**Universidade de São Paulo
Faculdade de Saúde Pública**

**Efeito da desigualdade de renda na mortalidade do
Município de São Paulo**

Alexandre Dias Porto Chiavegatto Filho

**Tese apresentada ao Programa de Pós-
Graduação em Saúde Pública para
obtenção do título de Doutor em Saúde
Pública.**

Área de Concentração: Epidemiologia.

**Orientadora: Profa. Dra. Sabina Léa
Davidson Gotlieb.**

**São Paulo
2010**

Efeito da desigualdade de renda na mortalidade do Município de São Paulo

Alexandre Dias Porto Chiavegatto Filho

Tese apresentada ao Programa de Pós-Graduação em Saúde Pública da Faculdade de Saúde Pública da Universidade de São Paulo para obtenção do título de Doutor em Saúde Pública.

Área de Concentração: Epidemiologia.

Orientadora: Profa. Dra. Sabina Léa Davidson Gotlieb.

**São Paulo
2010**

É expressamente proibida a comercialização deste documento, tanto na sua forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano da tese.

DEDICATÓRIA E AGRADECIMENTOS

Aos meus pais Tina e Alexandre, pela educação que me deram, além do apoio e confiança. À minha irmã Marina, uma das pessoas mais sábias que conheço. À minha avó Leda, meu exemplo pessoal de pioneirismo e coragem. À minha nonna Angela, que me ensinou muito sobre o Brasil ao definir-se uma vez para mim como brasileira.

A todos do Centro Brasileiro de Classificação de Doenças. Ao Prof. Ruy Laurenti (meu primeiro orientador e principal responsável pelo meu gosto pela ciência), à Dona Miriam (mãe de todos nós), à Profa. Maria Helena e, claro, à minha orientadora Sabina (melhor professora que conheço), sem ela nada disso seria possível. Agradeço também ao Mário por me ter indicado à vaga de iniciação científica em 2003.

Ao Ichiro Kawachi, meu co-orientador de Harvard, que há 5 anos é meu ídolo na área acadêmica, mesmo antes de ter trabalhado com ele. Às funcionárias do Departamento de *Society, Human Development and Health* da *Harvard School of Public Health*.

Aos professores do Departamento de Epidemiologia da Faculdade de Saúde Pública, principalmente àqueles que me deram a oportunidade de monitorar alguma disciplina. Aos funcionários da Pós-Graduação e às funcionárias da Biblioteca da Faculdade de Saúde Pública que sabem meu número USP de cor.

À Lívia, Milena e Dani, que tantos bandejões seguido de café compartilharam comigo, divertindo-se às custas das minhas histórias. À Paulinha, por ter crescido

junto comigo. À Ana María Nieto, que não acredita em ciência. À Carol, por me ter incentivado a não me contentar com pouco.

A todos os meus amigos da FEA, da FSP e de Harvard. Prefiro não mencionar cada um porque eu sempre tentei manter a tese o mais longe possível das conversas do dia e também porque sei que para eles poder dar algumas risadas comigo é o suficiente.

Agradeço também à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro dado pela bolsa de doutorado direto e pela bolsa sanduíche.

*Ill fares the land, to hastening ills a prey,
Where wealth accumulates, and men decay.*

Oliver Goldsmith, *The Deserted Village* (1770)

Who can say for sure that the deprivation which afflicts him with hunger is more painful than the deprivation which afflicts him with envy of his neighbor's new car?

John Kenneth Galbraith, *The Affluent Society* (1998).

RESUMO

Introdução - A maioria dos estudos sobre a teoria da renda relativa tem encontrado associação estatisticamente significativa entre alta desigualdade de renda e piores condições de saúde em grandes regiões como estados norte-americanos ou países. Por outro lado, análises realizadas em municípios e áreas fora dos EUA tem apresentado resultados no mínimo conflitantes. Os maiores obstáculos para que se atinja consenso na área são a inexistência de elevada desigualdade em regiões menores e a dificuldade de controlar pelo conjunto de variáveis de confusão que podem ter efeito na saúde além da desigualdade de renda. O presente estudo objetiva auxiliar na solução desse problema ao analisar o Município de São Paulo por meio de uma metodologia estatística chamada *propensity score matching*. **Metodologia** - A análise abordou os 96 distritos da cidade. Foram incluídas 16 variáveis no modelo para identificar distritos comparáveis. Do total de 96 distritos, 27 foram pareados com algum outro (alguns mais de uma vez), formando 17 pares, dos quais apenas um foi composto por distritos que fazem fronteira entre si. **Resultados** - Após a aplicação do *propensity score matching*, distritos mais desiguais apresentaram maior mortalidade geral ajustada por idade (41,58 por 10.000 hab; IC 95%: 8,85 – 73,3 por 10.000 hab). Foram também estatisticamente significativas as diferenças de mortalidade para homicídios (8,57 por 10.000; IC 95%: 2,60 – 14,53), doença isquêmica do coração (5,47 por 10.000; IC 95%: 0,76 – 10,17), aids (3,58 por 10.000; IC 95%: 0,58 – 6,57), doença respiratória (3,56 por 10.000; IC 95%: 0,18 – 6,94) e mortalidade infantil (2,8 por 10.000; IC 95%: 0,86 – 4,74). As dez causas

básicas mais frequentes foram responsáveis por 72,3% do total da diferença de mortalidade entre os distritos mais e menos desiguais. **Conclusões** - A metodologia estatística foi eficaz para diminuir as diferenças sociais e demográficas, possibilitando a comparação entre distritos semelhantes. A presença de alta desigualdade de renda no Município de São Paulo permitiu a análise do seu efeito na mortalidade.

Descritores: Epidemiologia social, Desigualdade de renda, Mortalidade, *Propensity score*.

ABSTRACT

Introduction - The majority of studies on the relative income theory has detected statistically significant associations between high income inequality and worse health conditions for larger areas such as countries or USA states. Nevertheless, for smaller areas such as municipalities or regions outside the USA, the results have been, at best, mixed. The biggest hindrances to a consensus are the lack of high inequalities within smaller areas and the difficulty to control for the many variables that may also affect health beyond the effect of income inequality. This analysis aims to help to solve these problems by applying a statistical analysis known as propensity score matching and by focusing on a very diverse and unequal city such as São Paulo.

Methodology - The analysis was done for the 96 *distritos* of the Municipality of São Paulo, Brazil. The statistical model included 16 variables to account for local heterogeneity. Of the 96 *distritos*, 27 were matched with a similar one (sometimes, more than once), making a total of 17 pairs (of those, only one pair was made out of two bordering *distritos*). **Results** - After the propensity score matching approach, higher inequality *distritos* had higher age-adjusted overall mortality rate (41.58 per 10,000, 95% CI: 8.85 – 73.3). The difference between high and low inequality was also 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), respiratory diseases (3.56 per 10,000; 95% CI: 0.18 – 6.94) and infant mortality (2.8 per 10,000; 95% CI: 0.86 – 4.74). The ten most common causes of death accounted for 72.3% of total difference in mortality.

Conclusion - The statistical methodology was effective to control for local social and demographic heterogeneity, allowing the comparison of only similar *distritos*. The use of a large and diverse city such as São Paulo made it possible to analyze the effects of income inequality on health.

Descriptors: Social epidemiology; Income Inequality; Mortality; Propensity Score.

APRESENTAÇÃO

Esta tese foi escrita e apresentada para a obtenção do título de doutor no Programa de Pós-Graduação em Saúde Pública, área de concentração Epidemiologia, no Departamento de Epidemiologia da Faculdade de Saúde Pública da Universidade de São Paulo.

O método escolhido para a apresentação da tese foi sob a forma de artigos científicos. Segundo as diretrizes deliberadas na Comissão de Pós-Graduação da FSP em sua sessão 9ª/2008, “a Tese de Doutorado deve incluir, no mínimo, três manuscritos resultantes do seu projeto de pesquisa de doutorado, sendo que um deles não deve ter sido submetido ou publicado em periódico ou capítulo de livro. Os outros dois artigos devem ter sido submetidos ou publicados em periódicos ou capítulos de livro arbitrados por pares, no período de sua matrícula no Programa” (ANEXO 1).

A pesquisa contou com o apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) de 2007 a 2010. Parte da pesquisa foi realizada no *Department of Society, Human Development and Health* da *School of Public Health* da *Harvard University*, sob a orientação do Prof. Ichiro Kawachi, com auxílio financeiro da bolsa sanduíche, também da CAPES.

Por se tratar de uma tese sob a forma de artigos, procurou-se evitar a repetição de informações e dados o máximo possível. Esse objetivo não foi sempre cumprido dentro de cada artigo científico por terem sido publicados em revistas e em

períodos diferentes, em vez de em sequência, como é o caso da apresentação desta tese.

Primeiramente é feita uma introdução ao tema, onde são esclarecidos os postulados da teoria da renda relativa, novos resultados de revisões sobre o tema, os potenciais fatores de risco da desigualdade para a saúde, críticas à teoria, situação da literatura no Brasil e possíveis mecanismos de ação.

O primeiro artigo ainda não foi enviado para publicação, segundo as regras mencionadas pela CPG. Pretende-se com ele identificar os padrões de desigualdade dentro do Município de São Paulo e analisar o uso do *propensity score matching* no pareamento dos distritos. Serão também apresentados mapas para a visualização da distribuição do índice de Gini, do valor de *propensity score* e dos coeficientes de mortalidade geral ajustados por idade nos 96 distritos.

O segundo artigo foi aprovado para publicação no *Journal of Epidemiology and Community Health* (Qualis A1 em Saúde Coletiva segundo a CAPES) e a previsão da publicação da versão impressa está atualmente para março de 2011. O artigo apresenta em detalhes os resultados do *propensity score matching* e aponta que a metodologia foi eficaz na diminuição da variabilidade dos distritos comparados. Os resultados indicam uma diferença estatisticamente significativa na mortalidade entre distritos desiguais e mais igualitários.

O terceiro artigo foi enviado para publicação na *Annals of Epidemiology* (Qualis A2 em Saúde Coletiva segundo a CAPES) e no momento do depósito desta tese aguarda parecer da revista. O artigo analisou as dez causas básicas mais comuns no Município de São Paulo, além da mortalidade infantil, da mortalidade de idosos e

da mortalidade segundo sexo. As causas básicas que apresentaram diferenças estatisticamente significativas entre distritos mais desiguais foram homicídios, doença isquêmica do coração, aids e doenças respiratórias.

A tese finaliza com alguns comentários, que tem o objetivo de complementar as conclusões presentes nos artigos.

ÍNDICE

1	INTRODUÇÃO	16
	1.1 DESIGUALDADE E VIDA EM SOCIEDADE	18
	1.2 DESIGUALDADE E FATORES DE RISCO	23
	1.3 DESIGUALDADE E SAÚDE: MECANISMOS DE AÇÃO	25
	1.4 DESIGUALDADE E SAÚDE NO BRASIL	27
2	OBJETIVOS	29
	2.1 GERAL	29
	2.1 ESPECÍFICOS	29
3	METODOLOGIA	30
4	RESULTADOS E DISCUSSÃO	31
	ARTIGO 1: O uso do propensity score matching na análise da distribuição espacial de características demográficas e sociais do Município de São Paulo.	32
	ARTIGO 2: Propensity score matching approach to test the association of income inequality and mortality in São Paulo, Brazil.	58
	ARTIGO 3: Cause-specific mortality difference between high and low inequality areas of São Paulo, Brazil, by the use of the propensity score matching approach.	79
5	CONCLUSÕES	98
6	CONSIDERAÇÕES FINAIS	101
7	REFERÊNCIAS	105
	ANEXO	
	Anexo 1: Diretrizes para outra forma de apresentação de Tese e Dissertação dos Programas de Pós Graduação da Faculdade de Saúde Pública da USP	113

1 INTRODUÇÃO

A partir do influente estudo de WILKINSON (1996), que concluiu que a situação de saúde dos países desenvolvidos estava mais associada com a distribuição relativa da renda do que com a renda mediana *per capita*, a “teoria da renda relativa” ganhou projeção no meio acadêmico e nas revistas epidemiológicas, fazendo com que um editorial da *British Medical Journal* (BMJ, 1996) a rotulasse de “*the big idea*”.

A teoria estabelece que a renda absoluta tem efeito predominante sobre a saúde da população até que seja atingida uma renda suficiente para a satisfação das necessidades básicas (como acesso a saneamento básico e alimentação de qualidade), momento a partir do qual a desigualdade relativa de renda passa a ter um efeito complementar na determinação das condições de saúde.

Duas revisões da literatura apontam que a maioria dos resultados da área tem sido favoráveis à teoria da renda relativa para áreas maiores que regiões metropolitanas (SUBRAMANIAN e KAWACHI, 2004; WILKINSON e PICKETT, 2006), enquanto outra considera que os resultados tem sido conflitantes (LYNCH et al., 2004). Segundo LYNCH et al. (2004), a maioria dos estudos favoráveis à teoria analisaram ou os 50 estados dos Estados Unidos, ou fizeram uma comparação entre países. Para os autores, as diferenças entre sistemas de governo, as características do mercado de trabalho e a presença ou não de sistema público de saúde ajudam a explicar a diferença de mortalidade entre países. A comparação inter-países seria assim um artefato, resultado da falta de controle de variáveis de confusão. No caso

dos EUA, a relação positiva entre desigualdade de renda e saúde pode ser uma consequência das distribuições raciais de cada estado (DEATON e LUBOTSKY, 2003).

A revisão de SUBRAMANIAN e KAWACHI (2004) de estudos multinível concluiu que estudos realizados fora dos EUA são os maiores responsáveis pela ausência de suporte para a teoria da renda relativa. Para os autores, isso se deve ao fato de esses estudos terem analisado predominantemente regiões menores como bairros ou distritos de cidades pequenas. Mesmo nos EUA, essas áreas tem apresentado resultados conflitantes, possivelmente devido à falta de poder estatístico. Ainda assim, os autores concluíram que a maioria dos estudos analisados pela revisão apresentaram associação estatisticamente significativa entre desigualdade de renda e saúde.

A terceira revisão (WILKINSON e PICKETT, 2006), publicada um ano depois das duas anteriores, identificou 168 estudos na área de desigualdade de renda e saúde e concluiu que 87 (51,8% do total) eram completamente favoráveis à teoria, 44 (26,2%) parcialmente favoráveis e 37 (22,0%) contrários à teoria. Como nas revisões anteriores, a maioria dos estudos contrários à teoria foram os realizados em áreas menores. Do total de 126 estudos que analisaram regiões metropolitanas ou áreas maiores, 105 (82,0%) apresentaram resultados favoráveis à teoria.

1.1 DESIGUALDADE E VIDA EM SOCIEDADE

Enquanto a maioria dos epidemiologistas que estudam desigualdade considera que tanto a diferença absoluta quanto a relativa de renda são relevantes para a saúde (mesmo não existindo ainda consenso sobre o grau de importância), alguns economistas defendem que melhorar a renda absoluta da população (mais especificamente, diminuir a proporção de pobres) deveria ser o único foco das políticas públicas. Essa vertente ganhou projeção com o artigo “*Reducing poverty, not inequality*” de FELDSTEIN (1999) publicado na revista *The Public Interest*. Segundo o autor, se cada leitor da revista ganhasse US\$1.000,00 e ninguém mais fosse afetado, o resultado seria positivo mesmo se a externalidade resultante dessa ação fosse aumentar a desigualdade. Preocupar-se com o aumento da renda dos mais ricos, segundo Feldstein, seria uma característica de “igualitaristas rancorosos” e, como tal, fora do escopo das políticas públicas.

Estudos recentes confirmam a teoria de Feldstein da presença de “igualitaristas rancorosos”, apesar de não usarem essa terminologia (FOWLER et al., 2004; GINTIS et al., 2008). Existe uma tendência de grande parte dos seres humanos em punir não só atitudes egoístas, mas também o aumento da desigualdade, mesmo que a um determinado custo pessoal. Escolhas altruístas em relação à distribuição de renda podem ser facilmente explicadas se influenciarem positivamente a reputação social dos envolvidos (WEDEKIND e MILINSKI, 2000), mas atitudes cooperativas tem sido verificadas mesmo em condições em que a identidade dos participantes é mantida em sigilo (FEHR e GÄCHTER, 2002).

Um exemplo prático vem da teoria dos jogos, mais especificamente do jogo do ultimato, que consiste em distribuir um valor de dinheiro para um dos participantes (por exemplo, 100 reais), oferecendo a possibilidade de esse voluntário dividir parte do dinheiro com outro participante anônimo. Esse segundo participante pode vetar a oferta se essa não for do seu agrado, fazendo com que ambos recebam nada. O esperado pelos modelos econômicos tradicionais é que o segundo participante aceite qualquer oferta do primeiro participante, ou seja, que atue somente em interesse próprio, independentemente de quanto dinheiro o primeiro participante ficar (FEHR e GINTIS, 2007). No entanto, estudos realizados em várias partes do mundo, tanto em países em desenvolvimento como a Indonésia, quanto em países desenvolvidos como os EUA, verificaram que ofertas abaixo de 20% do total são frequentemente rejeitadas pelo segundo participante, que prefere perder dinheiro a participar de uma distribuição extremamente desigual do dinheiro (FEHR e SCHMIDT, 1999).

Segundo KAWACHI (1999), a presença de altruísmo e laços de reciprocidade é um forte indício da situação de saúde de uma região. A partir da publicação do influente estudo de PUTNAM (2000), novas pesquisas apontaram a relação entre desigualdade e capital social, definido como as características da organização social que trazem recursos ao indivíduos e facilitam a ação coletiva (LOCHNER et al., 1999). Uma medida bastante utilizada na literatura para mensuração de capital social é o nível de confiança entre indivíduos, frequentemente analisado pela pergunta: “Você acha que a maioria das pessoas são confiáveis?”. Em uma análise de estados norte-americanos, KAWACHI et al. (1997) encontraram correlação entre alta desigualdade e alta desconfiança entre residentes ($r = 0,73$; $p <$

0,0001). Para ROTHSTEIN e USLANER (2005), capital social e desigualdade de renda atuam em conjunto, fazendo com que as regiões fiquem presas num ciclo vicioso de alta desigualdade, desconfiança e instituições políticas fracas. Segundo o autor, alta desconfiança leva a uma maior oposição a políticas redistributivas. Ao mesmo tempo, a alta desigualdade de renda favorece a segregação e diminui as oportunidades de interações sociais, aumentando a desconfiança entre residentes.

Estudos feitos na última década consistentemente encontraram que pessoas mais ricas se declaram mais felizes que pessoas mais pobres (FIREBAUGH e TACH, 2005). Uma explicação é que a renda absoluta traz mais acesso a bens de consumo e a serviços. Em países mais pobres, um aumento da renda per capita está associado a um aumento do número de pessoas que se declaram felizes ou muito felizes; por outro lado, países desenvolvidos que aumentaram sua renda nas últimas décadas não apresentaram aumento proporcional da felicidade (BOK, 2010). Uma teoria é que as pessoas julgam sua felicidade de forma relativa (EASTERLIN, 1995). Em um país que já possui amplo acesso a saneamento básico e a alimentação adequada, um aumento da renda está normalmente associado a um maior consumo de bens duráveis como eletrônicos e veículos. Esse aumento da renda será imperceptível ao indivíduo se todos ao seu redor consumirem esses mesmos bens.

Um clássico estudo de felicidade realizado por BRICKMAN et al. (1978) verificou que ganhadores da loteria no Estado de Illinois, EUA, não eram mais felizes que um grupo de controle com as mesmas características, tanto na felicidade passada e presente, quanto na felicidade esperada para o futuro. Para os autores, esse resultado se deve à capacidade de seres humanos de se adaptarem a novas situações.

Uma teoria bastante defendida por economistas ortodoxos (neoclássicos) é que o nível de desigualdade de renda pode inclusive ser benéfico para o crescimento econômico (KRUGMAN, 2007). Segundo esse pensamento, a presença de alta desigualdade de renda faria com que as pessoas se esforçassem mais, criando incentivos para maior produtividade e mobilidade social. Em uma sociedade extremamente igualitária, o aumento do esforço marginal traria poucos benefícios financeiros (devido ao fato de que boa parte dessa renda extra seria consumida pelos altos impostos).

Em contraste com essa teoria, estudos apontam que regiões mais igualitárias (países nórdicos como a Finlândia, Suécia, Noruega e Dinamarca) são também as que possuem maior mobilidade intergeracional (BLANDEN et al., 2005). Enquanto na Noruega a probabilidade de um filho cujo pai pertence ao último quintil de renda permanecer nesse quintil é de 24,7%, nos EUA essa probabilidade é de 42,2% (D'ADDIO, 2007). Segundo HILL et al. (2010), a presença de alta desigualdade é fundamental para a manutenção do *status quo*, por criar barreiras sociais e diminuir as oportunidades dos mais pobres de acesso a educação e a empregos de qualidade.

Apesar dos exemplos de bilionários filantropos, a presença de pessoas mais ricas pode também estar associada a menor doação de dinheiro para caridade (JAMES e SHARPE, 2007). Em uma análise recente que utilizou um jogo simples para medir generosidade, pessoas que se consideraram na parte de baixo da pirâmide social doaram mais para indivíduos anônimos do que aqueles que se colocaram no topo da escala de renda (PIFF et al., 2010). Os resultados foram estatisticamente significativos mesmo após o ajuste por idade, religiosidade e origem étnica. Segundo os autores, isso se deve ao fato de os mais pobres terem mais a ganhar com atitudes

pró-sociais que os mais ricos, por se encontrarem em uma situação econômica mais vulnerável.

Ao contrário do esperado pela teoria econômica tradicional, novos estudos mostram que a alta desigualdade pode ser também prejudicial ao crescimento econômico. Em uma análise de áreas rurais da China, a desigualdade de renda local esteve negativamente associado ao crescimento de renda no futuro (BENJAMIN et al., 2006). Mesmo uma pequena distribuição de renda já seria capaz de trazer benefícios à produtividade e ao empreendedorismo, ao permitir ampliar o acesso ao crédito e à formação de pequenas empresas (AGHION et al., 1999).

Países desiguais são também mais vulneráveis a momentos de instabilidade política. A manutenção da convivência diária entre alta pobreza e alta riqueza tem passado tradicionalmente pela exclusão e censura de grupos sociais, tanto dos diretamente afetados (pobres e minorias), quanto de seus representantes (ONGs e mídia) (KAWACHI e KENNEDY, 2002). O controle político fica nas mãos dos indivíduos mais ricos, interessados na manutenção da hierarquia social vigente. Um clássico estudo verificou que, nos Estados Unidos, os 3% mais ricos são responsáveis por 35% de todas as doações de dinheiro para campanhas políticas durante as eleições presidenciais (VERBA et al. 1995).

1.2 DESIGUALDADE E FATORES DE RISCO

Maior pobreza está quase que inevitavelmente associada a maiores índices de violência, mas estudos recentes procuram explicar por que as pessoas cometem atos de violência mesmo quando não há relação direta com a sobrevivência. WILKINSON (2004) ao analisar estudos sociológicos e epidemiológicos concluiu que a violência é muitas vezes apenas uma reação contra uma sensação de inferioridade, ou seja, uma busca individual por maior *status*. Essa necessidade de afirmação social é mais evidente quanto maior a diferença local de renda entre ricos e pobres. Um indivíduo pode sentir-se inferiorizado não apenas por uma pessoa que diretamente o “desrespeitou”, mas por um ambiente social que desfavorece suas oportunidades de encontrar relações afetivas ou admiração social. Em apoio a essa teoria, KAPLAN et al. (1996) encontraram, em um estudo de estados norte-americanos, alta correlação entre percentagem total de renda dos 50% mais pobres e homicídio ($r = 0,74$; $p < 0,0001$), crimes violentos ($r = 0,70$; $p < 0,0001$) e total da população cumprindo pena na cadeia ($r = 0,44$; $p < 0,01$).

Diversos estudos apontam associação entre desigualdade e indicadores de saúde. DIEZ-ROUX et al. (2000), ao analisarem o efeito da desigualdade sobre alguns dos fatores de risco para doenças cardiovasculares, encontraram que os estados norte-americanos mais desiguais apresentaram maior percentagem de sedentarismo e tabagismo entre os homens. Para as mulheres, além desses dois fatores, a associação foi também estatisticamente significativa para hipertensão e alto Índice de Massa Corpórea (IMC).

Um estudo de PICKETT et al. (2005) analisou gravidez na adolescência e verificou que tanto para os EUA, quanto para outros países em conjunto, regiões com maior desigualdade de renda apresentaram maior taxa de natalidade em adolescentes (15 a 19 anos). A mesma associação foi encontrada em um estudo dos distritos norte-americanos para adolescentes entre 15 e 17 anos, mas não para as mulheres entre 18 e 19 anos (GOLD et al., 2001).

A desigualdade pode também afetar o desempenho escolar de crianças. Em um estudo muito citado do Banco Mundial (HOFF e PANDEY, 2004), 321 estudantes indianos de castas superiores e 321 de castas inferiores foram colocados para resolver jogos de labirintos considerados simples. Na primeira rodada do exercício, os resultados foram semelhantes (até ligeiramente melhor para os de baixa casta). Na segunda rodada houve uma pequena alteração: os participantes foram informados em voz alta, para todos ouvirem, qual era a sua casta. Nesse caso, os resultados se inverteram e os estudantes de maior casta conseguiram resolver significativamente mais labirintos. Para os autores do estudo, esse resultado pode ser explicado por mudanças na autoconfiança, nas expectativas intrínsecas de sucesso e na confiança na recompensa pelo esforço.

Para WILKINSON e PICKETT (2009), um paralelo pode ser feito no caso da desigualdade de renda, no qual a rigidez das estruturas sociais faz com que a expectativa de recompensas justas seja pequena. Assim, manter um estilo de vida saudável (como praticar exercícios físicos ou consultar um médico regularmente) ou aumentar a produtividade (trabalhar ou estudar mais) traz menos oportunidades de melhoria de vida nessas regiões segregadas do que em regiões com maior mobilidade social.

Algumas análises experimentais do efeito da desigualdade na saúde foram possíveis pelo estudo do comportamento de macacos. Em uma colônia de macacos, a estrutura social é bastante hierarquizada (DE WAAL, 2009). Essa posição, contudo, não é estática. Mudanças na posição da escala social podem inclusive alterar a presença de arterosclerose. Macacos dominantes que se tornaram subordinados apresentaram um aumento de 500% dos indicadores de arterosclerose em relação aos que não mudaram de posição (SHIVELY e CLARKSON, 1994). Outro estudo com macacos mostrou que, quando disponível, mais cocaína é consumida pelos macacos subordinados do que pelos dominantes (MORGAN et al., 2002).

1.3 DESIGUALDADE E SAÚDE: MECANISMOS DE AÇÃO

LYNCH et al. (2000) identificaram três hipóteses presentes na literatura que procuram explicar a relação entre desigualdade de renda e piores condições de saúde: a neomaterial, a da renda individual e a do ambiente psicossocial.

A hipótese neomaterial defende que as experiências e exposição a fatores de risco são diferentes para cada classe social e isso se reflete diretamente nas condições de saúde. Indivíduos de classes mais baixas são historicamente discriminados no acesso a melhor educação, alimentação, emprego e transporte e são mantidos fora dos processos de decisão política. De acordo com essa hipótese, NAVARRO e SHI (2001), em um estudo de países europeus, concluíram que governos de partidos comprometidos com a diminuição das desigualdades foram os que mais investiram em

políticas públicas de saúde, educação e bem-estar; tal fato não só diminuiu a desigualdade, como também reduziu consideravelmente a mortalidade infantil. Para a hipótese neomaterial, os efeitos adversos da desigualdade de renda poderiam ser anulados se a presença de alta desigualdade não estivesse fortemente associada com baixo acesso aos recursos materiais promotores da saúde.

A hipótese da renda individual utiliza como justificativa a relação curvilínea entre renda e saúde. Como existe consenso que o efeito marginal de um aumento de renda na saúde é decrescente, ou seja, um aumento de R\$ 100,00 na renda de uma pessoa pobre tem maior efeito positivo na saúde da sociedade do que um mesmo aumento de R\$ 100,00 na renda de um rico, a desigualdade seria prejudicial à saúde de uma região por não permitir uma alocação perfeita da renda. Na análise de dados agregados, alta desigualdade levaria a piores condições de saúde devido aos efeitos assimétricos da renda.

A terceira hipótese, chamada de ambiente psicossocial, tem apresentado crescente influência entre pesquisadores da área, principalmente a partir da metade da década de 90. Segundo seus defensores, as percepções de posição relativa na estrutura social levam o indivíduo a desenvolver sentimentos de vergonha, angústia e desconfiança, que possuem consequências neurológicas e endócrinas prejudiciais à saúde. A desigualdade também promove efeitos comportamentais nos indivíduos, diminuindo o capital social (queda na cooperação e participação cívica). Ao contrário das duas hipóteses anteriores, a hipótese do ambiente psicossocial defende que a posição relativa do indivíduo na escala de renda é fundamental e que, mesmo em uma sociedade rica, a alta desigualdade traria efeitos prejudiciais à saúde. Para WILKINSON (2009), a importância da posição relativa na hierarquia social é uma

questão evolutiva. Indivíduos, ou animais, situados na parte de baixo da escala social são barrados de oportunidades reprodutivas e de crescimento social, o que leva à frustração crônica, independentemente de a região como um todo ser considerada rica ou pobre.

A maioria dos pesquisadores considera que as três hipóteses atuam em conjunto, mas discordam em relação ao grau de importância de cada uma. A literatura está apenas começando a entender os mecanismos de ação da desigualdade na saúde e ainda se encontra longe de chegar a um consenso (SUBRAMANIAN e KAWACHI, 2004).

1.4 DESIGUALDADE E SAÚDE NO BRASIL

Em um estudo realizado pelas Nações Unidas em 2005, o Brasil apresentou índice de desigualdade mais elevado que 95% dos 124 países analisados (WATKINS, 2005). Segundo o mesmo estudo, se o cálculo do Índice de Desenvolvimento Humano (IDH) brasileiro incluísse apenas a renda média dos 20% mais pobres, mantendo os outros indicadores usados no IDH inalterados, o país cairia 52 posições no ranking mundial (de 63º para 115º).

Apesar de pouco numerosos, alguns estudos realizados no Brasil procuraram analisar a relação entre a desigualdade relativa de renda e condições de saúde. SZWARCOWALD et al. (1999), utilizando dados do Censo de 1991 e de mortalidade para o triênio 1990/1992, no Município do Rio de Janeiro, encontraram correlações

significativas entre três indicadores de desigualdade de renda e alguns indicadores de saúde (coeficiente de mortalidade geral padronizado por idade, esperança de vida ao nascer e taxas de homicídio e de analfabetismo). A taxa de homicídios foi a variável mais fortemente associada com a desigualdade de renda ($r = 0,60$ a $0,64$, dependendo do indicador utilizado).

Outro estudo, também realizado no Rio de Janeiro (VICENTIN et al., 2002), encontrou alta correlação entre mortalidade por tuberculose e dois indicadores de desigualdade de renda (índice de Robin Hood e razão de renda entre os 10% mais ricos e os 40% mais pobres). A correlação foi também positiva, apesar de mais baixa, para o coeficiente de Gini.

Ao analisar o Município de São Paulo, ANTUNES (2004) encontrou correlação estatisticamente significativa entre o coeficiente de Gini e mortalidade por câncer de boca e faringe, esôfago e estômago.

Os estudos mencionados procuraram correlacionar desigualdade de renda diretamente com indicadores de saúde. Esse tipo de análise é vulnerável à falácia da desigualdade, na qual a alta desigualdade atua como um indicativo de alta pobreza. Assim, regiões desiguais apresentarão piores índices não pela presença da desigualdade *per se*, mas sim pela alta proporção de pobres. Tanto SZWARCWALD et al. (1999) quanto ALMEIDA (2004) reconhecem essa limitação em seus respectivos estudos.

2 OBJETIVOS

2.1 GERAL

O objetivo do estudo é verificar o efeito da presença de alta desigualdade de renda na mortalidade do Município de São Paulo.

2.2 ESPECÍFICOS

Testar a adequação da metodologia *propensity score matching*.

Verificar a possibilidade de divisão da cidade em áreas homogêneas para análises futuras.

Testar a existência de diferença estatisticamente significativa entre a mortalidade geral de distritos mais e menos igualitários.

Identificar as causas básicas de morte responsáveis pela diferença de mortalidade entre distritos mais e menos igualitários.

3 METODOLOGIA

A metodologia e técnicas utilizadas para a elaboração das análises encontram-se detalhadas nos manuscritos (artigos) em seguida.

4 RESULTADOS E DISCUSSÃO (ARTIGOS)

ARTIGO 1

O uso do *propensity score matching* na análise da
distribuição espacial de características demográficas e sociais
do Município de São Paulo.

Chiavegatto Filho, Alexandre Dias Porto ^a

Gotlieb, Sabina Léa Davidson ^a

de Almeida, Samuel Luna ^a

Kawachi, Ichiro ^b

^a Faculdade de Saúde Pública, Universidade de São Paulo.

^b Department of Society, Human Development and Health, Harvard School of
Public Health.

RESUMO

Introdução. O padrão espacial de distribuição de renda do Município de São Paulo, frequentemente generalizado como sendo “radial”, tem sido muito questionado pela literatura recente. São Paulo tem uma complexa distribuição de características sociais e demográficas entre seus distritos, o que dificulta a análise por meio de modelos estatísticos que permitem a inclusão somente de algumas variáveis de cada vez, como as regressões lineares. O presente estudo objetiva identificar os distritos do município que possam ser considerados como “comparáveis” por meio do uso de uma metodologia estatística conhecida como *propensity score matching*.

Metodologia. Os 96 distritos do Município de São Paulos foram analisados separadamente. Foram incluídas 16 variáveis no modelo, sendo o índice de Gini a variável que permitiu a separação de distritos entre expostos (alta desigualdade) ou não-expostos (baixa desigualdade). Do total de distritos, 27 foram considerados comparáveis com algum outro, isto é, possuíram valores de propensity score com uma distância menor de 0,1 de outro com tipo de exposição diferente. **Resultados.**

Das 16 variáveis incluídas, 9 apresentaram diferenças estatisticamente significativas entre os distritos incluídos e excluídos, fato esperado pela metodologia. Dos 17 pares de distritos formados, apenas três foram compostos por distritos de uma mesma região administrativa e apenas um por distritos que faziam fronteira entre si.

Conclusão. A análise da diferença no padrão de distribuição das variáveis, permitida pelo uso do *propensity score matching*, indica a dificuldade de dividir a cidade segundo regiões. Para entender São Paulo é preciso considerar suas particularidades e suas complexas distribuições espaciais.

Propensity score matching approach to analyze the spatial distribution of some social and demographic characteristics of the Municipality of São Paulo.

ABSTRACT

Introduction. The spatial pattern of income distribution in the Municipality of São Paulo, considered to be of a “radial” type, has been challenged by recent studies. São Paulo has a complex distribution of social and demographic characteristics between its *distritos* (a proxy for neighborhood). This demands an in-depth analysis that takes into consideration a multitude of variables, in order to control for local heterogeneity. This study aims to identify the *distritos* of São Paulo that can be defined as “comparable” to another one, by using a statistical methodology known as propensity score matching. **Methodology.** The 96 *distritos* of the Municipality of São Paulo were analyzed separately. 16 variables were included in the statistical model, and the Gini coefficient was the variable used to define “exposure” (high inequality) and “non-exposure” (low inequality). Of all *distritos*, 27 were considered “comparable”, i.e. presented a propensity score that had a maximum distance of 0.1 from another *distrito* with different exposure. **Results.** Of the 16 variables inserted in the model, nine presented a statistically significant difference between included and excluded *distritos*, which is expected by this methodology. Of the 17 pairs of *distritos* considered to be comparable, only three were composed of *distritos* situated in the same administrative region, and only one was composed of bordering *distritos*. **Conclusion.** The complex spatial distribution of the propensity

score in the Municipality of São Paulo indicates that it is very difficult to divide the city according to its geographical regions. In order to understand how the *distritos* of São Paulo affect the health of its residents, it is important to take into consideration its many particularities and how they are spatially distributed.

INTRODUÇÃO

Estudos sobre pobreza em São Paulo não podem apenas se restringir a uma descrição espacial da distribuição de renda. A presença de altas desigualdades sociais e demográficas em pequenas áreas faz com que a análise da situação de São Paulo necessite de uma abordagem mais complexa que a utilizada para descrever outras cidades.

Comparativamente, a Região Metropolitana de São Paulo abriga um número superior de famílias vivendo com menos de um quarto de salário mínimo *per capita* do que a população total de Porto Alegre.¹ E, apesar de contar com 47% da população do Estado de São Paulo, abriga 55% do total de pobres do Estado, contrariando a tendência nacional de maior concentração da pobreza em regiões rurais.¹

Uma das mais estabelecidas características de São Paulo é o seu alto nível de segregação, mesmo quando comparada a outras cidades brasileiras como o Rio de Janeiro.² Embora a população de baixa renda tenha diminuído proporcionalmente desde a década de 90 do último século, a sua concentração aumentou nas áreas mais pobres da cidade, o que explica o crescimento da segregação.³ O meio urbano paulistano é caracterizado por casas com muros altos, condomínios fechados e policiamento de fronteiras, o que vem reforçar as desigualdades. Para Caldeira⁴ (2000), São Paulo é a cidade que melhor representa a modernidade brasileira com todos os seus paradoxos: indústrias, favelas, metrô sofisticados, altas taxas de mortalidade infantil e arranha-céus.

O uso do conceito de linha de pobreza para a análise da exclusão social está longe de ser um consenso na literatura. A definição de “renda mínima necessária para a sobrevivência de uma família” não leva em conta os maiores custos de vida existentes nas regiões urbanas (moradia e transporte, principalmente), nem as múltiplas dimensões da pobreza, como a importância da educação, saúde e saneamento básico para a perpetuação das desigualdades.⁵ Indivíduos pobres que residem em regiões urbanas também possuem menor acesso às rendas não-monetárias (como a produção agrícola para auto-consumo), tornando-os mais vulneráveis às oscilações de renda.¹

Para o Município de São Paulo, a grande questão não é apenas encontrar qual o índice mais adequado para medir pobreza, mas conseguir analisar a distribuição de renda levando em consideração toda a sua heterogeneidade.⁶ Em uma cidade com poucos recursos, a pobreza não é distribuída aleatoriamente, mas está mais concentrada em regiões carentes e com menor acesso à infraestrutura urbana.⁷ O entorno habitacional tem efeitos diretos na saúde da população. Espaços mais pobres tem menor acesso a água encanada, coleta de lixo, tratamento de esgoto e educação de qualidade, características que tem uma influência bem estabelecida na deterioração da saúde dos residentes.⁸ O fato de simplesmente morar em uma região pobre de São Paulo também está associado a menores chances de conclusão do ensino médio, independentemente da renda familiar *per capita*.⁹

O padrão histórico da distribuição de renda na cidade de São Paulo é o radial, com os mais ricos ocupando as regiões centrais e um aumento da pobreza em direção à periferia,¹⁰ mas estudos recentes apontam a existência de uma forte heterogeneidade mesmo dentro dos diferentes estratos de renda.^{1,6} Marques e Torres¹

(2004) dividiram as áreas mais ricas em três tipos, as de classe-média em quatro e as pobres em três e verificaram “a existência de espaços periféricos bastante complexos, nos levando a destacar a existência de *periferias*, ao invés de periferia”.

Ao contrário da periferia de São Paulo, a região central abriga favelas de pequeno porte, localizadas em beiras de córregos e pequenas faixas remanescentes de obras públicas.¹¹ Mesmo assim, morar perto do centro significa estar mais próximo do mercado de trabalho e ter maior acesso a informações sobre empregos e cursos.¹² Na região de classe-média-alta do Morumbi está situada Paraisópolis, uma favela com mais de 80.000 habitantes cuja proximidade física com o entorno rico significa maior presença de ONGs e mercado de trabalho.¹³ O estigma da presença espacial de uma favela é importante para a saúde, não só em relação à sensação de inferioridade apresentada pelos seus moradores, mas, também, pelos efeitos adversos (possivelmente estressores) em seu entorno mais rico. Em uma pesquisa qualitativa feita por Caldeira⁴ (2000), foi verificado que muitos paulistanos de classe-média consideram a favela como sinônimo da presença de criminalidade e de falta de caráter, principalmente se situada próxima de uma região rica.

As variáveis que definem exclusão social nos distritos do Município de São Paulo variam conforme tempo de ocupação, presença do poder público e características do entorno físico. A metodologia mais utilizada na área de saúde pública, a de regressões multivariadas, corre o risco de não detectar evidências de significância estatística devido ao grande número de variáveis que devem ser inseridas no modelo. Uma sugestão apresentada, e com crescente uso na área de epidemiologia social, é o *propensity score matching*, uma análise estatística aplicada inicialmente em estudos cardiológicos, mas que começou recentemente a ganhar

exposição por toda a área epidemiológica.¹⁴

METODOLOGIA

Propensity Score

O conceito de *propensity score* (PS) foi apresentado pela primeira vez na literatura por Rosenbaum e Rubin em 1983,¹⁵ sendo definido como a probabilidade condicional de exposição levando em conta um grupo de variáveis.

O PS tem como objetivo identificar as regiões mais semelhantes entre si, levando-se em consideração um conjunto de variáveis e um fator de exposição (no caso, a desigualdade de renda). O PS é calculado por meio de uma regressão logística cujos valores variam no intervalo entre 0 e 1. Assim, como o resultado final é a probabilidade de exposição, a não obtenção de significância estatística por excesso de variáveis não se torna um problema para esse modelo.

As variáveis sociais e demográficas incluídas no modelo e a identificação das regiões como expostas ou não-expostas permitirão o cálculo da probabilidade de exposição. Por resumir um conjunto de variáveis em uma função escalar, o PS possibilita identificar se os dois grupos (expostos e não-expostos) se intercalam o suficiente para permitir uma comparação com menor variabilidade entre eles. Um mesmo valor de PS significa igual probabilidade de exposição, de acordo com as variáveis selecionadas. Entre indivíduos ou locais com PS semelhantes, alguns serão expostos e outros não expostos, o que permitirá que seja feita a comparação. Essa análise pode ser definida como uma “randomização depois da exposição”.¹⁶ Uma técnica bastante utilizada para testar se os *propensity scores* encontrados pelo modelo tem alto poder preditivo é a curva ROC, cuja área representa a probabilidade de as

regiões apresentarem um valor de *propensity score* condizente com a sua exposição (nesse caso, de regiões expostas apresentarem *propensity scores* mais altos e vice-versa).¹⁷

Após o cálculo do PS, a sequência mais comum é o *matching*, onde cada local ou indivíduo é pareado com outro de mesmo valor de PS, mas com diferente exposição. O processo de *matching* inicia-se com a identificação do exposto com menor valor de PS, que será pareado com um não-exposto desde que se situem dentro de uma diferença de valor máxima (*caliper width*). O *caliper width* mais utilizado na literatura é igual a 0,01. Assim, por exemplo, um local ou indivíduo exposto que tenha PS de 0,670 poderá ser agrupado com não-expostos que tenham valor de PS entre 0,660 e 0,680. Na presente análise foi utilizado o *propensity score matching with replacement* (com reposição), o que significa que cada indivíduo ou local exposto pode ser pareado com mais do que um não-exposto, desde que dentro do *caliper width*. Indivíduos ou locais que não forem agrupados são excluídos das análises subsequentes.

São Paulo

O Município de São Paulo, atualmente, é composto por 96 distritos administrativos, agrupados em 31 subprefeituras. Os distritos representam a menor área para a qual dados de saúde estão disponíveis, por meio do Sistema TabNET do Programa de Aprimoramento das Informações de Mortalidade (PRO-AIM).¹⁸

Os dados de desigualdade de renda foram provenientes dos setores censitários do Município de São Paulo (13.278 no total).¹⁹ Utilizou-se para o cálculo a renda

per capita de todos os residentes (incluindo domicílios cujo responsável apresentou renda igual a zero), mantendo os setores censitários como unidade de análise. O índice de Gini foi calculado utilizando-se o aplicativo *ineqdeco* do Stata 10®.

O índice de Gini é calculado a partir da curva de Lorentz, por meio da área formada pela distância entre a distribuição real e a perfeitamente igualitária das rendas.²⁰ Os valores do índice de Gini dos distritos variaram de 0,12 (Jaguará) a 0,55 (Vila Andrade). No presente estudo, um distrito foi considerado exposto se apresentasse índice de Gini maior ou igual a 0,25 (alta desigualdade) e não-exposto quando o indicador fosse menor do que 0,25 (baixa desigualdade). Como não existe consenso na literatura sobre qual o valor de Gini que pode considerado como alto, optou-se por definir o limite como 0,25, por estar próximo ao valor mediano dos índices Gini dos distritos paulistanos.

As variáveis escolhidas para o *propensity score matching* foram referentes às características absolutas dos distritos. Elas foram selecionadas pelos autores por abordarem algumas importantes características sociais, demográficas e educacionais que podem ter efeito na saúde além da desigualdade de renda.

As 16 variáveis selecionadas para o cálculo do PS para cada distrito foram: média de anos de estudo dos responsáveis pelo domicílio, densidade residencial, pobreza (proporção de pessoas vivendo com menos de um salário mínimo), renda mediana, proporção de pessoas vivendo em favelas, proporção de residências com abastecimento de água, proporção de residências com coleta de lixo, proporção de residências sem banheiro, proporção de responsáveis por domicílio com menos de 21 anos, proporção de responsáveis não-alfabetizados, proporção de crianças de 8 a 12 anos de idade não-alfabetizadas, número de professores por aluno da 5ª à 8ª série,

incidência de aids, proporção de menores de 1 ano de idade, proporção de idosos ($\geq$ 65 anos de idade) e proporção de mulheres. A maioria dos dados foi retirada do Censo de 2000.¹⁹ A quantidade de professores para cada aluno foi obtida no Censo Educacional de 2001 e a incidência de aids no Boletim Epidemiológico Municipal.^{21,22} Os coeficientes gerais de mortalidade foram padronizados por idade e são referentes ao período de 1998 a 2002 (calculados para valores anuais utilizando-se os dados de população do censo de 2000).

RESULTADOS

Os valores de PS variaram de 0 (distrito de Marsilac) a 1 (distrito de Vila Andrade), com valor mediano igual a 0,63. A área abaixo da curva ROC, ou *c-statistic*, foi de 0,907, o que indica alto poder preditivo (alta sensibilidade); ou seja, distritos expostos apresentaram valores de PS consistentemente mais elevados, o que é o esperado pela metodologia.²³

Do total de 96 distritos, somente 27 deles apresentaram um PS que satisfazia ao limite do *caliper width* de 0.01 de outro distrito com exposição diferente e foram incluídos na análise. As diferenças observadas entre os distritos incluídos e os excluídos estão apresentadas na Tabela 1. O esperado nesse caso é que haja diferenças significativas entre eles, já que a metodologia foi aplicada para excluir os distritos que não apresentavam pelo menos um outro distrito com características semelhantes, ou seja, possíveis *outliers*. Das 16 variáveis selecionadas, nove delas apresentaram diferenças estatisticamente significativas entre os incluídos e os excluídos ($p \leq 0,05$).

Tabela 1 - Distribuição dos valores (nº e %) e suas respectivas diferenças para as variáveis selecionadas segundo distritos (incluídos e excluídos)*, Município de São Paulo, 2000.

Variáveis	Incluídos	Excluídos	Diferença	IC 95%	p-Test
Favelas	4.71	7.3	-2.59	-6,41 - 1,22	
Pobreza	7.1	21.05	-13.94	-22.81 - -5.08	<0.01
Renda mediana per capita	482.48	425.77	56.7	-129.29 - 242.70	
Escolaridade	8.83	7.84	0.99	0.01 - 1.97	
Densidade	3.24	3.44	-0.2	-0.37 - -0.02	<0.05
Abastecimento de água	1	0.97	0.03	-0.02 - 0.08	
Coleta de lixo	1	0.99	0.01	-0.01 - 0.02	
Sem banheiro	0.002	0.003	-0.001	-0.002 - -0.0002	<0.01
Responsáveis < 21 anos	0.008	0.01	0.002	-0.004 - -0.001	<0.01
Responsáveis não-alfabetizados	0.04	0.06	0.02	-0.04 - -0.01	<0.01
Não-alfabetizados (8-12 anos)	0.03	0.04	0.02	-0.03 - -0.01	<0.01
Professores por aluno	0.045	0.045	0	-0.002 - 0.003	
Incidência de aids	53.44	33.52	19.93	6.02 - 33.84	<0.01
Menores de 1 ano de idade	0.013	0.016	0.003	-0.01 - -0.001	<0.01
Idosos	0.09	0.07	0.02	-0.006 - 0.04	<0.01
Mulheres	0.53	0.53	0	-0.003 - 0.012	

* Classificados por meio da metodologia *propensity score matching*.

A análise da relação entre coeficiente de mortalidade anual ajustado por idade e índice de Gini nos 96 distritos mostra que as regiões da periferia de São Paulo possuem, em geral, maior coeficiente de mortalidade, com a exceção dos distritos do centro (Mapa 1). A zona Leste e o extremo Sul da cidade apresentam índice de Gini mais baixos (ou seja, são mais igualitários), mas, por outro lado, são mais pobres que os demais; todos os 14 distritos com menor renda mediana têm índice de Gini abaixo do valor mediano do Município (0,25).

Mapa 1 – Distribuição dos coeficientes gerais de mortalidade anuais ajustados por idade segundo índice de Gini, Município de São Paulo, 1998-2002.

Regiões mais centrais do Município de São Paulo apresentam valores de *propensity score* mais altos, em consequência da distribuição das variáveis e da maior presença de desigualdade de renda (Mapa 2). É possível verificar, porém, a presença de alguns distritos com PS altos e mais igualitários nas regiões centrais do município, o que aumenta a possibilidade de pareamento.

Mapa 2 – Distribuição dos valores de propensity score segundo índice de Gini, Município de São Paulo, 2000.

O Mapa 3 apresenta os resultados espaciais do pareamento (*matching*) feito pela metodologia do *propensity score*. Do total de 96 distritos que foram incluídos na análise após o pareamento, 27 foram considerados “comparáveis” segundo as 16 variáveis selecionadas, formando um total de 17 pares. Dois distritos foram pareados quatro vezes (Vila Mariana e Saúde) e o distrito de Santana foi pareado duas vezes. Levando em consideração os nove zoneamentos administrativos oficiais do Município de São Paulo (noroeste, nordeste, oeste, centro-sul, centro, sudeste, sul, leste 1, leste 2) apenas três dos 17 pares pertencem a uma mesma zona administrativa (Santana-Casa Verde, Santo Amaro-Vila Mariana e Campo Limpo-Cidade Ademar). E do total de pares, apenas um é de fronteira: Santana-Casa Verde.

Mapa 3 – Identificação dos distritos incluídos na análise de propensity score

matching, Município de São Paulo, 2000.

DISCUSSÃO

O presente estudo permitiu analisar a complexidade das diferenças entre os distritos do Município de São Paulo, utilizando como base a distribuição de renda representada pelo índice de Gini. Dos 96 distritos da cidade, apenas 27 foram considerados comparáveis com algum outro, levando-se em consideração as 16 variáveis selecionadas. Em apenas um dos 17 pares foi detectada a presença de distritos que fazem fronteira entre si. A dificuldade de apontar um padrão de distribuição espacial de características demográficas e residenciais indica a necessidade de uma abordagem nova para analisar a saúde dos paulistanos.

Como bem identificou Marques e Torres¹ (2004) em São Paulo, o padrão histórico mencionado na literatura de distribuição dos grupos sociais em uma estrutura radial concêntrica “...é, na melhor das hipóteses, uma aproximação genérica”. Antes do uso do *propensity score* para controle das diferenças absolutas entre distritos, não foi encontrada associação estatisticamente significativa entre os coeficientes de mortalidade ajustados por idade e pobreza (medida pela proporção de indivíduos com renda abaixo de um salário mínimo), proporção de favelas ou índice de Gini.

Comparar padrões de mortalidade dos distritos de São Paulo baseando-se apenas em uma das inúmeras possíveis variáveis com influência na saúde é um erro que prejudica a validade dos resultados. Não se pode falar em comparar regiões pobres com regiões ricas sem se considerar a complexidade da distribuição social da cidade. Alguns distritos de São Paulo tem alta proporção de pobres e a renda

mediana é alta. Um exemplo dessa situação é o Morumbi que tem a terceira renda mediana mais alta entre os 96 distritos, mas, ao mesmo tempo, 10% dos residentes vivem com menos de um salário mínimo por mês. Por outro lado, existem distritos extremamente igualitários, mas que tem alta proporção de favelas. O distrito de Pedreira apresenta valor do índice de Gini abaixo do valor mediano do município apesar de ter a segunda maior proporção de residentes morando em favelas (38%).

A presença de diferença estatisticamente significativa em 9 das 16 variáveis verificada para distritos incluídos e excluídos da análise é um forte indício da importância da utilização do *propensity score matching* para a identificação e seleção dos distritos comparáveis.

A metodologia estatística do *propensity score* permitiu também a inclusão de um elevado número de variáveis, analisadas segundo um fator de exposição (no caso, distribuição de renda medida pelo índice de Gini). O uso dessa metodologia tem sido bastante incentivado para trabalhos epidemiológicos que lidam com indivíduos ou locais com alta heterogeneidade.¹⁴

A metodologia, porém, tem como limitação o fato de não conseguir controlar as variáveis não incluídas no modelo, ao contrário do que acontece com estudos randomizados tradicionais.²⁴ Além disso, a aplicação da metodologia em cidades menores pode ser mais difícil, devido à necessidade de uma amostra inicial grande (no presente estudo, apenas 27 dos 96 distritos foram incluídos na análise final).

São Paulo continua a ser a cidade dos contrastes. Analisá-la por meio de uma comparação direta entre distritos ou segundo zonas administrativas é assumir uma uniformidade de características, além da variável de interesse, que não existe. Para

entender São Paulo é preciso considerar suas particularidades e as complexas distribuições espaciais de seus distritos.

BIBLIOGRAFIA

- 1 - Marques EC, Torres HG. Pobreza e distribuição espacial de grupos sociais na metrópole de São Paulo. São Paulo; 2004. [acesso em 11 ago 2010]. Disponível em: <http://www.centrodametropole.org.br/v1/pdf/Adenauer%20ed%20har.pdf>
- 2 - Scalón C, Oliveira MA. Desigualdades de Classe e Desigualdades Espaciais. In: XIII Congresso da Sociedade Brasileira de Sociologia; 2007; Recife, BR.
- 3 - Torres H. Medindo a Segregação. In: Marques E, Torres H. São Paulo: segregação, pobreza e desigualdades sociais. São Paulo: Senac; 2004. p. 81-100.
- 4 – Caldeira TPR. Cidade de muros: crime, segregação e cidadania em São Paulo. São Paulo: EDUSP; 2000.
- 5 – Mingione E. Urban poverty and the underclass: a reader. Cambridge: Blackwell; 1996. Urban poverty in the advanced industrial world: concepts, analysis and debates. p. 1-40.
- 6 – Torres HG, Marques E, Ferreira MP, Bitar S. Pobreza e espaço: padrões de segregação em São Paulo. Estud Av. 2003;17(47):97-128.
- 7 – Marques E. Elementos conceituais da segregação, da pobreza urbana e da ação do Estado. In: Marques E, Torres H. São Paulo: segregação, pobreza e desigualdades sociais. São Paulo: Senac; 2004. p. 19-56.
- 8 – Wilkinson R, Marmot M. Social determinants of health: the solid facts. Copenhagen; 2003. (World Health Organization report). Disponível em: http://www.euro.who.int/__data/assets/pdf_file/0005/98438/e81384.pdf

- 9 – Torres H, Ferreira MP, Gomes S. Educação e segregação social: explorando o efeito das relações de vizinhança. In: Marques E, Torres H. São Paulo: segregação, pobreza e desigualdades sociais. São Paulo: Senac; 2004. p. 123-142.
- 10 – Singer P. Desenvolvimento econômico e evolução urbana. São Paulo: Companhia Editora Nacional; 1977.
- 11 – Saraiva C, Marques E. A dinâmica social das favelas da região metropolitana de São Paulo. In: Marques E, Torres H. São Paulo: segregação, pobreza e desigualdades sociais. São Paulo: Senac; 2004. p. 143-168.
- 12 – Gomes S, Amitrano C. Local de moradia na metrópole e vulnerabilidade ao (emprego e) desemprego. In: Marques E, Torres H. São Paulo: segregação, pobreza e desigualdades sociais. São Paulo: Senac; 2004. p. 169-194.
- 13 – Almeida R, D’Andrea T. Pobreza e redes sociais em uma favela paulistana. *Novos Estudos*. 2004;68:94-106.
- 14 – Oakes JM, Johnson PJ. Propensity score matching for social epidemiology. In: Oakes JM, Kaufman JS. *Methods in social epidemiology*. San Francisco: Jossey-Bass; 2006.
- 15 – Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70: 41-55.
- 16 – Yue LQ. Statistical and regulatory issues with the application of propensity score analysis to nonrandomized medical device clinical studies. *J Biopharm Stat*. 2007;17(1):1-13.
- 17 – Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29-36.

- 18 – Programa de Aprimoramento das Informações de Mortalidade no Município de São Paulo (PRO-AIM) [base de dados na internet]. São Paulo: Prefeitura Municipal de São Paulo. Disponível em:
ww2.prefeitura.sp.gov.br/cgi/deftohtm.exe?secretarias/saude/TABNET/SIM/obito.def.
- 19 – Instituto Brasileiro de Geografia e Estatística (IBGE). Censo Demográfico 2000 - Resultados do Universo: Agregados de setores censitários - São Paulo, Região Metropolitana [CD-ROM]. Brasília: IBGE; 2003.
- 20 - Sen AK. On economic inequality. Oxford: Clarendon Press; 1973.
- 21 – Fundação Seade, CEPID-FAPESP, Centro de Estudos da Metrópole/Cebrap. Censo escolar 2001 [base de dados na internet]. Disponível em:
<http://www.centrodametropole.org.br/cd/escolas/ESC2001.rar>.
- 22 – Secretaria Municipal da Saúde de São Paulo (SMSP). Boletim epidemiológico de aids do Município de São Paulo, ano VII, n. 7, dezembro de 2003: 28-31.
- 23 – Stürmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. *J Clin Epidemiol*. 2006;59(5):437-447.
- 24 – Luellen JK, Shadish WR, Clark MH. Propensity scores: an introduction and experimental test. *Eval Rev*. 2005;29(6):530-558.

ARTIGO 2

Propensity score matching approach to test the association of
income inequality and mortality in São Paulo, Brazil.

[in press at the Journal of Epidemiology and Community Health]

Alexandre Dias Porto Chiavegatto Filho

Rua Teodoro Sampaio, 363, apto 614

05405-000 Pinheiros

São Paulo – SP

Brazil

alexdiasporto@usp.br

+55 (11) 30818452

Chiavegatto Filho, Alexandre Dias Porto ^{a, b}

Kawachi, Ichiro ^b

Gotlieb, Sabina Léa Davidson ^a

^a Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, Brazil.

^b Department of Society, Human Development and Health, Harvard School of Public Health, Harvard University, Boston, Massachusetts.

MeSH: Propensity Score, Mortality, Socioeconomic Factors.

Word count: 1931.

ABSTRACT

Background. Support for the adverse effect of high income inequality on population health has come from studies that focus on larger areas, such as the US states, while studies at smaller geographical units (e.g. neighborhoods) have found mixed results.

Methods. We used propensity score matching to examine the relationship between income inequality and mortality rates across 96 neighborhoods (*distritos*) of the municipality of São Paulo, Brazil.

Results. Prior to matching, higher income inequality *distritos* ($\text{Gini} \geq 0.25$) had slightly lower overall mortality rates, 2.23 per 10.000 (95% CI: -23.92 – 19.46) compared to lower income inequality areas ($\text{Gini} < 0.25$). After propensity score matching, higher inequality was associated with a statistically significant higher mortality rate (41.58 per 10.000, 95% CI: 8.85 – 73.3).

Conclusion. In São Paulo, the more egalitarian communities are among some of the poorest, with the worst health profiles. Propensity score matching was used to avoid inappropriate comparisons between the health status of unequal (but wealthy) neighborhoods versus equal (but poor) neighborhoods. Our methods suggest that with proper accounting of heterogeneity between areas, income inequality is associated with worse population health in São Paulo.

What is already known:

- Robust support for the income inequality hypothesis has come from studies conducted across larger areas such as the US states.
- Evidence is less consistent at smaller units of spatial scale, such as neighborhoods within a city.

What this paper adds:

- We improved causal inference by using the propensity score matching approach, i.e. comparing the health status of neighborhoods that are potentially exchangeable with respect to assignment to the exposure of interest (income inequality).
- When the 96 neighborhoods of São Paulo, Brazil, were propensity-matched based on 16 variables, higher inequality was associated with increased overall mortality rates.

INTRODUCTION

Income inequality has been linked to population health outcomes in studies that span a number of different social contexts. A recent meta-analysis identified 27 multi-level studies published on the topic, representing about 60 million subjects.[1] According to the meta-analysis, each 0.05 unit increase in the Gini coefficient of income inequality was associated with a 7.8% excess mortality risk (95% CI: 5.9 – 9.8%). However, heterogeneity was noted across the different levels of spatial aggregation. The relationship between income inequality and worse health appears to be more robust at broader levels of aggregation (e.g. regions and states), whereas the findings are mixed for small areas (e.g. neighborhoods within metropolitan areas). The weaker findings for small area units are inconsistent with the theory of relative deprivation,[2] which posits that any damaging psychosocial effects of income inequality should be more strongly felt as the reference groups become more proximal to the individual, as the relevant social comparisons damaging to health are usually made between neighbors, acquaintances and family members.

A possible explanation for the inconsistent associations found at smaller geographic units is that neighborhoods are lacking in social heterogeneity,[3] i.e. neighborhoods in unequal societies end up being segregated into areas where the households are either uniformly poor or uniformly wealthy. If a neighborhood is at the same time very poor and has a very equal distribution of income, it will still

present a high mortality rate due to the fact that its residents lack basic health-related necessities such as clean water and adequate nutrition.

Hou and Myles[4] have further argued that high levels of inequality in the neighborhood may have a *beneficial* effect on poor residents -- which is the opposite of what is expected from the relative income theory -- for the reason that living in close proximity to affluent residents is often associated with better local public services. These spillover effects will be beneficial to the health of poorer citizens who share the same spaces.

As a third possibility, area homogeneity may also hide the true determinant of the effects of income inequality on health, which are the social comparisons across each person's reference group that may be situated outside their own neighborhood or metropolitan area.[5]

In this study, we provide a test of the income inequality hypothesis in the municipality of São Paulo, Brazil, by applying a statistical methodology – propensity score matching (PSM) – to identify the causal effect of income inequality on mortality rates among residents.

METHODOLOGY

Propensity Score Matching

The propensity score, defined as the probability of assignment to a particular treatment or exposure given a set of observed covariates, was first introduced by Rosenbaum and Rubin in 1983.[6] The idea behind the propensity approach is to identify neighborhoods that are as alike as possible to each other with respect to the probability of receiving the “treatment”, i.e. of having a high level of income inequality. These potentially “exchangeable” neighborhoods are matched on their propensity scores, so that differences in their health outcomes can be compared. In other words, each high inequality neighborhood in São Paulo is matched with another neighborhood that resembles its counterfactual had it not been exposed.

The propensity score is calculated by logistic regression with the dichotomous outcome being the variable treatment or exposure (1=High inequality, 0=Low inequality), using a set of observed covariates as the predictor variables. As it is concerned only with the probability of treatment or exposure, over-parameterizing the model is not an issue.

The matching procedure then pairs exposed and unexposed neighborhoods with similar propensity scores according to a pre-defined maximum range (the caliper width). If a neighborhood is not within a caliper’s width of any other with an

opposite exposure situation, it is dropped. In this paper, we use PSM with replacement, which allows exposed neighborhoods to be used more than once.

Study population

The municipality of São Paulo mirrors Brazil's social, racial and economic disparities.[7] Despite its residents living under the same tax rules (unified for the metropolitan area), basic costs of living and overall distance from food-producing areas, it exhibits a Gini coefficient of 0.51, higher than unequal countries such as the USA and China.[8]

Its 10.3 million residents are divided between 31 *subprefeituras*, which are further divided in 96 *distritos* (average population of 107,697 residents), a proxy for neighborhood, and the smallest area for which mortality data is available.

We used the following 16 *distrito*-level covariates to calculate the propensity score for each neighborhood: prevalence of *favelas* (slums), poverty rate, median income, education of head of household (measured by number of years), household density, proportion of residences with piped water, proportion of residences with garbage collection, proportion of residents without toilet, 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 incidence, proportion of infants (under 1 year old), proportion of elderly (over 64 years old) and proportion of women. Most of the data comes from the 2000 Demographic Census.[9] The 2001 Educational Census provided the data for the

proportion of teachers,[10] and the AIDS incidence is from the municipal epidemiological bulletin.[11]

The outcome variable was the age-adjusted mortality rate. Total mortality is available online from the municipal government.[12] In order to minimize random annual variations within *distritos*, we calculated the mortality rate from 1998 to 2002, the five years centered around 2000 (year of the most recent census).

RESULTS

All the 96 *distritos* of the municipality of São Paulo were used in the analysis. Exposure was defined by the Gini coefficient. *Distritos* with Gini lower than 0.25 were considered unexposed (low inequality) while those with 0.25 or higher were considered exposed (higher inequality). As there is no clear cut-off for high and low inequality, we selected 0.25 as it was close to the median Gini of the *distritos*. For the purpose of matching, we used a caliper of 0.01.

Figure 1a presents the distribution of the propensity scores according to exposure status. The area under the ROC curve, or c-statistic, was 0.907, indicating very high predictive power, i.e. exposed *distritos* have a consistently higher propensity score.[13] Figure 1b shows that the overlap between propensity scores occurred more frequently on higher values, which is confirmed by the fact that most matched *distritos* had a propensity score of over 0.5.

Figure 1: Distribution of the propensity scores (exposed = $Gini \geq 0.25$; unexposed = $Gini < 0.25$).

Of the total 96 *distritos* of São Paulo, 27 had a propensity score situated within caliper width of 0.01 of another *distrito* with opposite exposure, and were consequently included in the analysis. 24 of those were matched only once, one was matched twice, and two were matched four times.

A propensity score analysis is considered to be valid when it reduces covariate bias between exposed and unexposed subjects.[14] Table 1 presents the mean values of the covariates before matching. Of the 16 covariates, five had a statistically significant t-test ($p < 0.05$) suggesting different means between exposed and unexposed.

Table 1: Covariate imbalance before propensity score matching.

Covariates	Unexposed (mean)	Exposed (mean)	Standardised difference	t test	p Value
Favelas	0.05	0.08	12.25	−2.06	0.04
Poverty	0.22	0.13	−22.40	2.41	0.02
Median Income	390.15	487.22	478.56	−1.16	0.25
Education of head of household	7.54	8.62	73.44	−2.45	0.02
Household density	3.48	3.30	−28.53	2.24	0.03
Piped water	0.95	1.00	14.17	−1.83	0.07
Garbage collection	0.98	0.99	6.53	−1.71	0.09
Residences without toilet	0.002	0.002	0.00	0.00	1.00
Head of household under 21 years	0.01	0.01	−0.55	0.42	0.68
Illiteracy rate of heads of household	0.06	0.05	−5.80	1.56	0.12
Illiteracy rate of 8–12 years old	0.04	0.04	−0.89	0.30	0.77
Teachers per students (5th to 8th grade)	0.04	0.05	3.74	−2.34	0.02
HIV/AIDS incidence (100 000)	33.10	44.43	209.34	−1.75	0.08
Proportion of infants (<1 year old)	0.02	0.01	−2.38	1.72	0.09
Proportion of elderly (≥65 years old)	0.07	0.08	5.40	−1.32	0.19
Proportion of women	0.52	0.53	4.37	−1.67	0.10

After propensity score matching, none of the covariates' means remained statistically different ($p > 0.05$). Bias reduction was effective for 12 covariates (Table 2). Of those with increased post-matching bias, three had very close mean averages before matching and remained that way, even if less so. It has been suggested that standardized difference values after matching should stay under 10 per cent.[15] In our analysis, 12 lie within that range.

Table 2: Covariate imbalance after propensity score matching.

Covariates	Unexposed (mean)	Exposed (mean)	Standardised difference	t test	p Value	Bias reduction
Favelas	0.03	0.04	3.32	−0.51	0.61	72.86
Poverty	0.08	0.07	−3.01	−0.55	0.59	86.56
Median income	459.64	495.91	198.06	1.53	0.15	58.61
Education of head of household	8.59	8.96	26.56	1.43	0.17	63.83
Household density	3.34	3.19	−24.37	−0.23	0.82	14.57
Piped water	1.00	1.00	0.01	0.87	0.39	99.93
Garbage collection	0.99	1.00	2.59	−0.04	0.97	60.31
Residences without toilet	0.002	0.002	0.30	−0.98	0.34	−40780.10
Head of household under 21 years	0.01	0.01	2.02	−1.47	0.16	−269.53
Illiteracy rate of heads of household	0.04	0.03	−1.69	−1.17	0.26	70.78
Illiteracy rate of 8–12 years old	0.03	0.03	1.03	−1.71	0.11	−15.08
Teachers per students (5th to 8th grade)	0.05	0.04	−1.47	−0.20	0.84	60.65
HIV/AIDS incidence (100.000)	39.46	61.67	366.89	−1.38	0.19	−75.26
Proportion of infants (<1 year old)	0.01	0.01	0.46	−1.59	0.13	80.83
Proportion of elderly (≥65 years old)	0.09	0.09	0.76	1.12	0.28	85.94
Proportion of women	0.53	0.53	1.02	1.00	0.33	76.66

Figure 2 presents the difference in total age-adjusted mortality adjusted for age before and after propensity score matching. Before matching, higher income inequality *distritos* had slightly lower average mortality, 2.23 per 10.000 (95% CI: −23.92 – 19.46), which is contrary to the income inequality hypothesis. When the comparison is done after propensity score matching, higher inequality is associated with a significantly higher mortality rate (41.58 per 10.000, 95% CI: 8.85 – 73.3) (Figure 2).

Figure 2: Mortality difference between exposed and unexposed *distritos* before and after propensity score matching (PSM) (mean and 95% CI).

DISCUSSION

Before using propensity score matching, *distritos* with higher inequality had slightly lower overall mortality (though statistically non-significant). When we compared *distritos* with similar pre-selected background characteristics, differing only with regard to its inequality status, higher inequality had a statistically significant adverse impact on mortality.

An explanation for this result may be attributed to two characteristics of this study: it analyzed a very diverse and unequal Municipality, and it accounted for the complex heterogeneity within small areas by using the propensity score matching.

Brazil's cities have the greatest urban inequality in the world.[16] São Paulo is at the same time its cultural and economic center, the city with the largest number of poor, and where 63% of multinational corporations are settled.[17]

The heterogeneity of social status implies that social comparisons damaging to health should stay within São Paulo, as the extremes incomes (both very high and very low) that are a characteristic of the Brazilian society are present within the city. The best analogy for São Paulo is that it is not a city, but a country in itself with its residents crowded together. Studies that fail to corroborate the income inequality hypothesis have come, for the most part, from the analysis of a distinct group of small areas (e.g. neighborhoods and counties situated within broader areas).[4,18]

Our propensity score matching approach suggests that a misleading answer is obtained from the naïve comparison of high income inequality neighborhoods with

low income inequality neighborhoods. The naïve comparison suggests that high income inequality is somewhat beneficial for health – but that is likely an artifact of the correlation between extremely deprived neighborhoods and lower income inequality among its residents. When we compare unequal neighborhoods with more egalitarian neighborhoods matched on their propensity score, we obtain results that are more consistent with the prior literature, i.e. area income inequality is damaging to the health of residents. The causal inference problem that we have addressed in our analysis is distinct from the context/composition problem (i.e. whether the effect of area income inequality is confounded by individual poverty status) that has been addressed through multi-level regression analysis. Regression analysis does not explicitly deal with the “off support” inference that propensity score matching is designed to overcome.[19] The methodology made it possible to include 16 variables in our analysis, helping to overcome a continuous shortcoming on relative income studies, that is controlling for the extensive effect of area deprivation.

An alternative use of the propensity score is through stratification. This was not feasible in our study as the area under the ROC curve (0.907) was too high in relation to our size simple. That meant that each stratum was very homogenous for exposure, i.e. unexposed *distritos* had its propensity score situated mostly on the first two quintiles while exposed were predominant on the last two.

The choice of exposure was also restricted by our sample size. For better definitions of high and low inequality it might be more appropriate to exclude *distritos* situated on the middle, e.g. consider “exposed” the *distritos* with a Gini coefficient higher than 0.35 and “unexposed” the ones lower than 0.20. The

downside of this would be leaving even more *distritos* unmatched, decreasing our statistical power. Of the total 27 *distritos* that were effectively matched, 10 were unexposed and 17 exposed. That happened because the overlapping of *distritos* with similar propensity score was more frequent for values over 0.5 (which were for the most part considered exposed).

Propensity score analysis should be ideally used on larger samples in order to guarantee higher statistical power on standardized difference and bias reduction. This may be an issue in reproducing the findings in different areas. São Paulo is one of the largest cities in the world, and yet the smallest unit for which health data is available is limited to a sample of 96.

Longitudinal individual level analysis may be the best solution for assessing the relative income theory on smaller areas, using the covariates to provide background characteristics, exposure as living in an unequal area (neighborhood) and with the outcome as individual health indicators such as survival over time or morbidity diagnosis.

To our knowledge this is the first study to use the propensity score approach to analyze the relative income theory. New studies should follow, ideally using individual level data. Larger areas such as countries and US states have provided consistence evidence to support the relative income theory. The time has come to turn to the smaller areas, which have been relatively neglected despite fitting the original theory, which posits that social comparisons are more relevant at the local level.

Acknowledgments: The authors thank the Instituto Brasileiro de Geografia e Estatística (IBGE) for providing the data and assistance.

Funding: the study was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Competing interests: None.

REFERENCES

- 1 Kondo N, Sembajwe G, Kawachi I, et al. Income inequality, mortality, and self rated health: meta-analysis of multilevel studies. *BMJ* 2009;339:b4471.
- 2 Runciman WG. Relative deprivation and social justice. London: Routledge 1966.
- 3 Wilkinson RG. Commentary: Income inequality summarises the health burden of individual relative deprivation. *BMJ* 1997;314:1727.
- 4 Hou F, Myles J. Neighbourhood inequality, neighbourhood affluence and population health. *Soc Sci Med* 2005;60(7):1557-69.
- 5 Wilkinson RG, Pickett KE. Income inequality and population health: a review and explanation of the evidence. *Soc Sci Med* 2006;62:1768-84.
- 6 Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41-55.
- 7 Marques E, Torres H. São Paulo: segregação, pobreza e desigualdades sociais. São Paulo: Senac 2005: 9-15.
- 8 Central Intelligence Agency. The world factbook 2009.
www.cia.gov/library/publications/the-world-factbook/index.html (accessed 5 Jan 2010).
- 9 Instituto Brasileiro de Geografia e Estatística. Censo Demográfico 2000 -

Resultados do Universo: Agregados de setores censitários - São Paulo, Região Metropolitana. Brasília: IBGE 2003.

10 Fundação Seade, CEPID-FAPESP, Centro de Estudos da Metrópole/Cebrap.

Censo escolar 2001. www.centrodametropole.org.br/cd/escolas/ESC2001.rar

(accessed 5 Jan 2010).

11 Secretaria Municipal da Saúde de São Paulo. Boletim epidemiológico de aids do Município de São Paulo. São Paulo: Secretaria Municipal da Saúde 2003: 12.

12 Programa de Aprimoramento das Informações de Mortalidade no Município de São Paulo.

ww2.prefeitura.sp.gov.br/cgi/deftohtm.exe?secretarias/saude/TABNET/SIM/obito.def (accessed 5 Jan 2010).

13 Stürmer T, Joshi M, Glynn RJ, et al. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. *J Clin Epidemiol* 2006;59:437-47.

14 Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Statist Med* 2008;27:2037-49.

15 Normand ST, Landrum MB, Guadagnoli E, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores. *J Clin Epidemiol* 2001;54:387-398.

16 United Nations Human Settlements Programme. State of the world's cities

2008/2009 - harmonious cities. London: Earthscan 2008: 70.

17 Secretaria Municipal de São Paulo de Relações Internacionais. São Paulo: Cidade do mundo. portal.prefeitura.sp.gov.br/secretarias/relacoes_internacionais/0030 (accessed 5 Jan 2010).

18 Subramanian SV, Kawachi I. Income inequality and health: what have we learned so far? *Epidemiol Rev* 2004;26(1):78-91.

19 Oakes JM, Johnson PJ. Propensity Score Matching for social epidemiology. In J. M. Oakes, & J. S. Kaufman (Eds.), *Methods in social epidemiology*. San Francisco: Jossey-Bass 2006: 370-92.

ARTIGO 3

Elsevier Editorial System(tm) for Annals of Epidemiology

Manuscript Number: 10-292

Title:

Cause-specific mortality difference between high and low inequality areas of São Paulo, Brazil, by the use of the propensity score matching approach.

Article Type: Full Length Article

Keywords: income inequality; propensity score; mortality; social epidemiology; São Paulo.

Corresponding Author: Mr. Alexandre Dias Porto Chiavegatto Filho,

Corresponding Author's Institution: Faculdade de Saúde Pública, University of São Paulo

First Author: Alexandre D Chiavegatto Filho

Order of Authors: Alexandre D Chiavegatto Filho; Sabina L Gotlieb, MD, PhD;

Ichiro Kawachi, MD, PhD

Manuscript Region of Origin: BRAZIL

Word count: 2582.

ABSTRACT

Purpose. The relative income theory posits that an individual will be more susceptible to worse health if living in a highly unequal place. Results from previous studies are robust for larger areas but mixed for smaller ones such as neighborhoods. In highly segregated areas such as São Paulo, social and demographic variables do not follow a homogenous pattern, so comparable areas are difficult to identify. Propensity score matching is a statistical approach that restricts comparison to similar areas. **Methods.** We divided the 96 *distritos* (a proxy for neighborhood) of São Paulo, Brazil, into two groups: high inequality (Gini coefficient ≥ 0.25) and low (<0.25). 27 *distritos* had propensity score situated within the caliper width of 0.01 from a *distrito* with different exposure and were matched. **Results.** The difference between high and low inequality was statistically significant for homicide (8.57 per 10,000; 95% CI: 2.60 – 14.53), ischaemic 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). **Conclusion.** Income inequality studies in small areas need to better account for local heterogeneity and propensity score matching is an effective methodology to identify comparable areas.

INTRODUCTION

The association between income and health can be parsed into two distinct mechanisms: the absolute income effect, and the relative income effect. The absolute income mechanism is exemplified by the well established association between income poverty and death and illness. Living in poverty leads to lower access to health services, clean water, secure jobs, a decent education, as well as increased vulnerability to crime and natural disasters.¹

In contrast to the absolute income mechanism, the relative income hypothesis posits that an individual's health status is additionally determined by the relative social position conferred by a given level of income. The relative income hypothesis predicts that an individual with a given level income would experience worse health status by living in proximity to wealthier individuals compared to others with a similar standard of living. To paraphrase Seneca in *Epistles to Lucilius*, being poor in the midst of riches is the worst kind of poverty.

Among several theories that have been put forward to explain the relative income mechanism, one of the most influential is the one advanced by Wilkinson,² who has argued that the stress and shame induced by invidious social comparisons increase the individual's vulnerability to illness (the psychosocial hypothesis). Alternatively, or additionally, lower relative income can materially affect an individual's life opportunities.³ Thus, in many wealthy societies, it is not sufficient to own a telephone if everyone else has access to the Internet; or it is not enough to have access to the Internet via a modem if everyone else has access to broadband.

In the field of social epidemiology, one empirical test of the relative income hypothesis has been to examine whether the local distribution on income has an independent effect on mortality and morbidity above and beyond the effect of absolute income.

To date a number of empirical tests have been provided on the link between income distribution and health. A recent meta-analysis of multi-level studies found that there is a statistically significant association between higher income inequality and excess mortality or poor self-rated health.⁴ However, one of the major sources of heterogeneity is the area over which income inequality is defined.

According to an earlier review by Subramanian and Kawachi,⁵ the association between income inequality and health is more robust for studies conducted at larger spatial scales (in particular, U.S. states) than at smaller spatial scales (such as neighborhoods). This observation is contrary to the relative income theory when we consider that social comparisons are likely to be more salient when they are made across more proximal distances (i.e. one's immediate neighbors versus others living in the same state).

However, before rejecting the theory, we should note one important methodological problem in testing the relative income hypothesis at small spatial scales such as neighborhoods. The issue is that in highly segregated societies (such as the United States or Brazil), within-neighborhood distributions of income are lower than between-neighborhood distributions of income. Thus, incomes within *favelas* (urban slums) in Brazil are apt to be far more homogeneous (i.e. everybody is equally poor), compared to the disparity in incomes between the residents of slums versus the rest of the metropolitan area. In this case, the relevant social comparisons

may not be captured by focusing on the intra-neighborhood distribution of income.

A second challenge to the relative income theory is understanding the biological pathways by which inequality harms health. The psychosocial theory of inequality posits that the deleterious impacts of inequality are mediated by stress.² However, a recent review of the literature found that five out of nine studies that examined stress as a possible explanatory mechanism between SES and health found little or no effect of stress as a possible mediator.⁶ On the other hand, since “stress” is often poorly measured, we cannot reject the theory on the available evidence. An alternative approach is to examine the specific pattern of excess morbidity and mortality associated with inequality to check whether they are consistent with a stress mechanism. For example, accumulated evidence at the individual level points to an association between stress and cardiovascular disease; whereas there is less evidence for neoplasms. Wilkinson and Pickett analyzed seven different cause-specific mortality rates from 3139 counties in the United States and found that income inequality was related to ischaemic heart disease, respiratory disease, and homicide.⁷

The present paper proposes a contribution to the analysis of the relative income theory by focusing on the two challenges we have identified above. First we use a statistical methodology known as propensity score matching to address the problem of smaller spatial scales being correlated with a deceptively egalitarian distribution of income (when everyone is equally poor). Our study is set in a highly unequal and highly segregated metropolitan area of Brazil, São Paulo.⁸ Secondly, we seek to understand the cause-specific patterns of mortality associated with income inequality.

In addition to being one of the largest cities in the world, São Paulo is also one of the most unequal and segregated with regard to incomes. In the past, the city was considered to have a radial distribution of income, with the wealthiest residents living in the center and with poverty spreading around it, i.e. the farther away from the center the poorer its residents.⁹ With the increase of its population density, wealthier residents came to prefer more peripheral areas. The need for construction workers, maids, and other low-paying jobs moved workers to these areas. Extreme income inequality between closely-adjacent areas thus became one of the characteristics of São Paulo. With its more than 10 million residents, São Paulo is sometimes considered a country in itself.¹⁰ Even though a recent study in the US found that social comparisons with more distal reference groups (such as residents living in the rest of the state) were more strongly associated with health than social comparisons with neighbors,¹¹ our hypothesis in the case of São Paulo is that all of its residents' social comparisons stay within the city limits. That is, we assume that all possible comparison groups (either based on race or color or income) already exist within the city boundaries.

METHODS

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

We included in the model the following 16 *distrito*-level variables to account for area heterogeneity: prevalence of favelas (“slums”), poverty rate, median income, education of head of household (measured by number of years), household density, proportion of residences with piped water, proportion of residences with garbage collection, proportion of residents without toilet, 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 incidence, proportion of infants (under 1 year old), proportion of elderly (over 64 years old) and proportion of women. Most of the data comes from the 2000 Demographic Census,¹² the others being provided by the 2001 Educational Census (proportion of teachers per student)¹³ and the Municipal Epidemiological Bulletin (HIV/AIDS incidence).¹⁴

The variable used to define exposure on the propensity-score model, the Gini coefficient, was also calculated from the 2000 Demographic Census. The outcome variable, used for matching the *distritos*, was cause-specific (ten most frequent causes of death) and total mortality, available from the municipal government,¹⁵ which were subsequently adjusted for age. To minimize random annual variation within the *distritos*, we calculated mortality rates from 1998 to 2002, the five years

centered around 2000, the year of the last Census.

Propensity Score

Studies on the association between income distribution and health must control for the confounding effect of absolute income. Poverty at the local level may increase mortality due to lack of access to health-enhancing resources. As neighborhoods with high inequality may also present with high poverty, their effects may overlap.

In order to overcome the problem of confounding of area income inequality by area poverty, we used the technique of propensity score matching. The propensity score is defined as the probability of assignment to a particular treatment or exposure given a set of observed covariates, first introduced by Rosenbaum and Rubin.¹⁶

The probability of assignment to “treatment” (in this instance, a neighborhood having a high degree of income inequality) is estimated from logistic regression in which we enter all possible (and observed) causes of assignment to the treatment. The idea behind the propensity approach is to identify the subjects (in our case, the *distritos*) that are as alike as possible to each other with respect to the probability of being “exposed”.

We define exposure as having high income inequality (Gini coefficient > 0.25), and non-exposure as low income inequality (< 0.25). As there is no clear cut-off for high vs. low inequality, we selected 0.25 as it was close to the median Gini of the *distritos*.

The probability of exposure, i.e. the propensity score, ranges from 0 (high

probability of the *distrito* having low inequality) to 1 (high probability of high inequality). The variables selected in our model generated propensity scores considered to be “highly predictive” of exposure¹⁷ (the area under the ROC curve, or c-statistic, was 0.907).

The *distritos* with different exposures were subsequently matched on their propensity score, according to a pre-defined range (caliper width: <0.01). Of the total 96 *distritos* of São Paulo, 27 had a propensity score situated within caliper width of 0.01 of another *distrito* with opposite exposure, and were consequently included in the analysis.

RESULTS

Before conducting the matching process by propensity score value, there was no significant mortality difference between high inequality and low inequality *distritos* for either infant, elderly, men or women mortality (Table 1). When we analyzed the underlying causes of death, only neoplasm (higher on high inequality *distritos*) and diabetes (higher on low inequality *distritos*) had statistically significant different age-adjusted mortality.

Table 1 – Age-adjusted mortality (per 10,000) before propensity score matching.

	High Inequality	Low Inequality	Difference	95% CI	p-Test
Homicides	20.15	22.45	-2.3	-6.67 - 2.07	
HIV/AIDS	6.95	6.1	0.86	-0.60 - 2.31	
Neoplasm	53.55	49.4	4.14	0.90 - 7.39	<0.05
Ischaemic Heart Disease	39.52	39.33	0.19	-2.61 - 2.98	
Hypertensive Diseases	7.34	8.46	-1.13	-2.34 - 0.09	
Cerebrovascular Diseases	25.17	27.55	-2.38	-5.02 - 0.27	
Diabetes	9.56	11.67	-2.11	-3.48 - -0.74	<0.01
Respiratory Diseases	31.02	31.45	-0.43	-2.86 - 2.01	
Car Accidents	5.3	5.63	-0.34	-0.96 - 2.91	
Cardiac Insufficiency	6.97	8.07	-1.1	-4.46 - 2.27	
Infant Mortality	15	14.12	0.88	-0.55 - 2.32	
65 and Older	148.92	145.8	3.12	-5.91 - 12.15	
Men	166.14	166.3	-0.16	-13.33 - 13.01	
Women	129.68	127.91	-1.23	-10.08 - 7.63	

After propensity score matching, four of the ten most common underlying causes of death had significantly higher mortality on high income inequality *distritos* than on low inequality *distritos* (Table 2). The highest difference was for homicide

(8.57 per 10,000; 95% CI: 2.60 – 14.53), followed by ischaemic 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 of the ten underlying causes analyzed had higher mortality rates for high inequality *distritos*.

Table 2 – Age-adjusted mortality (per 10,000) after propensity score matching.

	High Inequality	Low Inequality	Difference	95% CI	p-Test
Homicides	19.62	11.05	8.57	2.60 - 14.53	<0.01
HIV/AIDS	8.62	5.04	3.58	0.58 - 6.57	<0.05
Neoplasm	55.68	53.3	2.38	-0.39 - 5.16	
Ischaemic Heart Disease	41.66	36.2	5.47	0.76 - 10.17	<0.05
Hypertensive Diseases	6.81	5.85	0.96	-0.62 - 2.53	
Cerebrovascular Diseases	25.53	22.57	2.96	-0.32 - 6.25	
Diabetes	9.46	8.94	0.53	-1.42 - 2.47	
Respiratory Diseases	32.2	28.64	3.56	0.18 - 6.94	<0.05
Car Accidents	5.52	4.91	0.61	-0.48 - 1.71	
Cardiac Insufficiency	6.37	4.49	1.44	-2.27 - 5.14	
Infant Mortality	15.42	12.61	2.8	0.86 - 4.74	<0.01
65 and Older	153.08	144.22	8.85	-0.72 - 18.44	
Men	172.52	145.15	27.37	6.19 - 48.55	<0.05
Women	129.6	114.53	15.07	3.65 - 26.48	<0.05

Infant mortality also presented significantly higher age-adjusted mortality for high inequality *distritos* (2.8 per 10,000; 95% CI: 0.86 – 4.74), as well as men's (27.37 per 10,000; 95% CI: 6.19 – 48.55) and women's mortality (15.07 per 10,000; 95% CI: 3.65 – 26.48). Elderly mortality was also higher for high income *distritos* (8.85 per 10,000), although non-significant (95% CI: -0.72 – 18.44).

When we analyzed the relative importance of each underlying cause to the total mortality difference between high and low inequality *distritos* (41.58 per

10,000), homicides account for 20,6% of excess mortality, ischaemic heart disease for 13,16% and HIV/AIDS for 8,61% (Figure 1). The ten most common causes of death are responsible for 72,3% of the total mortality difference.

Figure 1 – Cause-specific mortality as part of the overall mortality difference between high-inequality and low-inequality *distritos*, after PSM.

DISCUSSION

We sought to test the impact of income inequality on cause-specific mortality by employing propensity score matching to restrict the analysis to comparable areas. After propensity score matching, more unequal *distritos* had higher mortality rates for all of the ten underlying causes of death tested. Four of those were statistically significant (homicide, ischaemic heart disease, HIV/AIDS and respiratory diseases). Excess mortality was found for both men and women, as well as for infant mortality. Our results are consistent with the report by Wilkinson and Pickett⁷ on the association between income inequality and cause-specific mortality in US states. In their analyses, income inequality was statistically significantly associated with mortality from ischaemic heart disease, homicide and respiratory diseases, though not diabetes or cancer.

The psychosocial hypothesis posits that the relative distribution of area income adversely affects health because of the stress and shame induced by invidious social comparisons between residents. This may account for the excess mortality from ischaemic heart disease, which has a well established association with stress.¹⁸ Growing evidence suggests that stress is also linked to pulmonary function.¹⁹ Furthermore, the majority of respiratory deaths in our study is likely to be due to chronic obstructive pulmonary disease (COPD), which is in turn primarily linked to cigarette smoking.²⁰ In other words, excess deaths from respiratory disease can be plausibly linked to higher psychosocial stress and smoking in the population.

Homicide and HIV/AIDS can similarly reflect the behavioral responses to

residents' feelings of truncated life opportunities. Youth growing up in highly segregated (and consequently low-mobility) areas feel that there is no future for them, which in turn increases the probability of higher risk-taking behavior such as joining a street gang or not wearing a condom during sex.

São Paulo is a relevant area to study the effects of relative income on health because of its highly segregated nature. Even though it presents a Gini coefficient of 0.51, the median value for the 96 *distritos* is 0.25. Living in poverty in São Paulo is not associated with living in an unequal neighborhood; to the contrary, the 14 poorest *distritos* have a Gini coefficient below 0.25, meaning that poorer neighborhoods have a more equal distribution of income, i.e. the 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 indicates that even for a notoriously unequal city such as São Paulo, this is still a factor. Only one *distrito* stood out as highly unequal (Vila Andrade, with a Gini coefficient of 0.55).

Three sets of advantages have been claimed for the propensity score matching approach.²¹ First, it allows the investigator to explicitly check for overlap in the confounder distribution prior to conducting analysis – a step that is skipped when conducting “black box” multivariable-adjusted regression. The exposed and unexposed groups are then explicitly matched, and comparisons are drawn only within the range of overlap in propensity scores, thereby averting the problem of “off support” inference. Secondly, the propensity score deals with the dimensionality problem in multivariable analysis by reducing a large set of covariates to a single scalar variable. Lastly, the propensity score approach does not rely on any particular

functional form (e.g. linearity) between the covariate and the outcome within each treatment group.

Against the advantages, the major limitation of the propensity score approach is that we can only account for the set of observed covariates to develop the model. In other words, it cannot overcome the problem of confounding by unobserved factors. Nor can the approach deal with the endogeneity of residential preferences, which is a problem in causal inference that applies to all observational studies dealing with area effects on the health of residents.²²

The multivariable approach was necessary for São Paulo due to its high heterogeneity characteristic. It fails to have an identifiable pattern of distribution of resources and demography that makes one part of the city easily comparable to another. 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 most of the unsupportive result between income inequality and health found on unequal areas would benefit from the introduction of a methodology such as propensity score matching.

REFERENCES

- 1 – Wilkinson R, Marmot M. Social determinants of health: The solid facts. Copenhagen, Denmark: World Health Organization; 2003.
- 2 – Wilkinson R. The impact of inequality: How to make sick societies healthier. New York: The New Press; 2005.
- 3 – Kawachi I, Kennedy BP. The health of nations: How inequality is harmful to your health. New York: The New Press; 2002.
- 4 – Kondo N, Sembajwe G, Kawachi I, van Dam RM, Subramanian SV, Yamagata Z. Income inequality, mortality and self rated health: meta-analysis of multilevel studies. BMJ 2009; 339: b4471.
- 5 –Subramanian SV, Kawachi I. Income inequality and health: what have we learned so far? Epidemiologic Reviews 2004; 26: 78-91.
- 6 – Matthews KA, Gallo LC, Taylor SE. Are psychological factors mediators of socioeconomic status and health connections? Annals of the New York Academy of Sciences 2010; 1186: 146-173.
- 7 – Wilkinson R, Pickett KE. Income inequality and socioeconomic gradients in mortality. American Journal of Public Health 2008; 98: 699-704.
- 8 – Torres HG, Marques E, Ferreira MP, Bitar S. Pobreza e espaço: padrões de segregação em São Paulo. Estudos Avançados 2003; 17: 97-128.
- 9 – Singer P. Desenvolvimento econômica e evolução urbana. São Paulo: Companhia

Editora Nacional; 1977.

10 – Caldeira TPR. Cidade de muros: Crime, segregação e cidadania em São Paulo. São Paulo: Editora 34; 2008.

11 – Wolff LS, Subramanian SV, Acevedo-Garcia D, Kawachi I. Compared to whom? Subjective social status, self-rated health, and referent group sensitivity in a diverse US sample. *Social Science & Medicine* 2010; 70: 2019-2028.

12 – IBGE: Instituto Brasileiro de Geografia e Estatística. Censo Demográfico 2000 - Resultados do Universo: Agregados de setores censitários - São Paulo, Região Metropolitana. Brasília: IBGE; 2003.

13 – Fundação Seade, CEPID-FAPESP, Centro de Estudos da Metrópole/Cebrap. Censo escolar 2001. Retrieved from: <http://www.centrodametropole.org.br/cd/escolas/ESC2001.rar>.

14 – SMSSP: Secretaria Municipal da Saúde de São Paulo. Boletim epidemiológico de aids do Município de São Paulo. São Paulo: Secretaria Municipal da Saúde; 2003.

15 – PROAIM: Programa de Aprimoramento das Informações de Mortalidade no Município de São Paulo; 2001. Retrieved from: ww2.prefeitura.sp.gov.br/cgi/deftohtm.exe?secretarias/saude/TABNET/SIM/obito.def.

16 – Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983; 70: 41-55.

17 – Stürmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional

multivariable methods. *J Clin Epidemiol* 2006; 59: 437-47.

18 – Kubzansky L, Kawachi I. Affective states and health. In L. F. Berkman, & I. Kawachi (Eds.), *Social epidemiology* (pp. 213-241). New York: Oxford University Press; 2000.

19 – Clougherty JE, Kubzansky LD. A framework for examining social stress and susceptibility to air pollution in respiratory health. *Environmental Health Perspectives* 2009; 117: 1351-1358.

20 – Dourado VZ, Tanni SE, Vale SA, Faganello MM, Sanchez FF, Godoy I. Manifestações sistêmicas na doença pulmonar obstrutiva crônica. *J Bras Pneumol* 2006; 32(2): 161-171.

21 – Rubin DB. Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine* 1997; 127: 757-763.

22 – Oakes JM, Johnson PJ. Propensity Score Matching for social epidemiology. In J. M. Oakes, & J. S. Kaufman (Eds.), *Methods in social epidemiology* (pp. 370-392). San Francisco: Jossey-Bass; 2006.

5 CONCLUSÕES

O presente estudo encontrou associação estatisticamente significativa entre desigualdade e mortalidade. Distritos mais desiguais apresentaram maior mortalidade geral ajustada por idade (41,58 por 10.000 hab; IC95%: 8,85 – 73,3 por 10.000 hab), após a aplicação da metodologia do *propensity score matching*. Outra possibilidade de apresentação desses dados (não incluída nos artigos) é sob a forma da regressão linear: $Y = 213,27 + 256,8 * \text{Gini}$ (no qual Y representa a mortalidade geral ajustada por idade). Assim, um aumento de 0,1 no índice de Gini esteve associado a um aumento de 25,68 na mortalidade por 10.000 habitantes.

Foram também estatisticamente significativas as diferenças de mortalidade para homicídios (8,57 por 10.000; IC 95%: 2,60 – 14,53), doença isquêmica do coração (5,47 por 10.000; IC 95%: 0,76 – 10,17), aids (3,58 por 10.000; IC 95%: 0,58 – 6,57), doença respiratória (3,56 por 10.000; IC 95%: 0,18 – 6,94) e mortalidade infantil (2,8 por 10.000; IC 95%: 0,86 – 4,74).

A teoria da renda relativa ainda está longe de ser uma unanimidade na literatura científica. Os diferentes tipos de análises estatísticas e as regiões utilizadas impedem o consenso. O presente estudo, porém, tem fortes argumentos em favor da teoria. Primeiramente, a área de estudo foi São Paulo, um município com alta diversidade racial, religiosa e econômica. A convivência diária com alta desigualdade nas ruas – e no céu (é possível identificar um certo orgulho paulistano no comentário de que “temos a segunda maior frota de helicópteros do mundo”,

apesar de tudo) – tem maior potencial de afetar as comparações sociais negativas do que simplesmente ler sobre isso nos jornais ou assistir na televisão, por exemplo. A capacidade de adaptação dos seres humanos a condições adversas não pode ser descartada, mas a indiferença à desigualdade paulistana é difícil, mesmo para quem sempre morou na cidade.

Em segundo lugar, as análises realizadas apontaram que o uso do *propensity score matching* foi eficiente para a diminuição do efeito da heterogeneidade dos distritos de São Paulo. O *propensity score* tem uma aplicação semelhante à da regressão multivariada, cujo objetivo também é controlar por variáveis de confusão; mas difere por introduzir esse controle na exposição e não na variável de interesse (no caso, a mortalidade ajustada por idade). A eficácia da utilização do método acontece quando se verifica uma diminuição da diferença entre as variáveis das regiões incluídas no modelo (AUSTIN, 2008), o que foi o caso para o presente estudo.

Apesar de possibilitar a introdução um grande número de variáveis, uma possível limitação dessa metodologia é não conseguir controlar pelas variáveis de confusão não incluídas no modelo, ao contrário dos estudos randomizados (LUELLEN, 2005).

Outra limitação da metodologia é a importância de uma amostra inicial grande devido à necessidade da sobreposição dos valores do *propensity score* (dentro do *caliper width*) entre expostos e não expostos. No caso do presente estudo, apenas 27 dos 96 distritos foram incluídos no modelo. Levando-se em consideração o tamanho de São Paulo, isso gera limitações para a aplicação da metodologia em cidades menores.

A metodologia utilizada também tem como limitação o fato de usar dados agregados, o que impede uma análise dos possíveis efeitos individuais da desigualdade. Assim, a definição do mecanismo de ação pelo qual a desigualdade afeta a mortalidade não pôde ser identificado pela pesquisa. Supõe-se, contudo, que as três hipóteses definidas por LYNCH et al. (2000) (a neomaterial, a da renda individual e a do ambiente psicossocial), mencionadas anteriormente, atuem em conjunto.

Este estudo procurou testar a eficácia da introdução de uma nova metodologia na área de desigualdade de renda e saúde, com o objetivo de complementar a utilização de outras análises já bem estabelecidas como *clusters*, modelos marginais, multiníveis e regressões multivariadas (SUBRAMANIAN e KAWACHI, 2004).

Outras possíveis limitações do estudo referem-se às variáveis incluídas no modelo. Primeiramente, a utilização de dados do censo segundo setores censitários pode simplificar a variabilidade existente dentro desse mesmo setor (CAUGHY et al., 1999). Além disso, o fato de alguns distritos serem pequenos pode fazer com que eles sejam mais afetados por valores extremos (nulos, principalmente) do que distritos maiores.

A análise da mortalidade segundo causa básica está também sujeita a possíveis erros de classificação. Laurenti et al. (2004) analisaram declarações de óbitos de mulheres em idade fértil e verificaram que, para as capitais brasileiras, tornou-se necessária a adoção de fator de correção igual a 1,44, para a mortalidade materna. Outro fator também limitante desse tipo de análise é que o PRO-AIM inclui apenas os óbitos ocorridos no Município de São Paulo (DRUMOND JR, 1999), o que leva à subestimação do total de óbitos dos residentes da cidade.

6 COMENTÁRIOS FINAIS

Existe um debate contínuo sobre as causas históricas da desigualdade nas Américas. A teoria mais divulgada atualmente é a sintetizada por ACEMOGLU et al. (2002) na qual os tipos de colonização européia definiram a desigualdade de renda inicial e o crescimento futuro. Segundo essa teoria, áreas com riqueza de recursos naturais e de mão-de-obra apresentaram uma colonização do tipo extrativa, nas quais o poder era controlado por uma pequena elite exportadora, que não possuía interesse na criação de instituições jurídicas e democráticas ou no desenvolvimento econômico local. Regiões inicialmente mais pobres, por outro lado, tiveram de criar regras que permitissem o investimento estrangeiro e o desenvolvimento local. Segundo os autores, isso criou uma “inversão institucional”: regiões pobres passaram a investir no desenvolvimento da propriedade privada, enquanto que regiões inicialmente mais ricas sofreram de uma paralisia institucional.

Segundo SALOMÃO FILHO et al. (2008), a pobreza e a desigualdade de renda na América Latina não são apenas uma externalidade dos processos coloniais, mas sim elementos definidores do próprio modelo econômico. Segundo os autores, a estrutura do pensamento colonialista no Brasil era baseada no incentivo à formação de monopólios e na exclusão deliberada de qualquer outro setor da economia, com o objetivo final de impedir o desenvolvimento de um mercado interno e de manter os salários no nível de subsistência. A desigualdade de renda foi assim mantida para permitir lucros altos, maiores jornadas de trabalho e a expropriação de todos os ganhos de produtividade dos trabalhadores. As elites se revezaram no poder nos

séculos seguintes, mas a estrutura econômica e jurídica sofreu apenas pequenas adaptações.

Para JUDT (2010), o debate sobre as origens primárias da desigualdade perdeu a relevância de antigamente. Independentemente do tipo de colonização, a maioria dos países do continente americano já não sofre mais de um problema de escassez de recursos, mas de distribuição desses recursos. A história recente aponta para o fato de que a desigualdade é, mais do que muitos governos gostariam de admitir, uma opção política, como mostra o seu crescimento nos EUA a partir do governo ideologicamente favorável aos mais ricos de Reagan e no Reino Unido a partir do governo Thatcher (KRUGMAN, 2007). No Brasil, o aumento da desigualdade a partir da metade do século passado foi incentivado pelo governo militar para diminuir o salário manual, com o objetivo de acelerar o desenvolvimento da indústria nacional (LUNA e KLEIN, 2007). A diminuição mais recente da desigualdade brasileira se deu, em um primeiro momento, pelo controle inflacionário ocorrido em meados da década de 90, que permitiu a estabilidade da renda, e depois pela ampliação dos programas de distribuição de renda (NERI, 2006), duas ações definitivamente políticas.

Um relatório recente (MARMOT, 2010), encomendado pelo ministério da saúde do Reino Unido, analisou as causas das desigualdades de saúde na Inglaterra, com o objetivo de diminuí-las a partir de 2010. Para os autores, as diferenças encontradas entre os indicadores de saúde de pobres e ricos são uma consequência da desigualdade social e podem ser evitadas se houver vontade política. O desenvolvimento de uma sociedade justa, sem segregação social e com oportunidades iguais para pobres e ricos foi considerado um importante passo para a

melhoria das condições de saúde. Segundo os autores, a diminuição da desigualdade pode também trazer benefícios econômicos como aumento de produtividade, diminuição de gastos sociais e maior arrecadação tributária.

A desigualdade, contudo, não é o fator mais importante na determinação da saúde da população. Alguns críticos da teoria da renda relativa costumam utilizar exemplos que sobrevalorizam a importância da desigualdade, com o objetivo de refutá-la. Para CANNING e BOWSER (2010), um exemplo da falha da teoria é o fato de os EUA terem aumentado seus níveis de desigualdade desde a década de 70 e mesmo assim terem melhorado sua expectativa de vida. Como mencionado pelos autores, os ganhos em tecnologia médica e renda absoluta no período são os prováveis responsáveis pela melhoria da saúde. A desigualdade de renda, na maior parte dos casos, atua como um fator complementar na determinação das condições de saúde.

É importante ainda apontar que a desigualdade não é apenas um problema de transferência direta de renda por parte do governo, por mais importante que isso seja. É preciso também compreender as razões que permitiram a persistência da alta desigualdade por tanto tempo. Críticos da teoria chegam a sugerir que os seus defensores deveriam simplesmente expulsar os ricos de seu país, o que traria uma melhora automática dos níveis de saúde (FISCHER, 2010). O que precisa ficar claro, contudo, é que a alta desigualdade é um sintoma mais profundo dos problemas da sociedade. Solucionar a questão da desigualdade de uma só vez, *deus ex machina*, deixaria intactos todos os males sociais que levaram inicialmente à alta desigualdade, além daqueles que foram posteriormente causados pela desigualdade.

O presente estudo utilizou dados do Censo de 2000, ano em que o coeficiente de Gini brasileiro era de 56,7. Desde então ocorreu uma queda de 4,9% para o último período disponível (2008/2009), chegando a 53,9, consequência do aumento da importância do rendimento do trabalho na renda total (IPEA, 2010). Essa tendência de queda tem sido constante nas últimas décadas (desde 1995 houve apenas um aumento anual) (DE BARROS et al., 2006). É difícil saber ainda se essa mudança, ainda muito pequena, pode ter influência no resultado de novas análises, mas será importante analisar se a associação entre desigualdade de renda e maior mortalidade verificada para São Paulo, mas não para a maioria dos estudos internacionais, é consequência apenas da alta desigualdade paulistana ou de características particulares das estruturas sociais e econômicas brasileiras.

6 REFERÊNCIAS

Acemoglu D, Johnson S, Robinson JA. Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Q J Econ*. 2002;117(4):1231-94.

Aghion P, Caroli E, Garcia-Peñalosa C. Inequality and economic growth: the perspective of the new growth theories. *J Econ Lit*. 1999;37(4):1615-60.

Antunes JLF. Mortalidade por câncer e desigualdade social em São Paulo [tese de livre-docência]. São Paulo: Faculdade de Odontologia da USP; 2004.

Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Statist Med* 2008;27:2037-49

Benjamin D, Brandt L, Giles J. Inequality and growth in rural China: does higher inequality impede growth? [Discussion paper]. Bonn: Institute for the Study of Labor; 2006 [acesso em 11 ago 2010]. Disponível em: http://homes.chass.utoronto.ca/~brandt/recentpapers/IZA_BBG.pdf

Blanden J, Gregg P, Machin S. Intergenerational mobility in Europe and North America. Report. London; 2005. (Centre for Economic Performance).

BMJ Editor's Choice. The big idea [editorial]. *BMJ*. 1996;312(7037):0.

Bok D. The politics of happiness: what government can learn from the new research

on well-being. Princeton: Princeton University Press; 2010.

Brickman P, Coates D, Janoff-Bulman R. Lottery winners and accident victims: is happiness relative? *J Pers Soc Psychol.* 1978;36(8):917-27.

Canning D, Bowser D. Investing in health to improve the wellbeing of the disadvantaged: reversing the argument of fair society, healthy lives (The Marmot Review). *Soc Sci Med.* In press 2010.

Caughy MO, O'Campo P, Brodsky AE. Neighborhoods, families, and children: implications for policy and practice. *J Community Psychol.* 1999;27(5):615–33.

D'Addio AC. Intergenerational transmission of disadvantage: mobility or immobility across generations? A review of the evidence for OECD countries. Report. Paris; 2007. (OCDE social, employment and migration working papers).

de Barros RP, de Carvalho M, Franco S, Mendonça R. Uma análise das principais causas da queda recente da desigualdade de renda brasileira. *Econômica.* 2006;8(1):117-47.

de Wall F. The age of empathy: nature's lessons for a kinder society. New York: Crown; 2009.

Deaton A, Lubotsky, D. Mortality, inequality and race in American cities and states. *Soc Sci Med.* 2003;56(6):1139-53.

Diez-Roux AV, Link BG, Northridge ME. A multilevel analysis of income inequality and cardiovascular disease risk factors. *Soc Sci Med.* 2000;50(5):673-87.

Drumond Jr, M. Homicídios e desigualdades sociais na cidade de São Paulo: uma visão epidemiológica. *Saúde Soc.* 1999;8(1):63-81.

Easterlin RA. Will raising the incomes of all increase the happiness of all? *J Econ Behav Organ.* 1995;27(1): 35-47.

Fehr E, Gächter S. Altruistic punishment in humans. *Nature.* 2002;415:137-40.

Fehr E, Gintis H. Human motivation and social cooperation: experimental and analytical foundations. *Annu Rev Sociol.* 2007;33:43-64.

Fehr E, Schmidt KM. A theory of fairness, competition, and cooperation. *Q J Econ.* 1999;114(3):817-68.

Feldstein M. Reducing poverty, not inequality. *The Public Interest.* Fall 1999;137:33-41.

Firebaugh G, Tach LM. *Income and happiness in the United States*. Philadelphia; 2005. [atualizado em 25 maio 2009; acesso em 10 agosto 2010]. Disponível em: http://www.allacademic.com/meta/p_mla_apa_research_citation/0/1/8/1/6/p18167_index.html

Fischer CS. Mind the gap. *Boston Review.* July/August 2010. Disponível em: <http://bostonreview.net/BR35.4/fischer.php>

Fowler JH, Johnson T, Smirnov O. Egalitarian motive and altruistic punishment. *Nature.* 2004;433:E1.

Gintis H, Henrich J, Bowles S, Boyd R, Fehr E. Strong reciprocity and the roots of

human morality. *Soc Just Res.* 2008;21(2):241-53.

Gold R, Kawachi I, Kennedy BP, Lynch JW, Connell FA. Ecological analysis of teen birth rates: association with community income and income inequality. *Matern Child Health J.* 2001;5(3):161-7.

Hill J, Brewer M, Jenkins S, Lister R, Lupton R, Machin S, et al. An anatomy of economic inequality in the UK. Report. London; 2010. (National Equality Panel).

Hoff K, Pandey P. Belief systems and durable inequalities: an experimental investigation of Indian caste. [Working paper] Washington: World Bank; 2004. Disponível em: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=610395

IPEA-Instituto de Pesquisa Econômica Aplicada. Distribuição funcional da renda pré e pós crise internacional no Brasil. Brasília; 2010. Disponível em: http://agencia.ipea.gov.br/images/stories/PDFs/100505_comunicaipea_47_distribuicaoreda.pdf

James RN, Sharpe DL. The nature and causes of the U-shaped charitable giving profile. *Nonprofit and Voluntary Sector Quarterly.* 2007;36(2):218-38.

Judt T. *Ill fares the land.* New York: Penguin Press; 2010.

Kaplan GA, Pamuk ER, Lynch JW, Cohen RD, Balfour JL. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *BMJ.* 1996;312(7037):999-1003.

Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D. Social capital, income

inequality, and mortality. *Am J Public Health*. 1997;87(9):1491-8.

Kawachi I, Kennedy BP. *The health of nations: why inequality is harmful to your health*. New York: The New Press; 2002.

Kawachi I. Social capital and community effects on population and individual health. *Ann N Y Acad Sci*. 1999;896:120-30.

Krugman P. *The conscience of a liberal*. New York: W. W. Norton & Company; 2007.

Laurenti R, Jorge MHPM, Gotlieb SLD. A mortalidade materna nas capitais brasileiras: algumas características e estimativa de um fator de ajuste. *Rev Bras Epidemiol*. 2004;7(4):449-60.

Lochner K, Kawachi I, Kennedy BP. Social capital: a guide to its measurement. *Health Place*. 1999;5(4):259-70.

Luellen JK, Shadish WR, Clark MH. Propensity scores: an introduction and experimental test. *Eval Rev*. 2005;29(6):530-558.

Luna FV, Klein HS. *O Brasil desde 1980*. São Paulo: A Girafa; 2007.

Lynch J, Smith GD, Harper S, Hillemeier M, Ross N, Kaplan GA, et al. Is income inequality a determinant of population health? Part 1. A systematic review. *Milbank Q*. 2004;82(1):5-99.

Lynch JW, Smith GD, Kaplan GA, House JS. Income inequality and mortality: importance to health of individual income, psychosocial environment, or material

conditions. *BMJ*. 2000;320(7243):1200-4.

Marmot M. *Fair Society, Healthy Lives*. London; 2010. Disponível em: <http://www.marmotreview.org/AssetLibrary/pdfs/Reports/FairSocietyHealthyLivesExecutiveSummary.pdf>

Morgan D, Grant KA, Gage HD, Mach RH, Kaplan JR, Prioleau O, et al. Social dominance in monkeys: dopamine D2 receptors and cocaine self-administration. *Nat Neurosci*. 2002;5(2):169-74.

Navarro N, Shi L. The political context of social inequalities and health. *Soc Sci Med*. 2001;52(3):481-91.

Neri MC. *Miséria, desigualdade e estabilidade: o segundo Real*. Rio de Janeiro; 2006. Disponível em: http://www.fgv.br/cps/pesquisas/site_ret_port/RET_Texto.pdf

Pickett KE, Mookherjee J, Wilkinson RG. Adolescent birth rates, total homicides, and income inequality in rich countries. *Am J Public Health*. 2005;95(7):1181-3.

Piff PK, Kraus MW, Côté S, Cheng BH, Keltner D. Having less, giving more: The influence of social class on prosocial behavior. *J Pers Soc Psychol*. 2010. In press.

Putnam RD. *Bowling alone: the collapse and revival of american community*. New York: Simon & Schuster; 2000.

Rothstein B, Uslander EM. All for all: equality, corruption, and social trust. *World Polit*. 2005;58(1):41-72.

Salomão Filho C, Ferrão BLM, Ribeiro IC. *Concentração, estruturas e desigualdade:*

as origens coloniais da pobreza e da má distribuição de renda. São Paulo: Grupo Direito e Pobreza; 2008.

Shively CA, Clarkson TB. Social status and coronary artery atherosclerosis in female monkeys. *Arterioscler Thromb*. 1994;14(5):721-6.

Subramanian SV, Kawachi, I. Income inequality and health: what have we learned so far? *Epidemiol Rev*. 2004;26(1):78-91.

Szwarcwald CL, Bastos FI, Esteves MAP, de Andrade CLT, Paez MS, Medici EV, et al. Desigualdade de renda e situação de saúde: o caso do Rio de Janeiro. *Cad Saude Publica*. 1999;15(1):15-28.

Verba S. Schlozman KL, Brady H. Voice and equality: civic voluntarism in american politics. Cambridge: Harvard University Press; 1995.

Vicentin G, Santo AH, Carvalho MS. Mortalidade por tuberculose e indicadores sociais no município do Rio de Janeiro. *Cien Saude Colet*. 2002;7(2):253-63.

Watkins K. Human Development Report 2005: International cooperation at a crossroads: aid, trade and security in an unequal world. Report. New York; 2005. (UNDP Human Development Report 2005).

Wedekind C, Milinski M. Cooperation through image scoring in humans. *Science*. 2000;288(5467):850-2.

Wilkinson R, Pickett K. The spirit level: why more equal societies almost always do better. London: Penguin Group; 2009.

Wilkinson R, Pickett KE. Income inequality and population health: a review and explanation of the evidence. *Soc Sci Med*. 2006;62(7):1768-84.

Wilkinson R. *Unhealthy societies: the afflictions of inequality*. New York: Routledge; 1996.

Wilkinson R. Why is violence more common where inequality is greater? *Ann N Y Acad Sci*. 2004;1036:1-12.

ANEXO 1

Deliberação da CPG em sua sessão 9ª/2008 de 05/06/2008

Diretrizes para outra forma de apresentação de Tese e Dissertação dos Programas de Pós Graduação da Faculdade de Saúde Pública da USP

- A **Dissertação de Mestrado** deve incluir, no mínimo, um manuscrito resultante do seu projeto de pesquisa de mestrado, submetido ou publicado em periódico arbitrado por pares, no período de sua matrícula no Programa.
- A **Tese de Doutorado** deve incluir, no mínimo, três manuscritos resultantes do seu projeto de pesquisa de doutorado, sendo que um deles não deve ter sido submetido ou publicado em periódico ou capítulo de livro. Os outros dois artigos devem ter sido submetidos ou publicados em periódicos ou capítulos de livro arbitrados por pares, no período de sua matrícula no Programa.
- São considerados como manuscritos: periódicos ou capítulos de livros. Em todos os manuscritos o candidato a Mestre ou Doutor deverá ser o primeiro autor.
- As Teses e Dissertações devem incluir obrigatoriamente:

Apresentação: descrição da estrutura do trabalho

I. Introdução: texto que sistematize o conhecimento existente e justifique o trabalho realizado.

II. Objetivos

III. Metodologia: descrição das definições, procedimentos e técnicas adotados para a realização do trabalho, não contempladas nos manuscritos apresentados.

IV. Resultados e discussão: onde serão inseridos os manuscritos que apresentam os resultados e discussão correspondente do trabalho realizado que permitam as conclusões e considerações finais.

V. Conclusões ou considerações finais: em que as contribuições do trabalho sejam sintetizadas.

VI. Referências bibliográficas: utilizadas no trabalho

VII. Anexos

- Os manuscritos devem ser formatados segundo as normas para publicação dos periódicos ou obras a que serão submetidos ou publicados.
- As referências bibliográficas devem ser apresentadas ao final de cada capítulo e, quando não se referirem aos manuscritos apresentados, devem seguir as normas gerais do Guia de Teses da FSP.

*Prof. Dr. Arlindo Philippi Jr.
Presidente da CPG*