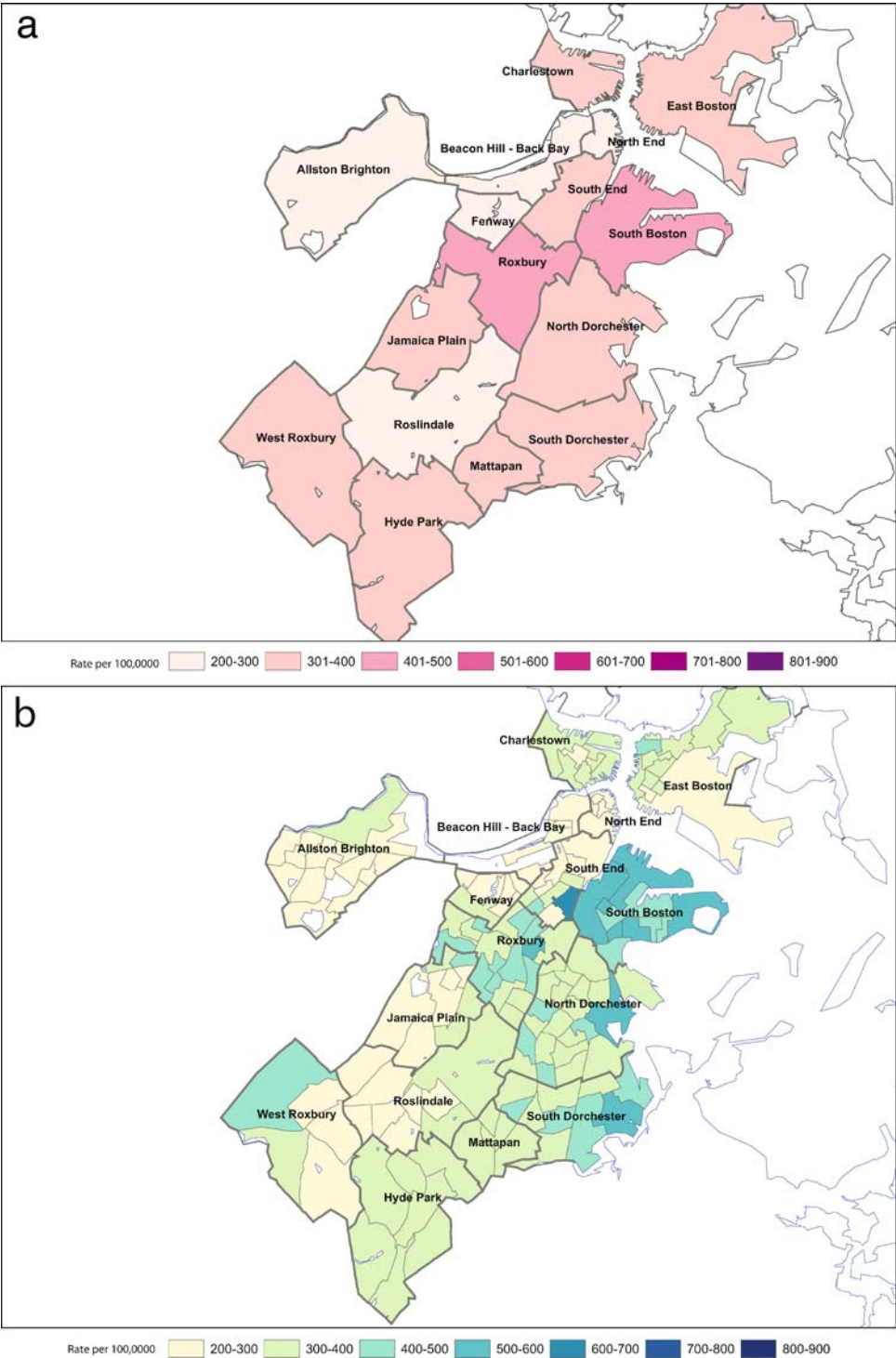
FIGURE 5. (**a**) Neighborhood model-based direct standardized rates AFTER adjustment for CT poverty, Boston, MA, 1999–2001. (**b**) Census tract model-based direct standardized rates AFTER adjustment for CT poverty, Boston, MA, 1999–2001.

Next, to generate neighborhood and CT level rates that are directly interpretable as being age standardized to the year 2000 standard million and, hence, comparable to other data using the same age standard, we calculated the model-based direct standardized rates as described above. The maps for the neighborhood and CT levels before adjustment for CT poverty are shown, respectively, in Fig. 4a and b. They depict a similar spatial patterning to Fig. 2a and b. However, rather than requiring interpretation relative to an unspecified mean rate for Boston, the quantities being mapped are the actual age standardized premature mortality rates, that is, the rates that would have been observed if the area in question had the same age distribution as the year 2000 standard million. Interpretation of the geographic patterning from low to high is further enhanced by the sequential scheme of shading, compared with the two-ended diverging scheme required for the standardized premature mortality ratio maps (Fig. 2a and b).

We next predicted the model-based direct standardized rates after adjustment for CT poverty (based on Model 2a) based on "standardizing" the CT poverty level of all CTs to the least impoverished category (<5%). The resulting maps are presented in Fig. 5a and b and show a substantial predicted reduction in premature mortality across all census tracts, although areas of elevated premature mortality still persist in South Boston, Roxbury, and North and South Dorchester. This persistence corresponds to the areas in Fig. 3a and b where the standardized premature mortality ratio increased after adjustment for CT poverty.

Finally, we mapped the model-based population attributable fraction in relation to CT poverty for Boston neighborhoods (Fig. 6a) and CTs (Fig. 6b). This fraction represents the proportion of premature deaths that would not have occurred if residents in every CT enjoyed the same age-specific premature mortality rates as residents of the least impoverished CTs. Thus, they provide a representation of the population impact of CT poverty on premature mortality, assuming that the association between CT poverty and premature mortality arises from a causal mechanism. This proportion exceeded 20% for 50% of Boston's 16 neighborhoods and 68% of Boston's 156 CTs. Within 14 of the 17 CTs (82%) in Roxbury and 12 of the 22 CTs (55%) in North Dorchester (two of Boston's poorest neighborhoods), the proportion of premature deaths that would not have occurred if residents had the same mortality experience as residents in the least impoverished CTs in Boston reached 25–30%.

DISCUSSION

Our results demonstrate how a multilevel model-based framework can be used to explore the social geography of premature mortality in an urban setting based on routinely collected public health surveillance data. The multilevel framework allows us to partition observed variation in premature mortality at the neighborhood, CT, and within-CT levels and to model this variation as a function of fixed and random effects. As a result, this framework allows for: (a) statistical smoothing in the estimation of small area rates; (b) estimation of variance at each of the specified levels; and (c) adjustment for multiple covariates. We have also presented a set of measures and corresponding 95% credible intervals based on Monte Carlo simulation, including model-based direct standardized rates, standardized premature mortality ratios, and population attributable fraction that can be easily calculated based on the model and mapped at the CT and neighborhood levels.
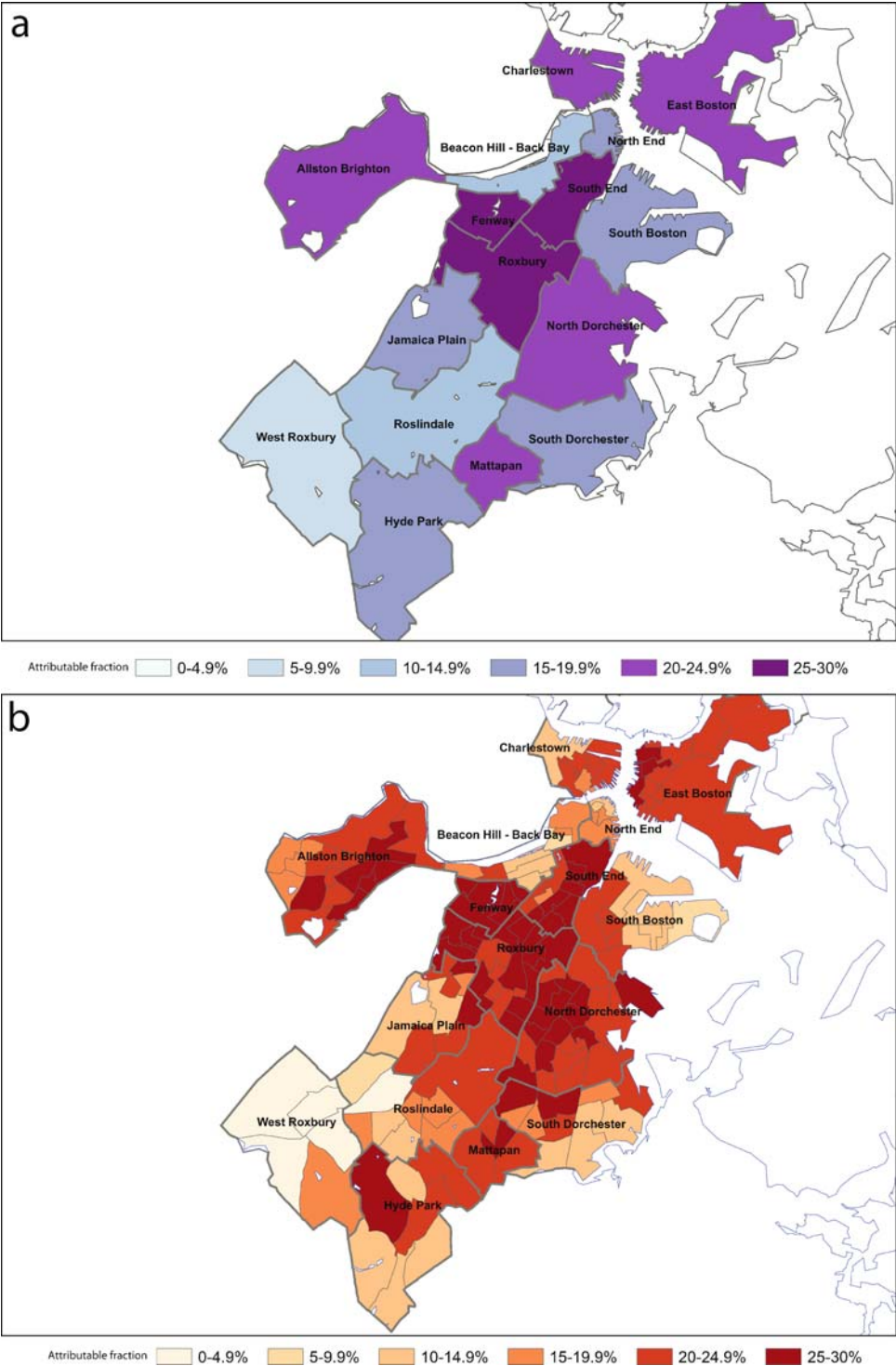
FIGURE 6. (**a**) Neighborhood model-based population attributable fraction due to CT poverty, Boston, MA, 1999–2001. (**b**) Census tract model-based population attributable fraction due to CT poverty, Boston, MA, 1999–2001.

Together, our results provide evidence of the role of CT poverty in patterning premature mortality within and across neighborhoods. In the most afflicted areas, whether neighborhoods or CTs, for every 100 premature deaths, 25 to 30 of the deaths would not have occurred if residents of these areas had the same age-specific premature mortality rates as residents of the city's least impoverished neighborhoods or CTs. To our knowledge, the population attributable fraction of premature mortality in relation to CT poverty has not previously been reported in the U.S. at either the neighborhood or CT level.

Before providing interpretation of additional aspects of our findings, however, we note a set of caveats relevant to interpretation of the causal effect of poverty on premature mortality. These include ecologic bias, the modifiable areal unit problem, etiologic period, and omitted variable bias.

Ecologic bias can occur when both the dependent and independent variables are group-level data, and confounding is introduced through the grouping process.[44,45] This can lead to fallacious inferences, particularly when the confounded group-level effect is used as a proxy for the corresponding individual level effect. We are careful to note, therefore, that our estimates of the CT poverty gradient are to be interpreted as a group-level association that likely reflects a complex combination of individual-level and area-level processes. However, we further note that the CT poverty effect is in the expected direction and is unlikely to reflect the sort of substantial cross-level confounding that can lead to a change in the direction of the association between poverty and premature mortality. Furthermore, in previous analyses for which we had individual as well as CT-level socioeconomic data, we have found that the direction of the socioeconomic gradient was the same, and also of similar magnitude, for both the individual-level and CT-level socioeconomic measures.[11,21,23]

The modifiable areal unit problem pertains to the sensitivity of spatial patterns to the choice of area level units in the analysis.[43] The multilevel approach we have used addresses one aspect of this problem, as it permits modeling of spatial variation in premature mortality at each of the three levels (age cell, CT, and neighborhood) at which we have data. Another aspect concerns whether or not CTs and neighborhoods are meaningful entities (rather than "arbitrary" spatial units) relevant to characterizing socioeconomic conditions and to shaping population health. We note that CT boundaries are drawn to be socially meaningful by local census committees[46] and that CTs serve as administrative areas relevant to resource allocation and for such programs as "Urban Empowerment Zones," "Medically Underserved Areas," and "Qualified Census Tracts" for the purpose of the low-income housing tax credit.[15,47,48] Moreover, in Boston, the Boston Public Health Commission uses CTs to define Boston neighborhoods for health programs, thereby having real-life implications for the health and quality of life of their residents.

Regarding etiologic period, in the case of premature mortality, CT characteristics at the time of death are likely to be temporally relevant to outcomes with a short etiologic period. These would include deaths among infants and children, as well as deaths among adults due to "preventable and immediate causes" or for which survival is drastically curtailed by inadequate access to health care.[6] Examples of the latter include deaths due to motor vehicle and other accidents; asthma, pneumonia and influenza; suicide, homicide and legal intervention; complications of pregnancy, diabetes, tuberculosis, and HIV/AIDS; as well as deaths due to temporally proximate triggers of heart attacks, e.g., pollution.[24] In 2000, these causes of death accounted for nearly 40% of potential years of life lost before age 75 in the U.S.[49]

Finally, regarding omitted variable bias, we note that while there are certainly covariates (e.g., smoking) that may also account for the observed disparities in premature mortality by CT poverty, these are likely to be in the causal pathways linking socioeconomic position to premature mortality, such that it may not be appropriate to control for them when the population impact of CT poverty is of interest.[44] We also emphasize that the primary aim of our analyses was to model the poverty–mortality relationship from a monitoring perspective, rather than engage in an exhaustive etiologic investigation of all potential determinants of premature mortality.

Assuming, then, that our findings present a plausible depiction of the patterning of premature mortality in Boston in relation to CT poverty, we offer some preliminary interpretation of additional aspects of our findings. First, our analysis confirms previously published observations of a socioeconomic gradient in premature mortality in the U.S., as reported in both the handful of other analyses using CT poverty,[7,14,23] as well as those employing economic indicators based on larger geographic areas (e.g., neighborhoods, metropolitan areas, and counties).[4–6] Only two of these prior studies used a multilevel analysis[7,23]; the rest used the more conventional approach of comparing rates based on aggregating the death and population data into strata defined by area-based socioeconomic position. Although this latter approach, by ignoring spatial variation, could potentially yield biased estimates, we note in our study that this did not occur, as shown by the similarity of the socioeconomic gradients estimated in Tables 1 and 2.

Interestingly, our estimates for Boston of a rate ratio of ~1.4, comparing premature mortality in census tracts with 20–100% poverty to those with <5% poverty, was equal to the analogous rate ratio we computed for Boston for 1988–1992, i.e., 1.4 (95% CI 1.3–1.5),[50] and less than the analogous rate ratio of ~2.2 we computed for this same time period for Massachusetts as a whole.[14] The smaller magnitude of the CT socioeconomic gradient for premature mortality in Boston compared to MA overall is likely due to their differing poverty levels. In 2000, the proportion of the population living in CTs with ≥20% below poverty (42%) was 2.5 times greater than for the state as a whole (16%). Given well-established patterns of wealthier suburban areas surrounding cities with more concentrated poverty,[51,52] it is possible that rates of premature mortality in Boston's least impoverished neighborhoods may nevertheless be higher than in their MA counterparts outside of Boston. If true, the net impact would be to reduce the Boston range of premature mortality associated with CT poverty level.

Second, we found that while adjusting for CT poverty overall tended to attenuate the estimated premature mortality ratio in most neighborhoods and CTs towards the null, as expected, in some cases it shifted the estimated premature mortality ratios *away* from the null. Thus, South Boston had an unusually high premature mortality rate despite its low poverty rate, whereas Allston/Brighton had an unusually low premature mortality rate despite its relatively high poverty rate.

Recognizing that interpretation of these anomalous results requires local knowledge of the social geography of Boston neighborhoods, we drew on the accumulated knowledge and experience of our local and state public health department partners to propose possible explanations. For the unusually high residual premature mortality rates in South Boston, hypotheses include: (1) a spike in heroin-related overdose deaths during the study period; (2) a high overall rate of mortality due to substance use, including both drugs and alcohol; (3) excess mortality due to lung cancer and cardiovascular disease, likely reflecting high rates of smoking;

and (4) the residential concentration of relatively well-paid police officers and firefighters, who may experience high premature mortality but do not live in impoverished CTs.[53] Conversely, Allston/Brighton is a neighborhood chiefly populated by students, who may not be gainfully employed but who do not experience high mortality. These interpretations, which point to the difference of student vs. family poverty and also the existence of relatively well-paid but risky jobs, are conjectural and require further study as possible explanations for the observed patterns. They also underscore the need for thoughtful interpretation of the relationship between CT poverty and risk of premature mortality, even as the overall findings clearly demonstrate that overall increasing CT poverty levels are associated with increased risk of premature mortality.

Lastly, an additional contribution of our study is to improve visual representation and interpretation of the social geography of premature mortality in relation to CT poverty. Previous literature on statistical mapping has noted that maps of standardized incidence ratios are interpretable only relative to the mean rate across the study population and, unlike directly standardized rates, are not readily comparable to externally published rates or policy relevant targets.[54] On the other hand, traditional direct standardization tends to perform poorly in small areas because it requires stable age-specific rates. At the small area level (e.g., CTs), this is made further problematic by sparseness of the data and "empty" age cells in the denominators. Moreover, traditional direct standardization does not incorporate any smoothing at the area-level.

As an alternative to the standardized premature mortality ratio map, we have proposed a model-based direct standardized rate that overcomes the problems with traditional direct standardization while permitting comparison with published standardized rates (e.g., the Healthy People 2010 objectives).[55] Our small area rate estimates incorporate model-based smoothing and are robust to sparse age cells. Moreover, as we show in Fig. 5a and b, the method can be extended to yield predicted rates given a "set" level for the covariate CT poverty. This is a useful tool, for example, if one wishes to predict the effect of an intervention by "setting" the covariate level equal to some target in the model-based calculation.[56] Nevertheless, we note that care should be taken in interpreting the estimated population attributable fraction as an "etiological" fraction, since this requires strong assumptions about the causal relationships between area socioeconomic conditions, premature mortality, and other unmeasured covariates. For this reason, we stress that the population attributable fraction should simply be regarded as an "excess" fraction.[57]

We also note that, in the current analysis, Model 1a has assumed a set of consistent age effects across all areas. Therefore, the spatial patterning of the standardized premature mortality ratios and model-based direct standardized rates at the CT level is the same, although the former are centered around the "null" ratio of 1.0, rather than expressed on the rate scale. This corresponds to the condition of "no age-area interactions" under which indirectly and directly standardized rates are identical.[54] However, in our model-based framework, this assumption could be relaxed to allow age x area interactions (a "random slopes" model for age), or even age x covariate interactions that would be reflected in the model-based standardized rate, thereby yielding estimates that would be more similar to the purely non-parametric directly standardized rates than the indirectly standardized rates. In future analyses, we plan to explore these alternative modeling strategies and also investigate whether the poverty–premature mortality relationship varies by neigh-

borhoods, race/ethnicity, gender, and specific causes of death and also whether it exhibits any non-linearities.

In conclusion, we believe that wider use of our proposed methodology can enable local and state health departments, as well as academic researchers, to use routinely available data to generate meaningful, policy-relevant analyses and depictions of socioeconomic inequities in premature mortality. As recognized in the U.S.,[8] Canada,[58] and the United Kingdom,[59–61] the premature mortality rate is a highly informative, easily calculated, and easily understood single measure that captures social disparities in community health. When considered in the geographical context, it also has enormous potential for identifying acute, small area clusters especially burdened by premature death and encouraging exploration of the specific correlates of area deprivation (including environmental conditions, housing, education, material resources, behavioral factors, access to care, etc.) that contribute to social disparities in health.

ACKNOWLEDGEMENT

REFERENCES

1. National Institutes of Health. *NIH Strategic Plan to Reduce and Ultimately Eliminate Health Disparities*. Available at: http://www.nih.gov/about/hd/strategicplan.pdf. Accessed January 4, 2006.
2. Villermé LR. De la mortalité dans les divers quartiers de la ville de Paris, et des causes qui la rendent très différente dans plusieurs d'entre eux, ainsi que dans les divers quartiers de beaucoup de grandes villes. 1830 Available at: http://gallica.bnf.fr/Catalogue/NoticesInd/FRBNF37303496.htm. Accessed November 11, 2005.
3. Krieger N. The making of public health data: paradigms, politics, and policy. *J Public Health Policy*. 1992;13:412–427.
4. Cooper RS, Kennelly JF, Durazo-Arvizu R, Oh HJ, Kaplan G, Lynch J. Relationship between premature mortality and socioeconomic factors in black and white populations of U.S. metropolitan areas. *Public Health Rep*. 2001;116:464–473.
5. Mansfield CJ, Wilson JL, Kobrinski EJ, Mitchell J. Premature mortality in the United States: the roles of geographic area, socioeconomic status, household type, and availability of medical care. *Am J Public Health*. 1999;89:893–898.
6. Ronzio CR, Pamuk E, Squires GD. The politics of preventable deaths: local spending, income inequality, and premature mortality in U.S. cities. *J Epidemiol Community Health*. 2004;58:175–179.
7. Franzini L, Spears W. Contributions of social context to inequalities in years of life lost to heart disease in Texas, USA. *Soc Sci Med*. 2003;57:1847–1861.
8. Fastrup J, Vinkenes M, O'Dell M. *Public Health: A Status Indicator for Targeting Federal Aid to States*. Washington, District of Columbia: U.S. General Accounting Office; 1996.
9. Bureau of Health Statistics, Research and Evaluation. *Massachusetts Deaths 2001: Premature Mortality*. Boston, Massachusetts: Massachusetts Department of Public Health, 2003. Available at: http://www.mass.gov/dph/bhsre/death/2001/premtrmortal2001.pdf. Accessed January 5, 2006.

10. Sorlie P, Johnson NJ. Validity of education information on the death certificate. *Epidemiology*. 1996;7:437–439.

11. Rehkopf DH, Haughton L, Chen JT, Waterman PD, Subramanian SV, Krieger N. Monitoring socioeconomic disparities in death; comparing individual-level education and area-based socioeconomic measures. *Am J Public Health*. (in press)

12. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of U.S. socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: the public health disparities geocoding project. *Am J Epidemiol*. 2002;156:471–482.

13. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Race/ethnicity, gender, and monitoring socioeconomic gradients in health: a comparison of area-based socioeconomic measures—the public health disparities geocoding project. *Am J Public Health*. 2003;93:1655–1671.

14. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Painting a truer picture of U.S. socioeconomic and racial/ethnic health inequalities: the public health disparities geocoding project. *Am J Public Health*. 2005;95:312–323.

15. U.S. Census Bureau. *Poverty Areas*. Available at: http://www.census.gov/population/socdemo/statbriefs/povarea.html. Accessed January 16, 2006.

16. Waller LA, Gotway CA. *Applied Spatial Statistics for Public Health Data*. Hoboken, New Jersey: Wiley; 2004.

17. Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. *J Am Stat Assoc*. 1993;88:9–25.

18. Clayton DG, Bernardinelli L. Bayesian methods for disease mapping. In: Elliott P, Cuzick J, English D, Stern R, eds. *Geographical and Environmental Epidemiology*. Oxford University Press; 1992:205–220.

19. Subramanian SV, Jones K, Duncan C. Multilevel methods for public health research. In: Kawachi I, Berkman L, eds. *Neighborhoods and Health*. New York: Oxford University Press; 2000:65–111.

20. Diez-Roux AV. Multilevel analysis in public health research. *Annu Rev Public Health*. 2000;21:171–192.

21. Subramanian SV, Chen JT, Rehkopf DR, Waterman PD, Krieger N. Comparing individual and area-based socioeconomic measures for surveillance of health disparities: a multilevel analysis of Massachusetts (U.S.) births, 1988–92. *Am J Epidemiol*. (in press)

22. Jannsen I, Boyce WF, Simpson K, Pickett W. Influence of individual- and area-level measures of socioeconomic status on obesity, unhealthy eating, and physical inactivity in Canadian adolescents. *Am J Clin Nutr*. 2006;83:139–145.

23. Subramanian SV, Chen JT, Rehkopf DH, Waterman PD, Krieger N. Racial disparities in context: a multilevel analysis of neighborhood variations in poverty and excess mortality among black populations in Massachusetts. *Am J Public Health*. 2005;95:260–265.

24. Tonne C, Schwartz J, Mittleman M, Melly S, Suh H, Goldberg R. Long term survival after acute myocardial infarctions is lower in more deprived neighborhoods. *Circulation*. 2005;111:3063–3070.

25. Stafford M, Martikainen P, Lahelma E, Marmot M. Neighborhoods and self rated health: a comparison of public sector employees in London and Helsinki. *J Epidemiol Community Health*. 2004;58:772–778.

26. Jarup L, Best N, Toledano MB, Wakefield J, Elliott P. Geographical epidemiology of prostate cancer in Great Britain. *Int J Cancer*. 2002;97:695–699.

27. Diez-Roux AV, Nieto FJ, Muntaner C, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol*. 1997;146:48–63.

28. Pickle LW, Mungiole M, Jones GK, White AA. *Atlas of United States Mortality*. Hyattsville, Maryland: National Center for Health Statistics; 1996.

29. Casper ML, Bernett E, Halverson JA, et al. *Women and Heart Disease: An Atlas of Racial and Ethnic Disparities in Mortality, 2nd edition*. Morgantown, West Virginia: Office of Social Environment and Health Research, West Virginia University; 2000.

30. Boston Public Health Commission; *The Disparities Project: Data Report*. Boston, Massachusetts: Boston Public Health Commission; 2005. Available at: http://www.bphc.org/director/pdfs/datareport/datareport.pdf. Accessed December 29, 2005

31. Bureau of Health Statistics, Research and Evaluation. *Massachusetts Deaths 2002*. Boston, Massachusetts: Massachusetts Department of Public Health; 2004. Available at: http://mass.gov/dph/bhsre/resep/resep.htm#death Accessed January 3, 2006.

32. Krieger N, Waterman P, Lemieux K, Zierler S, Hogan JW. On the wrong side of the tracts? Evaluating the accuracy of geocoding in public health research. *Am J Public Health*. 2001;91:1114–1116.

33. U.S. Census Bureau. *Census 2000 Summary File 1 Technical Documentation*. Available at: http://www.census.gov/prod/cen2000/doc/sf1.pdf. Accessed January 10, 2005.

34. U.S. Census Bureau. *Census 2000 Summary File 3 Technical Documentation*. Available at: http://www.census.gov/prod/cen2000/doc/sf3.pdf. Accessed January 10, 2005.

35. Anderson RN, Rosenberg HM. Age standardization of death rates: implementation of the year 2000 standard. *National Vital Statistics Report*. 1998;47:1–16.

36. Gelman A. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*. 2005;1:1–19.

37. Spiegelhalter DJ, Thomas A, Best NG, Gilks WR, Lunn D. BUGS: Bayesian inference using Gibbs sampling. MRC Biostatistics Unit, Cambridge, England, 2003. Available at: http://www.mrc-bsu.cam.ac.uk/bugs/; Accessed on January 8, 2006.

38. Sturtz S, Ligges U, Gelman A. R2WinBUGS: a package for running WinBUGS from R. *Journal of Statistical Software*. 2005;12:1–16.

39. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc*. 2002;64:583–639.

40. U.S. Census Bureau. *108th CD Census 2000 Tiger/Line®Files Technical Documentation*. Washington, District of Columbia: U.S. Census Bureau; 2003.

41. ESRI. ArcGIS. Available at: http://www.esri.com; Accessed on: January 8, 2006.

42. Brewer C. ColorBrewer. 2006 Available at: http://www.ColorBrewer.org. Accessed January 8, 2006.

43. Openshaw S. *The Modifiable Areal Unit Problem*. Norwich, United Kingdom: Geo; 1984.

44. Elliott P, Wakefield JC. Bias and confounding in spatial epidemiology. In: Elliott P, Wakefield J, Best N, Briggs D. *Spatial Epidemiology: Methods and Applications*. Oxford: Oxford University Press; 2000.

45. Greenland S, Morgenstern H. Ecological bias, confounding and effect modification. *Int J Epidemiol*. 1989;18:269–274.

46. U.S. Bureau of Census. *Census of Population and Housing, 1990: Summary Tape File 3 Technical Documentation*. Washington, District of Columbia: Bureau of the Census; 1991:A-5.

47. Dept of Housing and Urban Development. *Qualified Census Tracts and Difficult Development Areas*. Available at: http://www.huduser.org/datasets/qct/qct99home.html. Accessed January 18, 2006.

48. Health Resources and Services Administration. *Health Professional Shortage Areas*. Available at: http://bhpr.hrsa.gov/shortage/. Accessed January 18, 2006.

49. National Center for Health Statistics. *Health, United States 2004. With Chartbook on Trends in Health of Americans with Special Feature on Drugs*. Hyattsville, Maryland: NCHS; 2005.

50. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Geocoding and monitoring U.S. socioeconomic inequalities in health: an introduction to using area-based socioeconomic measures—The Public Health Disparities Geocoding Project monograph. Boston, Massachusetts: Harvard School of Public Health. Available at: http://www.hsph.harvard.edu/thegeocodingproject/. Accessed January 17, 2006.

51. Northridge ME, Sclar ED, Biswas P. Sorting out the connections between the built environment and health: a conceptual framework for navigating pathways and planning healthy cities. *J Urban Health*. 2003;80:556–658.

52. Galea S, Vlahov D. Urban health: evidence, challenges, and directions. *Annu Rev Public Health*. 2005;26:341–365.

53. Boston Public Health Commission Research Office. *The Health of Boston 2005*. Boston, Massachusetts: Boston Public Health Comission; 2005. Available at: http://www.bphc.org/news/default.asp. Accessed January 19, 2006.

54. Pickle LW, White AA. Effects of the choice of age-adjustment method on maps of death rates. *Stat Med*. 1995;14:615–627.

55. U.S. Department of Health and Human Services. *Healthy People 2010* (Conference edition, in two volumes). Washington, District of Columbia: U.S. Govt Printing Office; 2000. Available at: http://www.health.gov/healthpeople. Accessed January 19, 2006.

56. Greenland S. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. *Am J Epidemiol*. 2004;160:301–305.

57. Greenland S, Robins JM. Conceptual problems in the definition and interpretation of attributable fractions. *Am J Epidemiol*. 1988;128:1185–1197.

58. Eyles J, Birch S. A population needs-based approach to health care resource allocation and planning in Ontario: a link between policy goals and practice. *Can J Public Health*. 1993;84:112–117.

59. Carstairs V, Morris R. *Deprivation and Health in Scotland*. Aberdeen, Scotland: Aberdeen University Press; 1991.

60. Townsend P, Phillimore P, Beatte A. *Health and Deprivation: Inequality and the North*. London, UK: Croom Helm; 1988.

61. Shaw M, Dorling D, Gordon D, Davey Smith G. *The Widening Gap: Health Inequalities and Policy in Britain*. Bristol, UK: Policy; 1999.