



ELSEVIER

Social Science & Medicine 66 (2008) 691–703

SOCIAL
SCIENCE
&
MEDICINE

www.elsevier.com/locate/socscimed

Population vulnerabilities and capacities related to health: A test of a model

Jennifer Ahern^{a,b,c,*}, Sandro Galea^{b,c,d}, Alan Hubbard^e, Adam Karpati^f

^a*Division of Epidemiology, University of California, Berkeley School of Public Health, 101 Haviland Hall, Berkeley, CA 94720-7358, USA*

^b*Department of Epidemiology, University of Michigan School of Public Health, USA*

^c*Center for Urban Epidemiologic Studies, New York Academy of Medicine, USA*

^d*Department of Epidemiology, Columbia University Mailman School of Public Health, USA*

^e*Division of Biostatistics, University of California, Berkeley School of Public Health, USA*

^f*Department of Health and Mental Hygiene, New York City, USA*

Available online 19 November 2007

Abstract

Variability in the health of human populations is greater in economically vulnerable areas. We tested whether this variability reflects and can be explained by: (1) underlying vulnerabilities and capacities of populations and/or (2) differences in the distribution of individual socioeconomic status between populations. Health outcomes were rates of mortality from 12 causes (cardiovascular disease, malignant neoplasms, accidents, chronic lower respiratory disease, cerebrovascular disease, pneumonia and influenza, diseases of the nervous system, suicide, chronic liver disease and cirrhosis, diabetes, homicide, HIV/AIDS) for 59 New York City neighborhoods in 2000. Negative binomial regression models were fit with a measure of socioeconomic vulnerability, median income, predicting each mortality rate. Overdispersion of each model was used to assess whether variability in mortality rates increased with increasing neighborhood socioeconomic vulnerability. To assess the two hypotheses, we examined changes in the variability of mortality rates (as indicated by changes in overdispersion of the models) for outcomes with significant non-constant variability after accounting for (1) vulnerabilities and capacities (social control, quality of local schools, unemployment, low education), and (2) the distribution of individual socioeconomic status (low income, poverty, socioeconomic distribution, high income). Some variability in all mortality rates was explained by accounting for a range of potential vulnerabilities and capacities, supporting the first explanation. However, variability in some causes of mortality was also explained in part by accounting for the distribution of individual resources, supporting the second explanation. The results are consistent with a theory of underlying socioeconomic vulnerabilities of human populations. In areas with lower levels of income, other characteristics of those neighborhoods exacerbate or temper the economic vulnerability, leading to more or less healthy conditions. Understanding the vulnerabilities and capacities that characterize populations may help us better understand the production of population health, and may inform efforts aimed at improving population health.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Neighborhood; Population health; Vulnerability; Social environment; Inequality

*Corresponding author. Division of Epidemiology, University of California, Berkeley School of Public Health, 101 Haviland Hall, Berkeley, CA 94720-7358, USA. Tel.: +1 510 643 4350; fax: +1 510 643 5163.

E-mail address: jahern@berkeley.edu (J. Ahern).

Introduction

While much of modern epidemiology has focused on estimating the risk of disease for individuals, there have been calls in recent years to consider the importance of viewing human populations as units of interest in epidemiology (Levins & Lopez, 1999; Rockhill, 2005; Rose, 1992, 2001). There are compelling reasons for public health professionals to consider population health as a matter of interest. The majority of cases of most diseases arise in persons who would not be considered at “high-risk” for these diseases (Rose, 1992, 2001; Syme, 1996), thus interventions at the population level have the potential to make a stronger and more enduring public health impact as they influence all members of the population. Moreover, interventions that have not incorporated a broader understanding of the population in which they are being conducted (e.g., cultural context, social norms) have had limited success (Merzel & D’Afflitti, 2003). Therefore, increasing our understanding of the population conditions that enhance or worsen the health of populations may substantially improve our public health policies.

Interest in population health has sparked an increase in studies of population level exposures in relation to health outcomes (Pickett & Pearl, 2001; Sellstrom & Bremberg, 2006). Much of this research has focused on small geographically based populations that could be considered communities or neighborhoods. Research has examined population level exposures ranging from economic indicators such as median income, to social and structural characteristics such as community cohesion and elements of the built environment (McNeill, Kreuter, & Subramanian, 2006; Papas et al., 2007; Pickett & Pearl, 2001). As in most of epidemiologic research, research on neighborhood level exposures has primarily focused on the average associations between exposures and health outcomes. Certainly these associations are important as they suggest on average how health changes in response to changes in neighborhood exposures. Some of this research has also quantified between neighborhood variation in health outcomes, typically through the use of mixed models with random intercepts (Diez Roux, 2000). A small body of research suggests that variability may be an informative aspect of population health to examine on its own. Variability in health between neighborhoods may not be constant, and may change in magnitude (i.e., variability is low

between some neighborhoods while high between others) depending on aspects of neighborhoods such as socioeconomic characteristics (Galea, Ahern, & Karpati, 2005; Karpati, Galea, Awerbuch, & Levins, 2002).

Previous work has documented variability in health between neighborhoods that share the same level of a socioeconomic neighborhood exposure, and found that the variability was greater among communities with more economic vulnerability. Based on this observation and work of in the fields of ecology and infectious disease research (Koopman & Longini, 1994; Koopman & Lynch, 1999; Levins, 1975), we developed a theoretical model describing the production of health at the population level (Galea et al., 2005). The model proposes that population health reflects the inter-relationship between underlying socioeconomic vulnerabilities and capacities of populations and intermittent stressors and protective events that occur within populations. Variability in population health potentially is produced both by the random or uneven distribution of these intermittent stressors and by differences in underlying vulnerabilities and capacities that can be recruited to cope with stressors as they occur (Galea et al., 2005). Briefly, the rationale for this model is as follows. Populations that are characterized by socioeconomic vulnerabilities (e.g., low median income) have few material and psychosocial resources available to cope with intermittent stressors (e.g., closure of a large employer in a community). Conversely, populations that are characterized by socioeconomic capacities (e.g., high median income) may draw on substantial resources when faced with such a situation. Based on this model, we expect that populations characterized by underlying socioeconomic vulnerability will have greater variability in population health indicators, such as mortality rates, reflecting their greater sensitivity to stressors, and reflecting differences in the other vulnerabilities and capacities that characterize each population.

However, while the variability in population health that we observed among socioeconomically vulnerable areas is consistent with our proposed model, there is an alternate explanation that merits consideration. Indicators of socioeconomic vulnerability in a population, such as median income, are summaries of the individuals living in the area that necessarily ignore potential variation in the distributions of the underlying individuals. For example, one population characterized by a given level of

median income may have a substantial population of extremely poor residents, while another population with the same level of median income may have residents whose incomes are predominantly close to the median. The first population may have worse health because the extremely poor have worse health outcomes, while the second population may have better health as it contains fewer extremely poor residents. Following this alternate explanation, socioeconomically vulnerable populations may manifest more variability in health because: (a) they tend to have a wider range of variation in the underlying distribution of individual socioeconomic status than more wealthy populations, or (b) for a given level of variation in the underlying distribution of individual socioeconomic resources, neighborhoods with lower average wealth tend to include more of the extremely poor than wealthy populations.

In this analysis, we tested two hypotheses to explain the previously documented variability in health in socioeconomically vulnerable populations. Consistent with previous work, we considered urban neighborhoods as the populations of interest in this analysis. Based on our previously articulated model (Galea et al., 2005), we hypothesized that variability in population health between socioeconomically vulnerable neighborhoods would be explained by accounting for other vulnerabilities and capacities of those neighborhoods. Based on the alternate explanation, we hypothesized that variability in population health between vulnerable neighborhoods would be explained by accounting for the distribution of individual socioeconomic status in the neighborhoods of interest.

Methods

Units of analysis

The units of analysis for this study were neighborhoods in New York City, defined as the 59 New York City community districts (CDs). CDs are well-defined units, each with an administrative community board, which have political and social significance for their residents. Politically, the board of each CD serves as an advisory body with a formal role designated by the City Charter in matters including land use, determining local budget priorities and monitoring City service delivery. Meetings are held monthly and are open to the public. Each board establishes committees to focus

on specific issues of concern. As an example, CD1 which encompasses lower Manhattan has set as some of its priorities: revitalization of several commercial streets and park areas, establishment of a community cultural center, funding for community and arts organizations, streetscape improvements, and addition of a public library branch (Manhattan Community Board 1). The activities of these boards clearly have the potential to build and shape community capacities. Socially, CDs were initially defined by a resident consultative process organized by the Office of City Planning to reflect residents' own descriptions of neighborhoods in the 1970s and as a consequence represent recognizable neighborhood areas with which residents identify, such as the Upper East Side, or the South Bronx. As would be expected in a city as diverse as New York City, CDs are not demographically homogenous. However, they represent neighborhoods that are associated with particular resident behavior and health (Ahern & Galea, 2006; Galea et al., 2003; Hembree et al., 2005). Based on the 2000 Census, the 59 New York City CDs have a mean of 135,681 residents (median = 128,313; range = 34,420–242,952).

Data

Dependent variables

We studied age-adjusted mortality rates for 12 publicly available causes of death from 2000 (New York City Department of Health and Mental Hygiene, 2000) provided by the New York City Department of Health and Mental Hygiene, Bureau of Vital Statistics for each New York City neighborhood. Causes of mortality studied were: cardiovascular disease (ICD-10 I00-I09, I11, I13, I20-I51), malignant neoplasms (ICD-10 C00-C97), accidents (ICD10 V01-X59, Y85-Y86), chronic lower respiratory disease (ICD-10 J40-J47), cerebrovascular disease (ICD-10 I60-I69), pneumonia and influenza (ICD-10 J10-J18), diseases of the nervous system (ICD-10 G00-G98), suicide (ICD-10 X60-X84, Y87.0), chronic liver disease and cirrhosis (ICD-10 K70, K73-74), diabetes (ICD-10 E10-E14), homicide (ICD-10 X85-Y09, Y87.1) and HIV/AIDS (ICD-10 B20-B24). All data were in the public domain, did not contain personal identifiers, and were at the CD level.

Independent variables

Details on the independent variables used in this analysis and their sources are provided in Table 1.

We used the median household income of the neighborhood as a measure of socioeconomic vulnerability in this analysis. Median household income was selected as it is the only Census based measure that captures the domain of neighborhood income, and (since it is a median) does not also capture the spread of income among individuals, which was of separate interest in this analysis (Krieger et al., 2002, 2003).

To address our first hypothesis, we examined four additional variables which measure population vulnerabilities or capacities (Table 1). These particular four vulnerabilities and capacities were selected balancing the practical need to utilize variables available either from the United States (US) Census or from New York City agencies, with the theoretical interest in including variables that capture distinct aspects of the neighborhood environ-

ment that could be conceptualized as increasing the vulnerability or capacity of a community. (1) Social control was examined and measured as the percentage of sidewalks in each neighborhood that met an acceptable standard of cleanliness. Physical disorder such as litter is theorized to indicate low social control and to communicate low social control to residents, and has been associated with resident worries that could be caused by low social control such as fear of crime (Laraia et al., 2006; Perkins & Taylor, 1996). Social control was theorized to be a community capacity because in the face of a community stressor, communities with higher social control would potentially restrict harmful health behaviors such as substance use (Greiner, Li, Kawachi, Hunt, & Ahluwalia, 2004; Kawachi & Berkman, 2000). (2) Quality of local schools was examined and

Table 1

Definitions of variables measuring neighborhood socioeconomic vulnerability, vulnerabilities and capacities, and the distribution of individual income

Construct	Definition	Source
Socioeconomic vulnerability	Median household income	2000 US Census (US Census Bureau, 2000)
Vulnerabilities/capacities		
Social control	Percent of sidewalks that met an acceptable standard of cleanliness based on a seven point picture-based rating scale designed to reflect the public perception of acceptable cleanliness levels; annual average of twice monthly ratings of a city-wide sample	Survey conducted by the Department of Sanitation and data obtained from the Mayor's Office of Operations (New York City Mayor's Office of Operations Mayor's Management Report, 2002)
Quality of local schools	Percent of teachers in public elementary or middle schools with tenure of more than 2 years	Obtained from the biannual Mayor's Management Report (New York City Mayor's Office of Operations Mayor's Management Report, 2002)
Unemployment	Percent of persons who were unemployed in each neighborhood, among those who were of working age (≥ 16 years of age) and identified as part of the labor force	2000 US Census (US Census Bureau, 2000)
Low education	Percent of adults (≥ 25 years of age) with a less than high school education	2000 US Census (US Census Bureau, 2000)
Distribution of individual income		
Low income	Percent of households earning less than \$15,000 annually	2000 US Census (US Census Bureau, 2000)
Poverty	Percent of persons with income below the poverty limit, accounting for their household size and composition	2000 US Census (US Census Bureau, 2000)
Socioeconomic distribution	The Gini coefficient, a measure of the equality of the income distribution within an area (0 indicates complete equality, 1 indicates complete inequality) (Wagstaff et al., 1991)	2000 US Census (US Census Bureau, 2000)
High income	Percent of households earning \$150,000 or more annually	2000 US Census (US Census Bureau, 2000)

measured as the percent of teachers in public schools with tenure of more than 2 years. Research suggests that both recruitment of qualified teachers and retention of good teachers contribute to problems with teaching quality in urban schools in New York State (Jacob, 2007). Moreover, regardless of the quality of teachers lost, turnover is costly because educators are less experienced and more resources are used on training (Jacob, 2007). School quality can be considered a community capacity because quality of education affects children's social environment as well as their coping and problem solving skills, all of which may be important in the face of a stressful community event. (3) Unemployment was examined and measured as the percent of persons who were unemployed in each neighborhood. The level of unemployment may constitute a neighborhood vulnerability because in the face of a community stressor, the unemployed have worse mental health and fewer economic resources, and family members of the unemployed may be bearing more burden, creating a situation where there are fewer residents with resources to cope themselves or to assist others (Hamilton, Broman, Hoffman, & Renner, 1990; Hamilton, Hoffman, Broman, & Rauma, 1993). (4) Low education was examined and was measured as the percent of adults with a less than a high school education. A less educated population may be considered vulnerable because they may have worse jobs and fewer problem solving skills and may thus have limited group experience and training relevant to dealing with stressors in the neighborhood (Rouse & Barrow, 2006).

To address our second hypothesis, we considered four variables that measure the underlying distribution of individual socioeconomic status within each neighborhood (Table 1). We selected these particular variables because they capture the low and high ends of the distribution of individual income as well as the overall size of the distribution of individual income. All of these measures capture aspects of the constructs of neighborhood income and poverty and have been considered in previous analyses (Krieger et al., 2002). For the low end of the income distribution, we included both percent low income as well as percent below poverty, as poverty status also incorporates family size and composition and may thus capture a slightly different construct. (1) Low income was examined and measured as the percent of households earning less than \$15,000 annually, which is earning less than half of the national median income in the US. (2) Poverty was

included and measured as the percent of persons with income below the poverty limit, accounting for their household size and composition. (3) Socio-economic distribution was measured with the Gini coefficient, a measure of the equality of the income distribution within an area (0 indicates complete equality, 1 indicates complete inequality) (Wagstaff, Paci, & van Doorslaer, 1991). (4) High income was examined and measured as the percent of households earning \$150,000 or more annually, which is earning more than 400% of the national median income in the US.

Analysis

We described the neighborhood disease rates, vulnerabilities and capacities, and measures of the distribution of individual socioeconomic status using means and standard deviations (SD). We used Poisson and negative binomial regression to model the mortality rates and to quantify the heteroskedasticity (i.e., non-constant variance) of each rate. A Poisson probability model is based on the assumption that the mean and variance of the rate, within a given stratum of the covariate(s) in the model, are equal and quantified by the parameter λ (Hubbard & Jewell, 2007). This model inherently permits a level of heteroskedasticity—as the mean rate increases, the variance is allowed to increase such that it equals the mean. We treated Poisson as our “null” heteroskedasticity model. If the rates of each outcome fit the Poisson model well, this indicates that within each stratum of the covariates (e.g., among neighborhoods with the same level of median income, same proportion of residents with less than high school education, etc.) the underlying rate of disease is the same for all neighborhoods. If this were the case, there would be no heteroskedasticity in the model to explain with other covariates. A negative binomial probability model, in contrast, allows greater heteroskedasticity than Poisson analysis. The mean rate is still quantified by λ , however, the variance is $\lambda + (\lambda^2 \times 1/r)$ (Hubbard & Jewell, 2007). A large value of overdispersion (defined as $1/r$, where r is known as the dispersion) has several implications for our analysis: (1) the variance increases more dramatically with the mean than it does in a Poisson process, (2) for a given stratum defined by the covariate(s), the model assumes that there is a distribution of underlying rates of the outcome, and (3) there is heteroskedasticity in the rates of disease that is unexplained and

still worth exploring, i.e., we would want to understand why there are different underlying rates of the outcome within strata defined by the covariate(s).

The measure of economic vulnerability, median income, was the exposure in separate Poisson regression and negative binomial regression models predicting each mortality rate. The two models were compared using a likelihood ratio test, and a statistically significant test indicated that there was significant overdispersion for a particular outcome (i.e., $1/r > 0$, the negative binomial model was a better fit). The overdispersion ($1/r$) of the mortality rates were examined and the outcomes with statistically significant overdispersion were considered in further analysis to test the hypotheses of interest.

To address the first hypothesis, negative binomial models were fit as above, including median income and the four other potential vulnerabilities and capacities. To address the second hypothesis, negative binomial models were constructed with median income and the four measures of the distribution of individual socioeconomic status. The vulnerabilities and capacities, and the measures of socioeconomic distribution, were all divided into quarters and modeled with indicator variables to allow non-linear associations with the health outcomes. For each cause of mortality, the overdispersion parameters from the two adjusted models were compared with the overdispersion from the model with only the socioeconomic vulnerability measure, median income. If the model adjusting for vulnerabilities and capacities substantially reduced the overdispersion parameter, this was considered support for the first hypothesis. If the model adjusting for the distribution of individual socioeconomic status substantially reduced the overdispersion parameter, this was considered support for the second hypothesis. To illustrate these observations we created plots of actual and predicted rates of homicide mortality by the measure of socioeconomic vulnerability, median income.

Results

Means and SD of mortality rates for each cause of mortality and of the exposures are presented in Table 2. The highest rate of mortality was due to cardiovascular disease (310.48/100,000 person-years), and the lowest was due to suicide (5.15/100,000 person-years).

Table 2

Neighborhood mortality rates, vulnerabilities and capacities, and measures of the distribution of individual socioeconomic status: New York City, 2000

	Mean	SD	Range
Mortality rate^a			
Accidents	12.41	3.33	5.59–19.84
Cardiovascular disease	310.48	59.82	173.63–588.76
Cerebrovascular disease	26.59	9.50	10.00–46.24
Chronic liver disease and cirrhosis	8.02	6.10	0–33.67
Chronic lower respiratory disease	21.96	8.37	8.39–40.81
Diabetes	28.45	19.54	4.51–103.21
Disease of the nervous system	8.70	4.15	1.52–21.36
HIV/AIDS	29.41	28.16	0.67–125.24
Homicide	8.39	6.42	0–24.98
Malignant neoplasm	173.23	30.57	121.39–267.29
Pneumonia and influenza	30.10	12.11	13.41–62.11
Suicide	5.15	2.35	1.16–12.50
Socioeconomic vulnerability measure			
Median income	38713.81	15739.82	16000–79475
Vulnerabilities and capacities			
Percent acceptably clean sidewalks	87.86	7.79	70.10–99.00
Percent of teachers in schools > 2 years	63.32	6.51	49.00–78.00
Percent unemployed	10.88	5.54	3.71–23.81
Percent less than high school education	29.38	13.15	4.67–56.8
Distribution of individual income			
Percent < \$15,000 income	25.05	11.28	9.07–48.21
Percent below poverty	22.38	11.73	4.90–45.67
Gini coefficient	0.45	0.03	0.37–0.51
Percent ≥ \$150,000 income	5.20	6.21	0.71–25.88

^aRate per 100,000 person-years.

The overdispersion parameters ($1/r$, where mean rate = λ , variance is $\lambda + [\lambda^2 \times (1/r)]$) from the negative binomial models with the measure of socioeconomic vulnerability, median income, predicting each cause of mortality are presented in Table 3, column (1); likelihood ratio tests comparing each model with the equivalent Poisson regression model, which tests the significance of the overdispersion parameters, are also presented. There was statistically significant overdispersion for all mortality rates except for those caused by accidents and suicide, however, there was a wide range in the magnitude of overdispersion by the different outcomes. The two least

Table 3

Overdispersion parameters from negative binomial regression models with (1) median income predicting each mortality rate, (2) median income and vulnerabilities and capacities predicting each mortality rate, and (3) median income and the distribution of individual socioeconomic status predicting each mortality rate

Cause of mortality	(1) Median income		(2) Median income with vulnerabilities and capacities ^c		(3) Median income with the distribution of individual SES ^d	
	Overdispersion (1/r) ^a	p-value ^b	Overdispersion (1/r) ^a	p-value ^b	Overdispersion (1/r) ^a	p-value ^b
Accidents ^e	0.006	0.57	–	–	–	–
Malignant neoplasm	0.018	<0.01	0.008	<0.01	0.009	<0.01
Cardiovascular disease	0.026	<0.01	0.016	<0.01	0.022	<0.01
Suicide ^e	0.048	0.12	–	–	–	–
Cerebrovascular disease	0.062	<0.01	0.006	0.29	0.028	<0.01
Chronic lower respiratory disease	0.077	<0.01	0.053	<0.01	0.055	<0.01
Disease of the nervous system	0.13	<0.01	0.023	0.24	0.064	<0.01
Pneumonia and influenza	0.14	<0.01	0.087	<0.01	0.096	<0.01
Diabetes	0.17	<0.01	0.074	<0.01	0.10	<0.01
Chronic liver disease and cirrhosis	0.19	<0.01	0.081	<0.01	0.053	0.01
Homicide	0.23	<0.01	0.023	0.34	0.20	<0.01
HIV/AIDS	0.63	<0.01	0.13	<0.01	0.30	<0.01

^aIn negative binomial regression, mean = λ and variance = $\lambda + \lambda^2 \times (1/r)$.

^bp-value from a likelihood ratio test comparing a Poisson model with a negative binomial model, equivalent to a test of whether there is significant overdispersion ($1/r > 0$).

^cModel with median income, and quarters of neighborhood social control, school quality, unemployment, and low education.

^dModel with median income, and quarters of neighborhood low income, poverty, Gini coefficient, and high income.

^eMortality rates were not considered in further analysis (columns 2 and 3) because overdispersion was not significant in the initial model (column 1).

overdispersed outcomes (among those where overdispersion was statistically significant) were cancer mortality ($1/r = 0.018$) and heart disease mortality ($1/r = 0.026$), and the two most overdispersed outcomes were homicide mortality ($1/r = 0.23$), and HIV/AIDS mortality ($1/r = 0.63$). The two outcomes that had no significant overdispersion were not considered further in the analysis.

Summaries of the models addressing the study hypotheses are presented in Table 3. Column (2) presents the overdispersion parameters ($1/r$) from the negative binomial models including median income and the four potential other vulnerabilities and capacities of the neighborhoods as predictors of each outcome. Column (3) presents the overdispersion parameters ($1/r$) from the negative binomial models with median income and the four measures of the distribution of individual socioeconomic status as predictors of each outcome. After adjustment for neighborhood vulnerabilities and capacities, there was substantial reduction in the overdispersion of all causes of mortality studied, ranging from the reduction in overdispersion of flu

mortality of 35% (overdispersion parameter reduced from 0.14 to 0.09) to the reduction in overdispersion of homicide mortality of 90% (overdispersion parameter reduced from 0.23 to 0.02) (Table 3, column (2)). For three causes of mortality (cerebrovascular disease, diseases of the nervous system, homicide), there was no statistically significant overdispersion remaining after adjustment for the vulnerabilities and capacities of neighborhoods. After adjustment for the distribution of individual socioeconomic status, there was also reduction in the overdispersion for most causes of mortality studied, although in most cases this was a smaller reduction than the reduction achieved by adjusting for vulnerabilities and capacities. The smallest reduction in overdispersion was a 15% reduction for homicide mortality (overdispersion parameter reduced from 0.23 to 0.20) and the largest was a 72% reduction in overdispersion of liver disease mortality (overdispersion parameter reduced from 0.19 to 0.05) (Table 3, column (3)). There were no causes of mortality for which adjustment for the distribution

of individual socioeconomic status reduced overdispersion so that it was no longer statistically significant.

As an illustration, the details of the models that produced the results in Table 3 are presented for liver disease mortality and homicide mortality in Table 4. These two outcomes were selected because adjusting for vulnerabilities and capacities made the largest reduction in homicide overdispersion and adjusting for the distribution of individual socio-

economic status made the largest reduction in liver disease mortality overdispersion.

Examining liver disease mortality, adjustment for other vulnerabilities and capacities of neighborhoods reduced the overdispersion parameter from 0.19 to 0.08 (Table 4). High levels of social control were associated with lower rates of liver disease mortality ($\beta = -1.06$, $p < 0.01$). However, adjustment for the distribution of individual socioeconomic status also reduced the overdispersion

Table 4

Detailed results of negative binomial regression models predicting liver disease mortality and homicide mortality: models with (1) median income alone, (2) median income and other neighborhood vulnerabilities and capacities, and (3) median income and the distribution of individual socioeconomic status

Outcome parameter	Liver mortality			Homicide mortality		
	Estimate	SE ^b	p-value	Estimate	SE ^b	p-value
Model (1)						
Intercept	-8.53	0.19	<0.01	-7.92	0.22	<0.01
Median income	-0.26	0.05	<0.01	-0.43	0.06	<0.01
Overdispersion (1/r)	0.19	0.06	<0.01	0.23	0.06	<0.01
Model (2)						
Intercept ^a	-8.26	0.81	<0.01	-9.15	0.76	<0.01
Median income	-0.10	0.14	0.46	-0.23	0.12	0.06
Social control Q2	-0.17	0.18	0.34	-0.20	0.14	0.16
Social control Q3	-0.21	0.21	0.31	0.18	0.17	0.30
Social control Q4	-1.06	0.32	<0.01	0.22	0.35	0.53
School quality Q2	-0.06	0.18	0.73	0.00	0.14	1.00
School quality Q3	-0.16	0.24	0.52	-0.48	0.21	0.02
School quality Q4	-0.26	0.26	0.33	-0.58	0.24	0.01
Unemployment Q2	-0.16	0.26	0.53	0.73	0.24	<0.01
Unemployment Q3	-0.26	0.29	0.37	1.26	0.26	<0.01
Unemployment Q4	0.19	0.39	0.62	1.52	0.33	<0.01
Low education Q2	-0.53	0.37	0.15	-0.13	0.36	0.72
Low education Q3	-0.60	0.45	0.18	-0.42	0.42	0.31
Low education Q4	-0.44	0.47	0.35	-0.60	0.44	0.18
Overdispersion (1/r)	0.08	0.03	<0.01	0.02	0.03	0.34
Model (3)						
Intercept	-9.60	0.79	<0.01	-8.64	1.09	<0.01
Median income ^a	-0.05	0.15	0.74	-0.25	0.21	0.24
Low income Q2	-0.59	0.27	0.03	-0.06	0.40	0.87
Low income Q3	-0.73	0.39	0.06	0.06	0.57	0.92
Low income Q4	-1.15	0.57	0.04	-0.24	0.81	0.77
Poverty Q2	1.10	0.32	<0.01	0.02	0.41	0.96
Poverty Q3	1.35	0.44	<0.01	-0.07	0.54	0.90
Poverty Q4	2.21	0.53	<0.01	0.48	0.67	0.48
Gini coefficient Q2	-0.27	0.22	0.21	0.17	0.30	0.57
Gini coefficient Q3	-0.45	0.29	0.12	0.06	0.42	0.89
Gini coefficient Q4	0.04	0.30	0.89	0.30	0.43	0.48
High income Q2	-0.35	0.20	0.08	-0.12	0.29	0.69
High income Q3	-0.41	0.26	0.12	-0.02	0.36	0.95
High income Q4	0.02	0.40	0.96	-0.51	0.56	0.36
Overdispersion (1/r)	0.05	0.03	0.01	0.20	0.06	<0.01

^aMedian income divided by 10,000.

^bStandard error.

parameter from 0.19 to 0.05. High levels of poverty were associated with higher rates of liver disease mortality ($\beta = 2.21, p < 0.01$).

For homicide mortality, accounting for other vulnerabilities and capacities of the neighborhoods greatly reduced the overdispersion parameter from 0.23 to 0.02 (Table 4). High-quality schools were associated with lower rates of homicide mortality ($\beta = -0.58, p = 0.01$), and high unemployment was associated with higher rates of homicide mortality ($\beta = 1.52, p < 0.01$). By contrast, adjusting for indicators of the distribution of individual socioeconomic status did not substantially reduce the overdispersion parameter (from 0.23 to 0.20).

Using homicide as an example, Figs. 1a–d illustrate the results of the modeling process described above. Fig. 1a shows the actual distribution of homicide rates when plotted against the median income of the neighborhood. As captured by the

overdispersion parameter ($1/r = 0.23$; Table 4, model (1)), there is substantial heteroskedasticity in the homicide rate; the variance of the rate is much higher among neighborhoods with low income, and is larger than the mean rate. Fig. 1b shows the predicted rates of homicide from the simple negative binomial model that includes only median income. Although this one predictor model describes the average association between median income and homicide, it does not explain the variability in these rates around the average. The predicted homicide rates from the negative binomial model including median income and the four additional vulnerabilities and capacities (Table 4, model (2)) are presented in Fig. 1c; accounting for these additional neighborhood characteristics explains a substantial amount of the variability in the rates of homicide observed in the actual data. This is reflected in the fact that the overdispersion parameter has been reduced from 0.23 to 0.02 (Table 4, comparing models (1) and (2)).

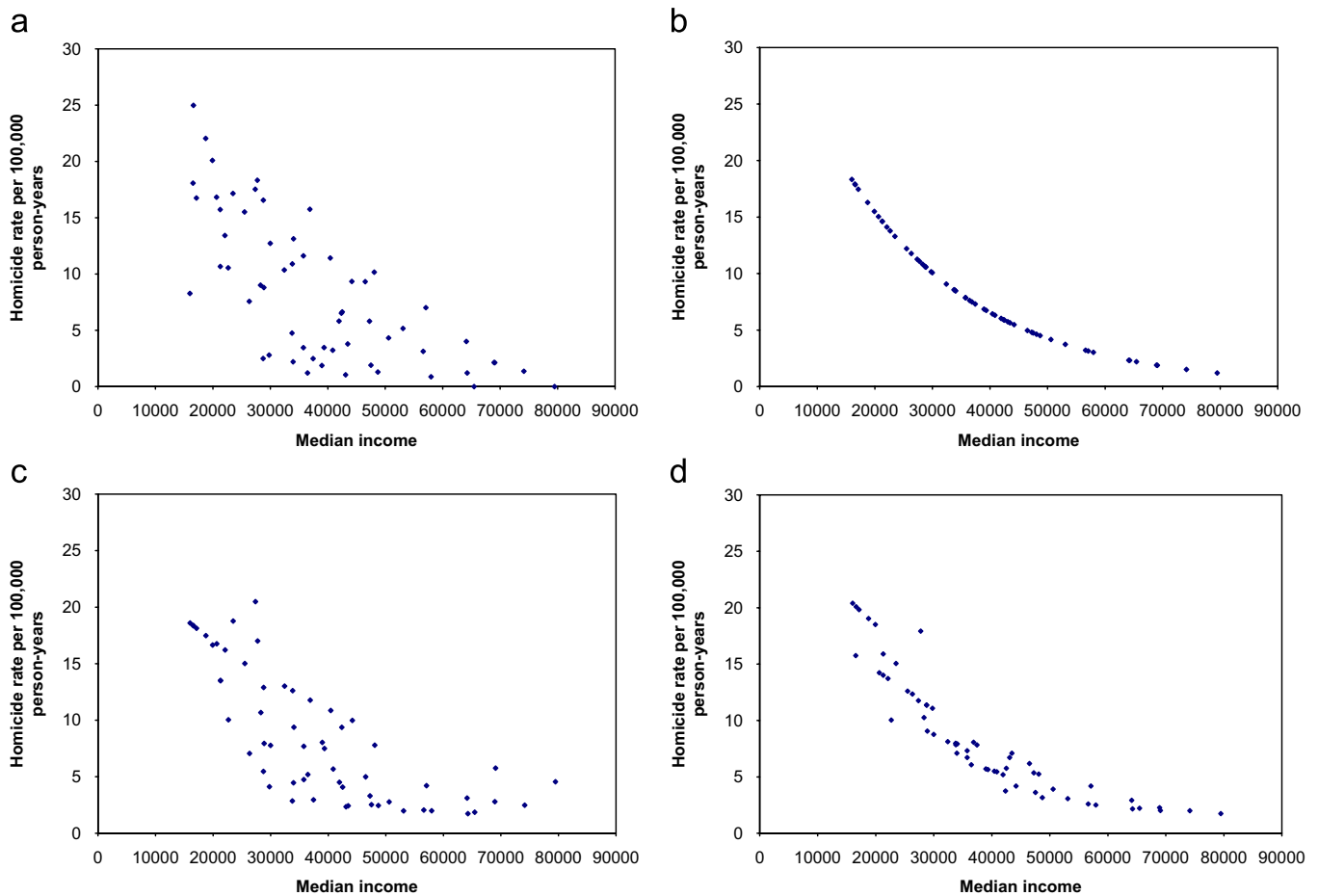


Fig. 1. (a) Actual homicide mortality rates by neighborhood median income (top left). (b) Predicted homicide rates from negative binomial model with only median income (top right). (c) Predicted homicide rates from negative binomial model with median income and vulnerabilities and capacities (bottom left). (d) Predicted homicide rates from negative binomial model with median income and socioeconomic distribution variables (bottom right).

The predicted homicide rates from the negative binomial model including median income and the four measures of the distribution of individual socioeconomic status (Table 4, model (3)) are presented in Fig. 1d; these predicted homicide rates suggest that accounting for the distribution of individual resources did not explain much of the variability in the homicide rates. This is confirmed by the minimal change in the overdispersion parameter from 0.23 to 0.20 (Table 4, comparing models (1) and (3)).

Discussion

We assessed whether variability in mortality rates, which is greater in neighborhoods characterized by socioeconomic vulnerability, was explained (1) by accounting for a range of other vulnerabilities and capacities of those neighborhoods, and/or (2) by adjusting for the distribution of individual socioeconomic status in those neighborhoods. We found that variability among socioeconomically vulnerable neighborhoods for all causes of mortality was explained to some extent by accounting for other vulnerabilities and capacities of the neighborhoods, while accounting for indicators of socioeconomic distribution explained some of the variability among the mortality rates, although to a lesser extent. These results clearly support the first hypothesis that variability in population health can be explained by accounting for other vulnerabilities and capacities of neighborhoods. However, these results also suggest that the distribution of individual socioeconomic resources may play a role in explaining variability in mortality.

A population view of health suggests that an individual—only perspective on health determination can lead to a failure to fully understand disease processes (Levins, 1998; Levins & Lopez, 1999). Understanding the constellation of vulnerabilities and capacities that characterize specific populations, and the contrasts between characteristics of different populations, may help us better understand what is driving health at a macro-level. For example, in areas with lower levels of income, other characteristics of those populations may exacerbate or temper the social environment, leading to more or less healthy conditions.

There are some examples where a focus on population level vulnerabilities and capacities that may affect health has been integrated into public health practice. For example, in New York City, the New York City Department of Health and Mental

Hygiene has created District Health Offices in some of the most vulnerable neighborhoods in the city that include an effort to work with local community members on decreasing individual health risks as well as improving underlying resources that might constitute important capacities for long-term population health (New York City Department of Health and Mental Hygiene, 2003). Comparable efforts in other places suggest that this type of thinking about population health may inform public health interventions such that they may be able to tackle structural and underlying factors at the same time as they are considering individual risks and behaviors, although there are certainly many challenges that must be faced in these endeavors (Merzel & D'Afflitti, 2003). While this analysis does not show that changing population level vulnerabilities and capacities causes changes in disease rates, it certainly suggests that this is plausible and that analyses in this area should be pursued.

Several considerations are important for interpretation of this study. We used an ecologic approach to examine the associations between neighborhood exposures and variability in health outcomes in this analysis, meaning that the exposures and outcomes were both at the neighborhood rather than individual levels, because only aggregated mortality data were available. One concern about the interpretation of ecologic associations is a question about the role of composition versus context (Duncan, Jones, & Moon, 1998; Greenland, 2001). Different populations may have different population level exposures, such as the rate of unemployment (context), but they are also composed of individuals with different exposures, such as individual employment status (composition), and an ecologic study does not allow these components to be distinguished. However, others have suggested that the distinction between a compositional and contextual effect may not be so clear because the composition is shaped by the context; for example, the level of unemployment is an indicator of the local labor market which certainly determines any individual's chances of employment (Macintyre, Ellaway, & Cummins, 2002). The ecologic nature of this study is essential to the interpretation of the findings; when we find that adjusting for the percent of residents who are unemployed accounts for variability in HIV/AIDS mortality between neighborhoods, we cannot determine whether that is because those who are unemployed have an increased risk of HIV/AIDS

mortality, or whether the level of unemployment puts the whole neighborhood population at increased risk of HIV/AIDS mortality. However, the findings are interesting nonetheless because from the population perspective it still suggests vulnerability, whether the whole population or a population subgroup is at increased risk when unemployment is high. It will also certainly be important for future multilevel studies to explore what can be learned about variability in population health when both population and individual characteristics are considered.

Due to the ecologic nature of the analysis, the findings related to the distribution of individual income in explaining variability in health outcomes have two potential interpretations. For example, our analyses suggest that among neighborhoods with a given level of median income, those with a larger group of very low-income individuals have higher rates of mortality from the causes studied here. These higher mortality rates may be observed because low-income groups within the populations have higher mortality rates, or because of some general population impact of having that very low-income group as part of the overall population. A general population effect would suggest the socioeconomic distribution affects the health of all residents, as suggested in the literature on income inequality as a predictor of health (Lynch, Davey Smith, Kaplan, & House, 2000). A multilevel study with data at the individual and group levels would be required to tease out these two possibilities.

In a cross-sectional analysis such as this, we cannot say that interventions such as bringing job opportunities and improving teacher salaries would improve health. A more complete understanding of the historical and current dynamics that have led to a certain situation in one population would certainly require longitudinal analysis, and would be greatly informed by in depth qualitative assessment of that population (Levins & Lopez, 1999; Merzel & D'Afflitti, 2003; Philippe & Mansi, 1998). Longitudinal research that examines changes in population level exposures as well as outcomes may be an important future direction for research in this area.

In addition, we defined our populations of interest as urban neighborhoods. Individuals belong to many groups, each of which may constitute relevant populations (e.g., social networks, work communities). A full understanding of population health may require understanding how character-

istics of these different population groups overlap and interrelate.

Our conceptual model states that underlying vulnerabilities and capacities will be drawn upon in the face of intermittent stressors and protective events that occur in populations. In this analysis, we have only examined vulnerabilities and capacities and do not have information on the stressors that might have occurred in the populations under study, and as such, we may only be seeing a shadow of the larger process proposed in our conceptual model. Future research that measures both the vulnerabilities and capacities of populations and stressors and protective events such as natural disasters or the availability of new job opportunities is important for more complete assessment of the proposed model.

For the mortality rates studied in this analysis there is likely some dependence of the outcomes between individuals, in which case a traditional analysis of ecological or individual level data does not accurately reflect how much of a change in mortality would be expected from a change in exposure (Koopman & Longini, 1994). A complementary analytic approach that merits consideration for future analyses of variability in population health would be one that can account for these dynamic processes, such as the systems modeling approaches that have been used to model infectious diseases (Koopman & Longini, 1994; Ness, Koopman, & Roberts, 2007).

Finally, we included in our analysis four potential vulnerabilities and capacities, and four indicators of the distribution of individual socioeconomic status. There may be other population characteristics of interest and other ways to operationalize the characteristics studied. We based our selection of characteristics on previous conceptual work on Census-based indicators of neighborhood socioeconomic status, income distribution, education, and employment (Krieger et al., 2002; Wagstaff et al., 1991), as well as on available measures from data collected by New York City agencies that had conceptually and empirically supported associations with population vulnerabilities (Greiner et al., 2004; Jacob, 2007; Kawachi & Berkman, 2000). For the Census-based vulnerability and capacity measures, we used education and employment which have been shown to load on a different factor than economic resources such as income and poverty, suggesting these could be considered to measure distinct constructs (Krieger et al., 2002). Nevertheless, one

challenge faced in this type of work is to tease apart those characteristics that may be fundamental vulnerabilities and capacities, from those that may be collinear with the important underlying characteristics (Rose, 1992). Further work to conceptualize underlying vulnerabilities and capacities, to optimize their assessment, and to understand their interrelationship will be essential if we hope to develop population level public health interventions aimed at improving vulnerabilities or bolstering capacities in the future.

We showed that underlying socioeconomic vulnerabilities and capacities of populations shaped heterogeneity in mortality rates from a variety of causes. In addition, the distribution of individual socioeconomic resources in populations contributed to heterogeneity in some causes of mortality. Understanding the constellation of underlying factors that come together in a population may provide more information on how population health conditions develop, how they will respond to perturbations such as stressful or protective events, and eventually inform us about ways in which we may improve population health.

Funding sources

Funded in part by grants R01 DA 017642 and DA 022720 from the National Institutes of Health.

References

- Ahern, J., & Galea, S. (2006). Social context and depression after a disaster: The role of income inequality. *Journal of Epidemiology and Community Health*, 60(9), 766–770.
- Diez Roux, A. V. (2000). Multilevel analysis in public health research. *Annual Reviews of Public Health*, 21, 171–192.
- Duncan, C., Jones, K., & Moon, G. (1998). Context, composition and heterogeneity: Using multilevel models in health research. *Social Science & Medicine*, 46, 97–117.
- Galea, S., Ahern, J., & Karpati, A. (2005). A model of underlying socioeconomic vulnerability in human populations: Evidence from variability in population health and implications for public health. *Social Science & Medicine*, 60, 2417–2430.
- Galea, S., Ahern, J., Vlahov, D., Coffin, P. O., Fuller, C., Leon, A. C., et al. (2003). Income distribution and risk of fatal drug overdose in New York City neighborhoods. *Drug and Alcohol Dependence*, 70(2), 139–148.
- Greenland, S. (2001). Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *International Journal of Epidemiology*, 30, 1343–1350.
- Greiner, K. A., Li, C., Kawachi, I., Hunt, D. C., & Ahluwalia, J. S. (2004). The relationships of social participation and community ratings to health and health behaviors in areas with high and low population density. *Social Science & Medicine*, 59, 2303–2312.
- Hamilton, V. L., Broman, C. L., Hoffman, W. S., & Renner, D. S. (1990). Hard times and vulnerable people: Initial effects of plant closing on autoworkers' mental health. *Journal of Health and Social Behavior*, 31(2), 123–140.
- Hamilton, V. L., Hoffman, W. S., Broman, C. L., & Rauma, D. (1993). Unemployment, distress, and coping: A panel study of autoworkers. *Journal of Personality and Social Psychology*, 65(2), 234–247.
- Hembree, C., Galea, S., Ahern, J., Tracy, M., Markham Piper, T., Miller, J., et al. (2005). The urban built environment and overdose mortality in New York City neighborhoods. *Health and Place*, 11(2), 147–156.
- Hubbard, A., & Jewell, N. P. (2007). *Analysis of longitudinal studies in epidemiology*. Chapman & Hall.
- Jacob, B. A. (2007). The challenges of staffing urban schools with effective teachers. *The Future of Children*, 17(1), 129–153.
- Karpati, A., Galea, S., Awerbuch, T., & Levins, R. (2002). Variability and vulnerability at the ecologic level: Implications for understanding the social determinants of health. *American Journal of Public Health*, 92, 1768–1772.
- Kawachi, I., & Berkman, L. (2000). Social cohesion, social capital, and health. In L. Berkman, & I. Kawachi (Eds.), *Social epidemiology* (pp. 174–190). New York: Oxford University Press.
- Koopman, J., & Longini, I. (1994). The ecological effect of individual exposures and nonlinear disease dynamics in populations. *American Journal of Public Health*, 84, 836–842.
- Koopman, J., & Lynch, J. (1999). Individual causal models and population system models in epidemiology. *American Journal of Public Health*, 89, 1170–1174.
- Krieger, N., Chen, J. T., Waterman, P. D., Soobader, M. J., Subramanian, S. V., & Carson, R. (2002). Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: Does the choice of area-based measure and geographic level matter? The Public Health Disparities Geocoding Project. *American Journal of Epidemiology*, 156(5), 471–482.
- Krieger, N., Zierler, N., Hogan, J. W., Waterman, P., Chen, J., Lemieux, K., et al. (2003). geocoding and measurement of neighborhood socioeconomic position: A US perspective. In I. Kawachi, & L. F. berkman (Eds.), *Neighborhoods and health*. Oxford: Oxford university press.
- Laraia, B. A., Messer, L., Kaufman, J. S., Dole, N., Caughy, M., O'Campo, P., et al. (2006). Direct observation of neighborhood attributes in an urban area of the US south: Characterizing the social context of pregnancy. *International Journal of Health Geographics*, 5, 11.
- Levins, R. (1975). Evolution in communities near equilibrium. In M. Cody, & J. Diamond (Eds.), *Ecology and evolution of communities* (pp. 16–50). Cambridge, UK: The Belknap Press of Harvard University Press.
- Levins, R. (1998). The internal and external in explanatory theories. *Science as Culture*, 7, 557–582.
- Levins, R., & Lopez, C. (1999). Toward an ecosocial view of health. *International Journal of Health Services*, 29, 261–293.
- Lynch, J. W., Davey Smith, G., Kaplan, G. A., & House, J. S. (2000). Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. *British Medical Journal*, 320, 1200–1204.

- Macintyre, S., Ellaway, A., & Cummins, S. (2002). Place effects on health: how can we conceptualise, operationalise and measure them? *Social Science & Medicine*, 55, 125–139.
- Manhattan Community Board 1 <<http://www.cb1.org>>.
- McNeill, L. H., Kreuter, M. W., & Subramanian, S. V. (2006). Social environment and physical activity: A review of concepts and evidence. *Social Science & Medicine*, 63(4), 1011–1022.
- Merzel, C., & D'Afflitti, J. (2003). Reconsidering community-based health promotion: Promise, performance, potential. *American Journal of Public Health*, 93, 557–574.
- Ness, R. B., Koopman, J. S., & Roberts, M. S. (2007). Causal system modeling in chronic disease epidemiology: A proposal. *Annals of Epidemiology*.
- New York City Department of Health and Mental Hygiene. 2000. Office of Vital Statistics. *Summary of Vital Statistics 2000: The City of New York*.
- New York City Department of Health and Mental Hygiene. 2003. *Mayor Michael R. Bloomberg and Health Commissioner Thomas R. Frieden announce launch of district public health offices in neighborhoods with most pressing health problems*.
- New York City Mayor's Office of Operations. 2002. *Mayor's Management Report, Fiscal*.
- Papas, M. A., Alberg, A. J., Ewing, R., Helzlsouer, K. J., Gary, T. L., & Klassen, A. C. (2007). The built environment and obesity. *Epidemiologic Reviews*.
- Perkins, D. D., & Taylor, R. B. (1996). Ecological assessments of community disorder: Their relationship to fear of crime and theoretical implications. *American Journal of Community Psychology*, 24(1), 63–107.
- Philippe, P., & Mansi, O. (1998). Nonlinearity in the epidemiology of complex health and disease processes. *Theoretical Medicine and Bioethics*, 19, 591–607.
- Pickett, K. E., & Pearl, M. (2001). Multilevel analyses of neighbourhood socioeconomic context and health outcomes: A critical review. *Journal of Epidemiology and Community Health*, 55, 111–122.
- Rockhill, B. (2005). Theorizing about causes at the individual level while estimating effects at the population level. *Epidemiology*, 16, 124–129.
- Rose, G. (1992). *The strategy of preventive medicine*. Oxford: Oxford University Press.
- Rose, G. (2001). Sick individuals and sick populations. *International Journal of Epidemiology*, 30, 427–432.
- Rouse, C. E., & Barrow, L. (2006). US elementary and secondary schools: Equalizing opportunity or replicating the status quo? *The Future of Children*, 16(2), 99–123.
- Sellstrom, E., & Bremberg, S. (2006). The significance of neighbourhood context to child and adolescent health and well-being: A systematic review of multilevel studies. *Scandinavian Journal of Public Health*, 34(5), 544–554.
- Syme, S. L. (1996). Rethinking disease: Where do we go from here? *Annals of Epidemiology*, 6, 463–468.
- US Census Bureau (2000). *Census summary tape, file 3A (STF3A)*. Washington, DC: US Department of Commerce.
- Wagstaff, A., Paci, P., & van Doorslaer, E. (1991). On the measurement of inequalities in health. *Social Science & Medicine*, 33(5), 545–557.