Cause-specific mortality and income inequality in São Paulo, Brazil

Mortalidade segundo causas básicas e desigualdade de renda no Município de São Paulo

ABSTRACT

OBJECTIVE: To analyze cause-specific mortality rates according to the relative income hypothesis.

METHODS: All 96 administrative areas of the city of São Paulo, southeastern Brazil, were divided into two groups based on the Gini coefficient of income inequality: high (≥ 0.25) and low (<0.25). The propensity score matching method was applied to control for confounders associated with socioeconomic differences among areas.

RESULTS: The difference between high and low income inequality areas was statistically significant for homicide (8.57 per 10,000; 95%CI: 2.60;14.53); ischemic heart disease (5.47 per 10,000 [95%CI 0.76;10.17]; HIV/AIDS (3.58 per 10,000 [95%CI 0.58;6.57]); and respiratory diseases (3.56 per 10,000 [95%CI 0.18;6.94]). The ten most common causes of death accounted for 72.30% of the mortality difference. Infant mortality also had significantly higher age-adjusted rates in high inequality areas (2.80 per 10,000 [95%CI 0.86;4.74]), as well as among males (27.37 per 10,000 [95%CI 6.19;48.55]) and females (15.07 per 10,000 [95%CI 3.65;26.48]).

CONCLUSIONS: The study results support the relative income hypothesis. After propensity score matching cause-specific mortality rates was higher in more unequal areas. Studies on income inequality in smaller areas should take proper accounting of heterogeneity of social and demographic characteristics.

The association between income and health can be assessed through two distinct mechanisms: the absolute income effect and the relative income effect. The absolute income effect is exemplified by the well-established association between income, poverty and death and illness. Those living in poverty have lower access to health services, clean water, secure jobs, decent education, as well as are more vulnerable to violence and natural disasters.18

Conversely, the relative income hypothesis posits that an individual’s health status is additionally determined by his/her relative social position, which depends on the income level. Those living in poverty have lower access to health services, clean water, secure jobs, decent education, as well as are more vulnerable to violence and natural disasters.18

From the perspective of social epidemiology, one empirical test of the relative income hypothesis is whether the local income distribution has an independent effect on mortality and morbidity above and beyond the effect of absolute income. To date, empirical tests have linked income distribution and health.15,19 A recent meta-analysis of multilevel studies found a statistically significant association between higher income inequality and excess mortality and poor self-rated health; each 0.05 unit increase in the Gini coefficient
was associated with a 7.8% excess mortality risk (95% CI 5.9;9.8). However, the size of the area for which income inequality is relevantly associated with health outcome is still a matter of debate. According to an earlier review by Subramanian & Kawachi, the association between income inequality and health is more robust for studies conducted in larger (in particular, U.S. states) than smaller areas (neighborhoods). This observation is inconsistent with the relative income theory which posits that social comparisons are likely to be more significantly felt as the individual is proximate to reference groups (i.e. when compared against one’s immediate neighbors versus other people living in the same state).

Before refuting this theory we should note an important methodological issue of testing the relative income hypothesis in smaller areas such as neighborhoods. In societies with high income segregation such as in Brazil, the distribution of income within a neighborhood is more equal than that between neighborhoods. Thus, income distribution within a favela (urban slum) tend to be far more equal (i.e. everybody is equally poor) compared to the rest of the metropolitan area. Thus, relevant social comparisons may not be fully valid by focusing on the distribution of income within a neighborhood.

Another challenge to the relative income theory is the mechanism by which inequality may have an effect on health. The psychosocial theory of inequality posits that the harmful effects of inequality are mediated by stress. But according to a recent literature review, five out of nine studies found little or no effect of stress as a potential mediator between socioeconomic status and health. This inconsistency in the literature may partly be explained by the crude manner in which “stress” is often measured. An alternative approach is to examine the specific pattern of excess morbidity and mortality associated with inequality to assess whether it is consistent with a stress-related mechanism. For example, accumulated evidence at the individual level points to an association between stress and cardiovascular disease but there is less evidence for cancers.

Wilkinson & Pickett analyzed seven different cause-specific mortality rates from 3,139 counties in the United States and found that income inequality was associated to ischemic heart disease, respiratory disease, and homicide.

The objective of the present study was to analyze cause-specific mortality differences according to the relative income theory using a methodological approach that takes into account local heterogeneity.

**METHODS**

This analysis was developed by focusing on the two aforementioned challenges. First, we applied a statistical method known as propensity score matching to control for potential confounders of the association between income inequality and health, as highly deprived neighborhoods can have relatively equal distribution of income (because everyone is equally poor). In such a case, it would be misleading to compare people’s health in unequal vs. equal neighborhoods regardless of absolute local characteristics. Second, we identified cause-specific patterns of mortality associated with income inequality. Preliminary results of the main study are published elsewhere.

São Paulo, southeastern Brazil, is a very unequal and segregated city. In the past, the city was considered to have a radial distribution of income, the farther away from the center the poorer its residents. As population density increased wealthier residents came to prefer more peripheral areas; and the need for construction workers, and domestic and other low-paid laborers made them move to these areas as well. Extreme income inequality between closely-adjacent areas became an important characteristic of São Paulo.

São Paulo is currently divided into 31 administrative areas, known as subprefeituras. They are further divided into administrative areas, a proxy for neighborhood and the smallest areas for which health data is available. São Paulo has a total of 96 administrative areas (median population of 98,649 residents) covering the entire area of the city.

We included 16 administrative area-level variables in the model: prevalence of favelas (proportion of households classified as “subnormais” or slums); poverty rate (proportion of residents living in households with an income of less than half of the monthly minimum wage per capita); median per capita income; education of head of household (measured by years of schooling); household density (average number of people living in the same household); proportion of households with tap water; proportion of households with garbage collection; proportion of households with no sewage system; proportion of heads of household under 21 years old; illiteracy rate of heads of household; illiteracy rate of 8–12 years old; proportion of teachers per student (5th to 8th grade); HIV/AIDS rate; proportion of infants (under 1 year old); proportion of elderly (over 64 years old); and proportion of women. Data was obtained from the 2000 Population Census, as well as from the 2001 School Census (proportion of teachers per student) and the São Paulo Epidemiological...
To avoid confusion between income inequality area and poverty area, we applied propensity score matching. A propensity score is the probability of assignment to a particular treatment or exposure given a set of observed covariates. It was first introduced by Rosenbaum & Rubin.\(^\text{xii}\)

The probability of assignment to a “treatment”, or “exposure” (in this instance, a neighborhood with high income inequality) is estimated using logistic regression in which we enter possible (and observed) causes of assignment to the treatment. The propensity score reduces a large set of covariates to a single scalar summary variable but it does not make assessments about the prediction of individual covariates to a single scalar summary variable.\(^*\) The propensity score approach is intended to identify administrative areas that are as alike as possible to each other with respect to the probability of being “exposed.”

The Gini coefficient was used to measure inequality of income distribution. It theoretically ranges from 0.0 (perfect equality, with every household earning exactly the same) to 1.0 (absolute inequality, with a single household earning a locality’s entire income). Mathematically, the Gini coefficient is equivalent to half the average absolute difference between the incomes of any two households randomly sampled for a population, and then normalized to the mean. In the present analyses, we defined our exposure as having relative high income inequality (Gini coefficient >0.25), and non-exposure as relative low income inequality (<0.25). As there is no clear cut-off for high vs. low income inequality, we set it at 0.25 as it was close to the median Gini coefficient of the areas studied. The overall Gini coefficient in our sample ranged from 0.12 (Jaguará) to 0.55 (Vila Andrade).

The propensity score ranges from 0 (high probability of low income inequality within an area) to 1 (high probability of high income inequality). The variables selected in our model generated propensity scores considered to be “highly predictive” of the exposure\(^*\) (the area under the ROC curve, or c-statistic, was 0.907).

The administrative areas with different exposures were subsequently matched on their propensity scores using a caliper width of 0.01—the most frequently used value in the literature. If an area was not within a caliper width of any other with an opposite exposure situation, it was excluded from the analysis. We chose to apply matching with replacement to minimize loss of areas in the analysis. Of a total of 96 administrative areas in the city of São Paulo, 27 different areas had propensity scores within a caliper width of 0.01 of another one with opposite exposure, and were included in the analysis. The “exchangeable” areas were matched based on their propensity scores according to exposure status.

RESULTS

Before using propensity score matching (which included all 96 administrative areas) there was no significant mortality difference between high and low income inequality areas among infants, elderly, males and females (Table 1). When we analyzed the underlying causes of death, there was statistically significant different age-adjusted mortality only from cancer (higher on high inequality areas) and diabetes (higher on low inequality areas).

After using propensity score matching four of the 10 most common underlying causes of death had significantly higher mortality rates on high than on low inequality income areas (Table 2). The greatest difference was for homicide (8.57 per 10,000 [95%CI 2.60; 14.53]), followed by ischemic heart disease (5.47 per 10,000 [95%CI 0.76; 10.17]), HIV/AIDS (3.58 per 10,000 [95%CI 0.58; 6.57]) and respiratory diseases (3.56 per 10,000 [95%CI 0.18; 6.94]). All 10 underlying causes analyzed had higher mortality rates in high inequality areas.

Infant mortality also showed significantly higher age-adjusted mortality in high inequality areas (2.80 per 10,000 [95%CI 0.86; 4.74]), as well as male mortality (27.37 per 10,000 [95%CI 6.19;48.55]) and female mortality (15.07 per 10,000 [95%CI 3.65;26.48]). Elderly mortality was also higher in high income inequality areas (8.85 per 10,000), although non-significant (95%CI -0.72;18.44).
The analysis of the relative importance of each underlying cause of death to the total mortality difference between high and low inequality areas (41.58 per 10,000) showed that homicides accounted for 20.60% of excess mortality, ischemic heart disease for 13.16% and HIV/AIDS for 8.61% (Figure). The 10 most common causes of death accounted for 72.30% of the total mortality difference.

DISCUSSION

More unequal areas had higher mortality rates from all 10 underlying causes of death analyzed. Four causes of death were statistically significant (homicide, ischemic heart disease, HIV/AIDS, and respiratory diseases). Excess mortality of both males and females as well as excess infant mortality were found. Our results are consistent with those reported by Wilkinson & Pickett, in which income inequality was statistically significantly associated with mortality from ischemic heart disease, homicide and respiratory diseases, though not from diabetes or cancer.

The psychosocial hypothesis posits that the relative distribution of income within an area adversely affects health because of stress and shame induced by invidious social comparisons between residents. This may explain the excess mortality from ischemic heart disease, which has a well-established association with stress. Growing evidence suggests that stress is also linked to pulmonary function. Furthermore, the majority of respiratory deaths in our study are likely to be due to chronic obstructive pulmonary diseases, which are in turn primarily associated to cigarette smoking. Excess deaths from respiratory disease can be plausibly linked to higher psychosocial stress and smoking rates in the population studied.

Homicide and HIV/AIDS can similarly reflect behavioral responses to people’s feelings of missed life opportunities. Youth growing up in areas with high income segregation (and consequently low mobility) feel they have no future in life, which in turn increases the probability of higher risk-taking behavior such as joining street gangs or unprotected sex.

São Paulo is a relevant area to study the effects of relative income on health because it is a city with high income segregation. Even though it has a Gini coefficient of 0.51, the median value for its 96 administrative areas is 0.25. Living in poverty in São Paulo is not associated with living in an unequal neighborhood. On the contrary, the 14 poorest areas had a Gini coefficient below 0.25, i.e., poorer neighborhoods have a more equal distribution of income, i.e. their residents are equally poor.

A reason for the missing consensus on the adverse health effect of income inequality for small areas such as neighborhoods may be the homogeneity of income within small geographic scales. Our analysis showed that even for a notoriously unequal city such as São Paulo, this is still an issue. Only one area stood out as highly unequal (Vila Andrade), with a Gini coefficient of 0.55.

The propensity score matching approach showed some advantages. First, it allows to explicitly check for overlap in the confounder distribution prior to the analysis—a step that is sometimes skipped when conducting “black box” multivariable-adjusted regression. The exposed and unexposed groups are then matched, and comparisons are drawn only within the range of overlap in propensity scores, thereby avoiding off-support inference. Secondly, the propensity score deals with the dimensionality aspect in the multivariable analysis by reducing a large set of covariates to a single scalar variable. Lastly, this approach does not rely on any particular functional form (e.g. linearity) between the covariate and the outcome within each treatment group.
A major limitation of the propensity score approach is that it deals only with the set of observed covariates to build the model. It does not address the issue of confounding by unobserved factors, and in the worst case, the approach may even lead to increased bias. Nor can the approach deal with the endogeneity of residential preferences, which is an issue in causal inference that applies to all observational studies on effects of area of residence on health. The replication of the results in other cities may also be an issue. Although São Paulo is one of the largest cities in the world, matches were found only for 27 of the 96 administrative areas. The use of binary exposure status (which is a limitation of the propensity score matching approach) may also caused loss of information. The categorization of the level of income inequality into three levels (high, low, and medium) rather than using the median might have strengthened the associations. On the other hand, precision would have been compromised because of further reduction in the sample size of matched administrative areas.

The use of aggregate data is also an important limitation because it does not allow an analysis of individual biological pathways by which income inequality affects health. Another potential limitation is the use of a politically defined area such as administrative areas (distritos), which does not take into account spatial autocorrelation (i.e. rich areas may border slums but...
the two would be classified into different administrative areas). One of the 16 variables included in the propensity score model was reported HIV/AIDS incidence rates for each administrative area. This variable was included because HIV/AIDS is a marker of deprivation and this condition affects an individual’s earning ability and hence can be a contributing factor to income inequality. In the analyses in which HIV/AIDS mortality was used as the outcome, HIV/AIDS was represented in both the left-hand and right-hand side of the regression equation. In these models, we interpreted the difference in HIV/AIDS mortality rate as conditional on the area-level differences in HIV/AIDS rates.

The multivariable approach was required due to the characteristically high heterogeneity in São Paulo. There is no identifiable pattern of distribution of resources and population that makes one part of the city easily comparable to another one. This is also true for many other metropolitan areas situated in highly unequal countries (such as the United States and most of the developing world). We argue that the null findings linking income inequality and health can benefit from the application of a methodology such as the propensity score matching that takes into account the confounding effects of variables other than income inequality that could affect health.

REFERENCES


6. Conniffe D, Gash V, O’Connell PJ. Evaluating state


The authors declare no conflicts of interest.