Income distribution and mortality: cross sectional ecological study of the Robin Hood index in the United States

Bruce P Kennedy, Ichiro Kawachi, Deborah Prothrow-Stith

See editorial, pp 999, 1015

Abstract

***Objective*—To determine the effect of income inequality as measured by the Robin Hood index and the Gini coefficient on all cause and cause specific mortality in the United States.**

***Design*—Cross sectional ecological study.**

***Setting*—Households in the United States.**

***Main outcome measures*—Disease specific mortality, income, household size, poverty, and smoking rates for each state.**

***Results*—The Robin Hood index was positively correlated with total mortality adjusted for age ($r=0{\cdot}54$; $P<0{\cdot}05$). This association remained after adjustment for poverty ($P<0{\cdot}007$), where each percentage increase in the index was associated with an increase in the total mortality of 21·68 deaths per 100 000. Effects of the index were also found for infant mortality ($P=0{\cdot}013$); coronary heart disease ($P=0{\cdot}004$); malignant neoplasms ($P=0{\cdot}023$); and homicide ($P<0{\cdot}001$). Strong associations were also found between the index and causes of death amenable to medical intervention. The Gini coefficient showed very little correlation with any of the causes of death.**

***Conclusion*—Variations between states in the inequality of income were associated with increased mortality from several causes. The size of the gap between the wealthy and less well off—as distinct from the absolute standard of living enjoyed by the poor—seems to matter in its own right. The findings suggest that policies that deal with the growing inequities in income distribution may have an important impact on the health of the population.**

Introduction

A small number of cross national studies have suggested a relation between income distribution and life expectancy: the greater the gap in income between the rich and poor in any given society the lower the average life expectancy.[1-5] In one study of 11 countries in the Organisation of Economic Cooperation and Development a strong inverse correlation was found between income inequality—as measured by Gini coefficients of income after tax standardised for differences in household size—and average life expectancy.[2] This association seems to be independent of that between absolute income and life expectancy. In other words, it matters not only how well off a country is but also how economic gains are distributed among its members.

The mechanisms underlying the association between income distribution and mortality are poorly understood.[6,7] For instance, it is not clear whether income distribution is related to particular causes of death, such as infant mortality,[1,8] more than other causes. Published studies to date have focused almost exclusively on average life expectancy or overall mortality and have failed to report data on specific causes of death. Previous studies also have entailed comparisons across different countries, raising the question of comparability and completeness of income data. We examined the relation between income distribution and all cause and cause specific mortality within the United States.

Methods

SOURCES OF DATA

Data on income, household size, and poverty were obtained from the 1990 United States census population and housing summary tape file 3A. This file provides annual data on household incomes for 25 income intervals. Counts of the number of households that fall into each income interval along with the total aggregate income and the median household income were obtained for each state. These data were used to calculate the Gini coefficient and the Robin Hood index. The Gini coefficient and the tenths of income distribution used in deriving the Robin Hood index were calculated by using the Gini and income distribution software developed by E Welniak (unpublished software, United States Census Bureau, 1988). This program was developed specifically to be used with aggregate census data to generate Gini coefficients and income distributions.

The file also contains statistics on household size and poverty in which households are classified as being above or below the poverty level based on the revised federal poverty index originally developed by the Social Security Administration in 1964. The current poverty index is based purely on income from wages and does not reflect other sources of income such as non-cash benefits from food stamps, Medicaid, and public housing. Poverty thresholds are updated annually to reflect changes in the consumer price index. The poverty variable we used represents the percentage of households in a given state that were below the federal poverty level. In 1990, this represented an income of less than $13 359 for households with four family members.[9]

All of the data on mortality adjusted for age for 1990 for each state were obtained from the compressed mortality files compiled by the National Center for Health Statistics, Centers for Disease Control and Prevention (CDC). The data were obtained from their database with the CDC WONDER/PC software.[10]

Data on prevalence of smoking in each state were obtained from the centre's smoking information page, available on the Internet. The data are from the current population survey (1989) and reflect the percentage of adults aged 18 years and older who are current smokers.

MEASURES OF INCOME DISTRIBUTION

We examined the relations of total and cause specific mortality with two alternative measures of income distribution: the Robin Hood index[11] and the Gini coefficient. The Robin Hood index was estimated from state specific data on the shares of total household income arranged by tenths of the distribution (see appendix, table A1 for the example of Massachusetts). The index is calculated by taking those groups whose share of the total income exceeds 10% then adding the excess of these shares over that level. In the example of Massachusetts, the value of the index is 30·26%

Department of Health Policy and Management, Harvard School of Public Health, Boston, MA 02115, USA
Bruce P Kennedy, *instructor*
Deborah Prothrow-Stith, *professor*

Department of Health and Social Behavior, Harvard School of Public Health, Boston
Ichiro Kawachi, *assistant professor*

Correspondence to:
Dr Kennedy.

BMJ 1996;312:1004-7

(appendix, table A1). This value approximates the share of total income that has to be taken from those above the mean and transferred to those below the mean to achieve equality in the distribution of incomes.[11] The higher the value of the index the less egalitarian is the distribution of income. The Gini coefficient is another commonly used summary measure of income inequality where higher values indicate greater inequalities in income distribution (see appendix for its derivation).

ALL CAUSE AND CAUSE SPECIFIC MORTALITY

All mortalities were directly standardised for age to the United States population and expressed as the number of deaths per 100 000 (except in the case of infant mortality, where death rates are expressed per 1000 live births). In addition to all cause mortality we examined the following major causes of death: coronary heart disease (defined by codes 410-414 of the *International Classification of Diseases*, 9th revision (ICD-9)); cerebrovascular disease (ICD-9 codes 430-438); malignant neoplasms (ICD-9 codes 140-239); and homicide (ICD-9 codes E960-969).

The association of income inequality with mortality may be partly mediated by lack of access to medical care. We tested this hypothesis indirectly by examining the relations of income inequality to causes of death amenable to medical intervention.[12-14] The specific causes of death examined were infectious and parasitic diseases (ICD-9 codes 001-139); tuberculosis (ICD-9 codes 010-018 and 137); pneumonia and bronchitis (ICD-9 codes 480-486 and 490); and hypertensive disease (ICD-9 codes 401-405).

DATA ANALYSIS

Ordinary least squares regression was used to examine separately the relations of the Robin Hood index and the Gini coefficient to measures of mortality. Two sets of models were examined for each outcome of interest. In the first set of models we adjusted for the proportion of households in each state with incomes below the federal poverty level (which was defined as an annual household income of less than $13 359 for a family of four in 1990). In the second set of models we adjusted for poverty as well as the state specific prevalence of smoking, median household income, and household size. Wherever appropriate, we examined mortality separately by race (white *v* black).

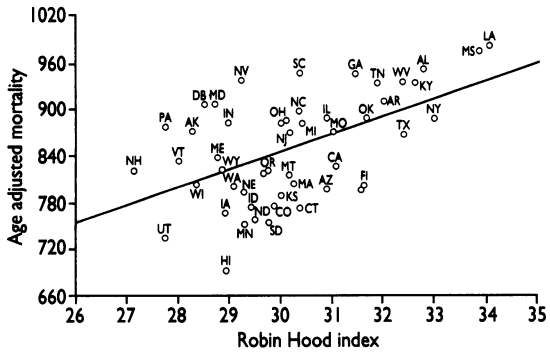

Fig 1—*Mortality by inequality (Robin Hood index) in United States (abbreviations are for each state)*

Table 1—*Effects of Robin Hood index adjusted for poverty*

Cause of death (ICD-9 code)	β (SE)	t (P value)	Adjusted r^2	$F_{2,47}$	P value
Total mortality	21·68 (7·68)	2·82 (0·007)	0·27	9·91	<0·008
Infant mortality	0·45 (0·17)	2·59 (0·013)	0·20	7·31	<0·002
Heart disease (410-414)	9·96 (3·27)	3·06 (0·004)	0·18	6·20	<0·004
Malignant neoplasms (140-239)	5·09 (2·18)	2·34 (0·023)	0·07	2·73	<0·075
Cerebrovascular disease (430-438)	0·77 (1·01)	0·762 (0·449)	0·08	3·06	<0·056
Homicide (960-969)	2·22 (0·38)	5·78 (0·000)	0·52	27·09	<0·000

Table 2—*Effects of Robin Hood index adjusted for poverty and smoking*

Cause of death	β (SE)	t (P value)	Adjusted r^2	$F_{3,46}$	P value
Total mortality	11·83 (6·14)	1·93 (0·060)	0·567	22·36	<0·000
Infant mortality	0·31 (0·17)	1·87 (0·068)	0·32	8·52	<0·000
Heart disease	8·44 (3·33)	2·53 (0·015)	0·21	5·24	<0·003
Malignant neoplasms	2·36 (1·77)	1·33 (0·189)	0·43	13·22	<0·000
Cerebrovascular disease	0·38 (1·04)	0·366 (0·716)	0·09	2·68	<0·058
Homicide	2·11 (0·39)	5·26 (0·000)	0·52	18·60	<0·000

Results

ROBIN HOOD INDEX

The Robin Hood index for the United States overall in 1990 was 30·22% (range 27·13% for New Hampshire to 34·05% for Louisiana). The index had a significant correlation with total mortality adjusted for age (r=0·54; P<0·05) (fig 1). The association of the index to total mortality remained highly significant after adjustment for poverty in our regression model: each percentage increase in the index was associated with an increase in total mortality of 21·68 deaths per 100 000 (95% confidence interval 6·63 to 36·71) (table 1). The bivariate association of the index with total mortality was similar for both black people (r=0·39; P<0·05) and white people (r=0·46; P<0·05). When the effects of poverty in each state were controlled for, the relation of the Robin Hood index to total mortality in black people (β=44·57; 95% confidence interval 12·57 to 76·57) was greater than to mortality in white people (β=15·04; 1·69 to 28·40). Adjustment for median household income and household size in each state did not materially alter these results (data not shown).

Strong associations with the index were also found for infant mortality (P=0·013); coronary heart disease (P=0·004); malignant neoplasms (P=0·023); and homicide (P<0·001) (table 1). In the case of homicide, the index variable alone explained 52·4% of the variance between states. The index was strongly associated with homicide rates among both black people (β=6·51; 2·82 to 10·18) and white people (β=1·81; 1·14 to 2·48).

Adjustment for smoking prevalence in addition to poverty generally attenuated the coefficients for total and cause specific mortality (table 2). As expected, smoking was an independent predictor of total mortality (β=12·37; P<0·0001) as well as deaths from cancer (β=3·42; P<0·0001). The association of the Robin Hood index with total mortality was of borderline significance (β=11·84; P=0·06). The index continued to be a powerful predictor of overall mortality from coronary heart disease (β=8·44; P=0·0148), although the association was confined to white people (β=9·36; P=0·009), the value being β=4·57 (P=0·471) for black people. Less egalitarian states continued to show higher rates of homicide, both among white people (β=1·82; P<0·0001) and black people (β=6·29; P=0·002).

GINI COEFFICIENT

The Gini coefficient for the United States overall was 0·43 in 1990 (range 0·38 for Minnesota to 0·50 for Louisiana). Contrary to previous reports based on comparisons among European nations,[2] the Gini coefficient showed little correlation with any of the mortality outcomes in these data, with the exception of homicide (table 3).

The apparent discrepancy in findings with the Gini and the Robin Hood index was partly accounted for by differences in the meaning of the two measures. The Gini coefficient in these data was sensitive to the income accruing to the extremes of the distribution: a correlation of −0·92 with the proportion of income

Table 3—*Effects of Robin Hood index on treatable causes adjusted for poverty and smoking*

Disease (ICD-9)	β (SE)	t (P value)	Adjusted r^2	$F_{3,46}$	P value
Infectious diseases (001-139)	4·57 (0·79)	5·80 (0·000)	0·39	11·49	<0·000
Hypertensive disease (401-405)	1·32 (0·53)	2·49 (0·016)	0·22	5·52	<0·003
Tuberculosis (010-018, 137)	0·15 (0·04)	4·04 (0·000)	0·44	13·55	<0·000
Pneumonia and bronchitis (480-486, 490)	1·38 (0·48)	2·87 (0·006)	0·10	2·81	<0·050

earned by households in the bottom 10% of the population and 0·93 with the proportion of income earned by the households above the 90th centile. On the other hand, the Gini coefficient correlated only modestly with the proportions of income earned by the bottom 50% and 60% of households (r=0·27 and −0·29, respectively).

The situation with the Robin Hood index was the reverse of that with the Gini coefficient: the index was highly correlated with the proportions of income earned by the bottom 50%, 60%, and 70% of households (r=0·99) but not with the proportion of income earned by the most poor (bottom 10%). The correlation between the Robin Hood index and the Gini coefficient was modest (r=0·29).

TREATABLE CAUSES OF MORTALITY

Strong associations were found between the Robin Hood index and all of the indicators of treatable causes of mortality, which were independent of poverty and prevalence of smoking (table 4). No associations were found between the Gini coefficient and treatable causes of death (data not shown).

Table 4—*Correlations between cause of death and Gini coefficient*

Cause of death	r	P value
Total mortality	−0·03	0·83
Infant mortality	0·06	0·67
Heart disease	0·09	0·52
Malignant neoplasms	−0·19	0·18
Cerebrovascular disease	0·16	0·26
Homicide	0·28	0·04
Infectious diseases	0·08	0·60
Hypertensive disease	0·10	0·47
Tuberculosis	0·24	0·09
Pneumonia and bronchitis	−0·05	0·70

Discussion

The relation between absolute living standards and mortality is well established.[15] The effects of relative deprivation on mortality, however, have been less well understood. Our study extends the findings of previous reports[2,4,5] in showing the association of income inequality (at least as measured by the Robin Hood index) to total and cause specific mortality within one country.

The mechanisms of the association between income inequality and mortality have not been completely elucidated.[7] Although the Robin Hood index correlated with poverty (r=0·73) and smoking (r=0·30), suggesting a potential problem due to multicolinearity, examination of the tolerance statistics and standard errors of the index regression coefficient when the poverty and smoking variables were entered into the model did not indicate that this was a serious problem. Several of the associations with cause specific mortality —in particular, coronary heart disease and homicide— remained significant after adjustment for these variables. We also estimated the regression models by adjusting for median household income and household size (data not shown) with essentially the same results. Although some researchers advocate the use of equivalency scales to take into account differences in household size, such scales ignore the effects of economies of scale. In our analyses there was no evidence that cross state variations in household size were related to mortality or measures of inequality.

Regions with a higher proportion of black residents tend to be overrepresented among the states with a high Robin Hood index (fig 1); none the less, when we stratified the analyses by race the association of the index with outcomes such as homicide remained just as strong for white people as for black people. In the case of coronary heart disease the association of the index with mortality was actually confined to white people.

Income distribution may be a proxy for other social indicators, such as the degree of investment in human capital. Communities that tolerate large degrees of inequality in income may be the same ones that tend to underinvest in social goods such as public education or accessible health care.[16] Our findings with regard to treatable causes of mortality suggest that lack of access to medical care may indeed be part of the mechanism by which income inequality produces higher mortality (although the contribution of treatable causes of death to overall mortality was rather small).

A limitation of the present study is its cross sectional design so that caution must be exercised in the interpretation of the observed associations. Some states may have a high proportion of sick people for reasons other than the hypothesis under investigation, and the less egalitarian distribution of income in such states merely reflects the reduced earning capacity of sick people, who are also at higher risk of dying. Further work should attempt to incorporate time series analyses of income inequality and mortality trends.

Another limitation of the present ecological study is its potential susceptibility to aggregation bias and unknown sources of confounding.[17] To some extent, aggregation of data is unavoidable in studies of this type since the main predictor of interest—namely, income dispersion—is itself an ecological variable. None the less, the ideal study design would incorporate collection of data at the individual level on other predictors of health, including health behaviours (such as smoking and drinking), access to health care, and social class.

CHOICE OF MEASURE OF INCOME INEQUALITY

An unexpected finding of this study was that the choice of the measure of income inequality affected the relation with mortality. Thus strong associations were found between the Robin Hood index and cause specific mortality but not by using the Gini coefficient.

Key messages

- Inequality in the distribution of income explains a significant proportion of the cross state variance in several causes of mortality independent of poverty and smoking
- The size of the gap between the wealthy and less well off—as distinct from the absolute standard of living enjoyed by the poor—seems to be related to mortality
- Policies that deal with the growing inequities in income distribution may have a considerable impact on the health of the population

Previous studies have used different measures of income inequality (such as the Gini coefficient or the proportion of income earned by the bottom 60% of households) without detailed justification for the choice of measure.[18] In practice, it is recognised that there is a wide choice of indices to measure income inequality, but there is no consensus that a single measure, such as the Gini coefficient, ought to be standard.[19] Instead, the selection of the measure of income dispersion should be dictated by the underlying theory of cause of disease.

We found that the Gini coefficient was highly correlated with the proportion of income earned by the bottom 10% of households and hence acts as a proxy for extreme deprivation. By contrast, the Robin Hood index correlated much more with the share of income earned by most of the population. Although the wide choice of inequality indices creates the hazard that researchers will use the measure that proves the result they wish to find, our findings suggest that, at least in the United States, the use of the Gini coefficient may result in more of a test of the effects on health of extreme deprivation rather than relative deprivation. As a measure of income inequality the Robin Hood index has a plausible interpretation. For instance, the findings for mortality from coronary heart disease (adjusted for poverty and smoking) imply that a redistribution of incomes in the United States to achieve a reduction in the Robin Hood index (from 30% to 25%, which is roughly equivalent to the Robin Hood index in England) would be associated on average with about a 25% decline in age adjusted mortality for that disease (from 183 to 139 per 100 000). Furthermore, if there is a causal relation we might expect a reduction in total mortality of 7%.

CONCLUSION

Our findings provide some support for the notion that the size of the gap between the wealthy and less well off—as distinct from the absolute standard of living enjoyed by the poor—matters in its own right. This finding in no way diminishes the importance of measures to alleviate the burden of poverty. None the less, in an affluent society such as the United States, reliance on trickle down policies may not be enough—society must pay attention to the growing gap between the rich and the poor.

Funding: No special funding.
Conflict of interest: None.

Appendix: Derivation of the Gini coefficient

The Gini coefficient is derived from the Lorenz curve, which is a graphic device for representing the cumulative share of the total income accruing to successive income intervals (fig A1). The curve shows the share of income accruing to households in the bottom income interval, then the share going to households in the next income interval (which includes the previous income interval), and so on. If all incomes were equal the Lorenz curve would follow the 45° diagonal. As the degree of inequality increases so does the curvature of the Lorenz curve, and thus the area between the curve and the 45° line becomes larger. The Gini coefficient is calculated as the ratio of the area between the Lorenz curve and the 45° line divided by the whole area below the 45° line.

In figure A1 the Robin Hood index, also known as the Pietra ratio,[20] is equivalent to the maximum vertical distance between the Lorenz curve and the line of equal incomes (CD–CP). A more straightforward derivation of the index can be obtained from the tenths of an income distribution as shown in the example in table A1.

The Robin Hood index may also be calculated by summing the percentage of income for each tenth of an income distribution where the percentage exceeds 10% and subtracting from this the product of the number of tenths that meet this criterion times 10%. In this case four of the tenths (7-10) exceed 10%, so the Robin Hood index

$=(10{\cdot}83\%+13{\cdot}09\%+16{\cdot}41\%+29{\cdot}93\%)-(4\times10\%)$
$=70{\cdot}26\%-40\%$
$=30{\cdot}26\%$.

TableA1—*Data on derivation of Robin Hood index for Massachusetts*

Tenth	Percentage of total income
1	1·08
2	2·48
3	4·13
4	5·74
5	7·33
6	8·97
7	10·83
8	13·09
9	16·41
10	29·93

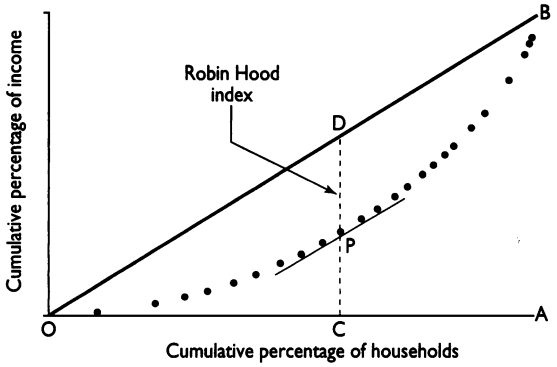

Fig A1—*Derivation of the Robin Hood index from the Lorenz curve and the Gini coefficient*

1 Rodgers GB. Income and inequality as determinants of mortality: an international cross-section analysis. *Population Studies* 1979;33:343-51.
2 Wilkinson RG. Income and mortality. In: Wilkinson RG, ed. *Class and health: research and longitudinal data.* London: Tavistock, 1986.
3 Wilkinson RG. Income distribution and mortality: a "natural" experiment. *Sociology of Health and Illness* 1990;12:391-412.
4 Wilkinson RG. Income distribution and life expectancy. *BMJ* 1992;304: 165-8.
5 LeGrand J. Inequalities in health. Some international comparisons. *European Economic Review* 1987;31:182-91.
6 Kawachi I, Levine S, Miller SM, Lasch K, Amick B. *Income inequality and life expectancy—theory, research, and policy.* Boston: The Health Institute, New England Medical Center, 1994. (Working paper No 94-2.)
7 Wilkinson RG. The epidemiological transition: from material scarcity to social disadvantage? *Daedalus* 1994;123:61-77.
8 Flegg AT. Inequality of income, illiteracy and medical care as determinants of infant mortality in underdeveloped countries. *Population Studies* 1982;36: 441-58.
9 US Bureau of the Census. *CD-ROM, income and poverty.* Washington, DC: US Bureau of the Census, 1993.
10 Friede A, Reid JA Ory HW. CDC WONDER: a comprehensive on-line public health information system of the Centers for Disease Control and Prevention. *Am J Public Health* 1993;83:1289-94.
11 Atkinson AB, Micklewright J. *Economic transformation in Eastern Europe and the distribution of income.* Cambridge: Cambridge University Press, 1992.
12 Charlton JRH, Velez R. Some international comparisons of mortality amenable to medical intervention. *BMJ* 1986;292:295-301.
13 Mackenbach JP, Bouvier-Colle MH, Jougla E. "Avoidable" mortality and health services: a review of aggregate data studies. *J Epidemiol Community Health* 1990;44:106-11.
14 Marshall SW, Kawachi I, Pearce N, Borman B. Social class differences in mortality from diseases amenable to medical intervention in New Zealand. *Int J Epidemiol* 1993;22:255-61.
15 Townsend P, Davidson N. *Inequalities in health: the Black report.* Harmondsworth: Penguin, 1982.
16 Birdsall N, Ross D, Sabot R. *Inequality and growth reconsidered.* Boston: American Economic Association, 1994.
17 Morgenstern H. Uses of ecologic analysis in epidemiologic research. *Am J Public Health* 1982;139:1336-44.
18 Judge K. Income distribution and life expectancy: a critical appraisal. *BMJ* 1995;311:1282-5.
19 Buhmann B, Rainwater L, Schmaus G, Smeeding TM. Equivalence scales, well-being, inequality, and poverty: sensitivity estimates across ten countries using the Luxembourg income study (LIS) database. *Review of Income and Wealth* 1988;34:115-42.
20 Kondor Y. An old-new measure of income inequality. *Econometrica* 1971;39: 1041-2.

(Accepted 24 January 1996)

Correction

Alcohol consumption, serum low density lipoprotein cholesterol concentration, and risk of coronary heart disease

Two editorial errors and a typesetting error occurred in this paper by Hans Ole Hein and colleagues (23 March, pp 736-41). The last sentence in the penultimate paragraph under results (p 739) should have read: "This agrees with Rimm *et al*, who found that the use of strong spirits seemed to be more strongly associated than the use of wine with a decreased [not an increased] risk of ischaemic heart disease." In table 3 the number (percentage) of men with no event should have read 287 (11) [not 287 (4)], and in table 5 the number (percentage) of men who died from all causes and regularly took medicine should have read 127 (48) [not 127 (40)].