BMJ 1998;316:382-385 (31 January)

Education and debate

How much of the relation between population mortality and unequal distribution of income is a statistical artefact?

Hugh Gravelle, *professor* [a]

[a] National Primary Care Research Development Centre, University of York, York YO1 5DD
hg8@york.ac.uk

Introduction

The absolute income hypothesis—that holding other factors constant, the higher an individual's income the better is their health—is supported by a considerable body of evidence.[1 2 3] However, according to the more recent relative income hypothesis, an individual's health is also affected by the distribution of income within society. Someone with a given income would have worse health if he or she lived in a society with greater inequality of income than in a society in which income is more equally distributed.[4] Several recent papers examining the relation between population mortality and income inequality seem to support the relative income hypothesis.[5 6 7 8 9 10 11] They suggest that greater inequality is associated with higher population mortality and that this relation persists even when account is taken of the average income of the population.

However, some scepticism has been expressed about the relative income hypothesis.[12] To quote one of the papers cited above, the "mechanisms underlying the association between income distribution and mortality are poorly understood."[7]

A statistical artefact may explain the relation

There may be a very simple explanation for some, or all, of the reported associations between inequality of income and population health used to support the relative income hypothesis. They may be, at least partly, a statistical artefact caused by using population data rather than individual data. A positive correlation between population mortality and income inequality can arise at aggregate level even if inequality has no effect on the individual risk of mortality. Thus, we do not need the relative income hypothesis to explain the observed associations between population health and income inequality—the absolute income hypothesis will serve.

Mortality risk and absolute income

The absolute income explanation can be illustrated with the help of the figure (the

mortality depends only on the income of the individual, as shown by the heavy line. As income increases, the risk of mortality falls, but it does so at a declining rate. Thus, an increase in income reduces the risk of mortality by a smaller amount at high incomes than at low incomes. Note that this model assumes that there is no relation between the distribution of income and the health of any member of the population. The risk of mortality depends on the absolute income, not the relative income, of the individual.

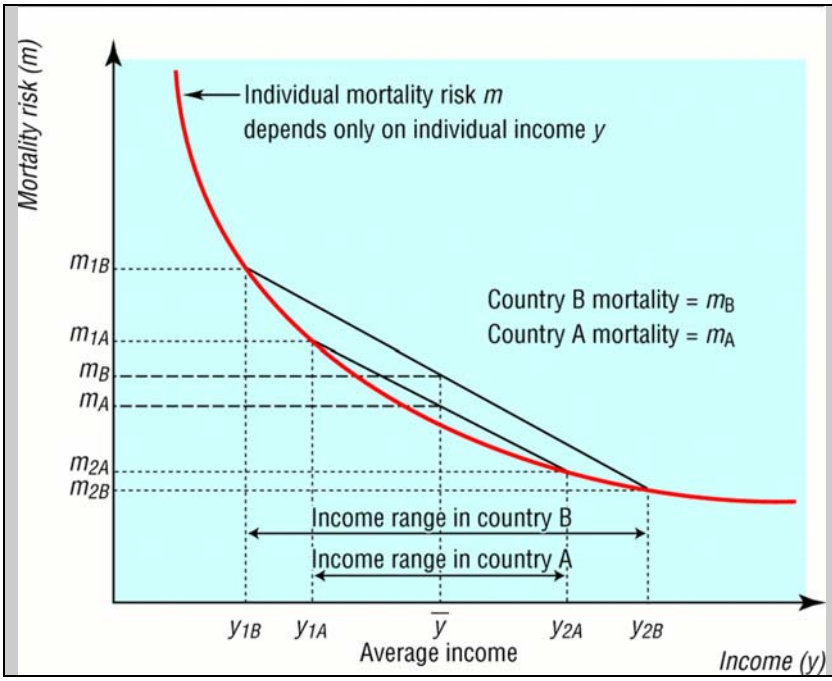
Summary points

The absolute income hypothesis, which states that the higher an individual's income, the lower his or her risk of mortality, is supported by a considerable body of evidence

However, the relative income hypothesis—that the distribution of income in a society affects the individual's risk of mortality—is being used increasingly in empirical work

Associations between unequal income distribution and population health may be a statistical artefact resulting from the use of aggregate rather than individual data—an example of the "ecological fallacy"

Because studies using population data cannot distinguish between the absolute and relative income hypotheses, the effects of income redistribution policies on population health can only be judged from individual data, interpreted by models of behaviours that affect health



Effect of increased inequality of income on population mortality

Mortality risk and inequality income

Now compare two countries where the average income is the same *y* but the distribution of income is different. To avoid cluttering the figure, suppose that in country A half the population has a low income of *y1A* and a high risk of mortality *m1A*. The other half has a higher income of *y2A* and therefore a lower mortality risk of *m2A*. The population mortality in country A is *mA* (the average of *m1A* and *m2A*). In country B income inequality is greater—half the population has an income of *y1B* (and a mortality risk of *m1B*) and the other half an income of *y2B* (and a mortality risk of *m2B*). Although the difference in incomes between rich and poor is greater in country B, it has the same average income as country A. However, population mortality in country B, *mB* (the average of *m1B* and *m2B*), is greater than the population mortality of country A.

Individual mortality and individual risk

The greater population mortality in the country that has a less equal income distribution (country B), results entirely from the shape of the relation between individual income and the individual risk of mortality. The higher income of rich people in country B compared with rich people in country A reduces their risk of mortality by *m2A-m2B* compared with rich people in country A. However, the lower income of poor people in country B compared with poor people in country A increases their risk of mortality by *m1B-m1A* compared with poor people in country A. Because the impact of income on mortality is smaller at higher incomes, the reduced mortality of the rich is more than offset by the increased mortality of the poor and population mortality is therefore higher in country B.

If mortality declines with income, but at a decreasing rate, transferring income from the poor to the rich will increase the mortality risk of the poor more than it reduces the mortality risk of the rich. Overall population mortality increases when inequality increases, even though every individual's risk of mortality depends only on their own income level and not on the income level of anyone else.

Aggregate estimates of data may distort individual risk

In a cross sectional study using population data for different countries or areas, the population mortality will be correlated positively with the degree of income inequality, even if income inequality does not affect the risk of mortality for individuals. The spurious or artefactual correlation at population level between population mortality and income dispersion will always occur if the effect of individual income on the individual risk of mortality is smaller at higher incomes than at lower incomes. This will be so even if there is no underlying relation between the distribution of income and the risk of mortality at the level of the individual.

The "artefactual argument" does not depend on the precise shape of the risk curve for individual mortality, provided only that the risk of mortality declines with income, but at a decreasing rate. Nor does it depend on the shapes of the income distributions or on using a particular measure of income inequality. The artefactual phenomenon can arise in very general circumstances.[13] [14]

The spurious or artefactual correlation at population level between population mortality and income dispersion will always occur if the effect of individual income on the individual risk of mortality is smaller at higher incomes than at lower incomes. This will be so even if there is no underlying relation between the distribution of income and the risk of mortality at the level of the individual.

Income and health—a non-linear relation?

Income affects health because it influences individuals' consumption of commodities that affect health. Housing and tobacco are two obvious examples. The relation between income and health depends on the relation between income and the commodities that affect health, and on the relation between these commodities and health. What matters for the artefactual argument is that the resulting relation between health and income is not a straight line—it produces a curve like that in the figure. This could result from non-linear (or J shaped) relations between health and commodities or environmental factors that affect health,[15] or it could arise from non-linear associations between income commodities or environmental factors that affect health.[16 17]

Evidence for the relative income hypothesis

The hypothesis that mortality is affected more by relative than absolute living standards has been supported by various types of evidence.[10] Several studies comparing different countries show that population mortality and income inequality are positively associated, even after differences in average income and other socioeconomic factors have been allowed for. But this is precisely what is predicted by the artefactual argument. The sheer volume of studies reporting population level correlations of health and inequality measures cannot in itself support either the relative income hypothesis or the artefactual argument.

Income and mortality are more closely related within countries

The second type of evidence used to support the hypothesis that income distribution affects directly the individual risk of mortality is the regular gradient between income and mortality within countries, which contrasts sharply with the weak relation between countries. This, it is argued, is because income differences between groups within a society are associated with social stratification (and its detrimental effect on health) while per capita income differences between countries are not. However, if the artefactual argument is correct, we would expect to find that the income-mortality correlation within countries is stronger than the average income-mortality correlation between countries. The within country correlations would, in effect, be plotting the relation between individual mortality risk and individual income shown in the heavy line in the figure, provided the data for a single country were derived from the level of the individual or were aggregated across relatively homogeneous income groups. The correlation of population mortality against average income between countries would be confounded by the omission of measures of the variability of income in the different countries, resulting in a weaker correlation.

Marginal effect of income is smaller in richer countries

The third argument is that in more developed (richer) societies the marginal effect of income on mortality is smaller, and the curve relating income to mortality is flatter. But this is what is required for the artefactual hypothesis to hold. If the artefactual hypothesis were an invalid explanation of the association between inequality and population mortality, the income-mortality curve would have to be a straight line.[4]

More general artefactual problems

Hypotheses about mechanisms

The artefactual problem can also arise in attempts to test hypotheses about the mechanism by which greater inequality affects individual health. For example, increased inequality might lead to greater stress in individuals, to greater alcohol consumption, and thus to worse health outcomes. One seemingly obvious way of testing this argument is to examine the relation between average alcohol consumption and income inequality. However, studies suggest that the relation between an individual's demand for alcohol and their income is non-linear—increases in income lead to greater than proportional increases in alcohol consumption.[15] If this is the case, an increase in income inequality will lead to an increase in average alcohol consumption because the increased consumption of the rich will more than outweigh the reduced consumption of the poor. Thus, any correlation between average alcohol consumption and income inequality would be at least partly artefactual, arising from the non-linearity of the relation at the level of the individual. Inequality could have no true effect on individual consumption.

Problems of non-linear relations at individual level

Similar problems arise in examining the relation between inequality of income and health problems associated with alcohol. Studies suggest that the relation between alcohol consumption and health effects is non-linear—the deleterious effects increase more than proportionately with consumption, and consumption may even be protective at low levels.[16] This will be sufficient to generate a non-linear relation between income and alcohol related problems at the level of the individual, and therefore artefactual correlations between income inequality and alcohol related problems in the population.

The ecological fallacy

The artefactual argument is another example of the difficulties—known in epidemiology texts as the "ecological fallacy"—in inferring relations at the individual level from associations between variables at the population level.[18] [19] The problem has been discussed in econometric reports on consumer behaviour, where it is known as the aggregation problem.[20] It was also noted in early work on income and health,[1] and was emphasised in Rodgers's pioneering paper, to which reference is made in many recent reports on the link between mortality and income distribution.[21] I argue here that the problem is more general and therefore potentially more troubling for empirical work.

Implications of the artefactual argument

Since the artefactual argument casts doubt on the methods used and the interpretation of a large and growing body of empirical work, being clear about what is and what is not being suggested is important. I do not suggest that an individual's health is not affected by the overall distribution of income as well as by their own income. The point I am making is that correlations between population level measures of mortality and inequality provide biased estimates of the importance of any relative income effect, even after allowing for the levels of other potentially important determinants of mortality. If the relation between the individual risk of mortality and individual income is non-linear, at least part and possibly all of the correlation will be artefactual.

Both the absolute income and relative income hypotheses predict that a reduction in inequality of income can improve the health of a population. However, if policies that alter the distribution of income are to be judged at least partly by their effects on population health, knowing how large these effects are is important. Studies using population level data do not answer this question since they can not distinguish between the absolute income and relative income hypotheses. Individual data are necessary. Equally important are models of the behaviours that affect the health of individuals so that these data can be interpreted appropriately.[22] [23] [24]

Acknowledgements

I thank R Carr-Hill, G Davey Smith, T Sheldon, M Sutton, and R G Wilkinson for their comments and suggestions.

Funding: Support from the from the Department of Health to the National Primary Care Research and Development Centre is acknowledged. The views expressed are those of the author and not necessarily those of the Department of Health.

Conflict of interest: None.

Appendix: The mathematical argument

Suppose that the mortality risk of an individual with income y depends only on their income and is $m(y)$. Taking a second order approximation, we can express the mortality risk of an individual in terms of the individual's income y and the mean income $\bar{y}$ of the population where m' and m'' are the first and second derivatives of m.

$$m(y) \approx m(\bar{y}) + m'(\bar{y})(y-\bar{y}) + \tfrac{1}{2}m''(\bar{y})(y-\bar{y})^2$$

The mortality of the population as a whole is found by multiplying the risk of mortality for individuals with income y by their relative frequency in the population and summing over all income groups. If this operation is performed on both sides of the above expression, remembering that by definition the average deviation of incomes is zero, so that the second term disappears, the following equation is derived:

$$Em(y) \approx m(\bar{y}) + m'(\bar{y})E(y-\bar{y}) + \tfrac{1}{2}m''(\bar{y})E(y-\bar{y})^2 = m(\bar{y}) + \tfrac{1}{2}m''(\bar{y})\,Var(y)$$

where $Em(y)$ is the population mortality and $Var(y)$ is the variance of incomes across the population. Hence, provided only that $m''>0$, so that individual mortality risk declines with individual income at a decreasing rate (as in the 1), the population mortality will be positively correlated with mean income and the variance of incomes.

References

1. Preston SH. The changing relation between mortality and level of economic development. *Popul Stud* 1975;29:231-48.

2. Pritchett L, Summers LH. Wealthier is healthier. *J Hum Resources* 1996;31:841-68.

3. Adler NE, Boyce W, Chesney MH, Folkman S, Fymes L. Socioeconomic inequalities in health: no easy solution. *JAMA* 1993;269:3140-5.

4. Wilkinson RG. *Unhealthy societies: the afflictions of inequality*. London: Routledge, 1996.

5. Ben-Shlomo Y, White IR, Marmot M. Does the variation in the socioeconomic characteristics of an area affect mortality? *BMJ* 1996;312:1013-4.

6. Kaplan GA, Parnuk ER, Lynch JW, Cohen RD, Balfour JG. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *BMJ* 1996;312:999-1003.

7. Kennedy BP, Kawachi I, Prothrow-Smith D. Income distribution and mortality: cross sectional ecological study of the Robin Hood index in the United States. *BMJ* 1996;312:1004-7.

8. Wilkinson RG. Income distribution and life expectancy. *BMJ* 1994;304:165-8.

9. Wilkinson RG. A reply to Ken Judge: mistaken criticisms ignore overwhelming evidence. *BMJ* 1995;311:1285-7.

10. Wilkinson RG. Health inequalities: relative or absolute material standards? *BMJ* 1997;314:591-5.

11. Waldman RJ. Income distribution and infant mortality. *Q J Econ* 1992;107:1283-302.

12. Davey Smith G. Income inequality and mortality: why are they related? *BMJ* 1996;312:987-8.

13. Atkinson AB. On the measurement of income inequality. *J Econ Theory* 1970;2:244-63.

14. Rothschild M, Stiglitz JE. Increasing risk I: a definition. *J Econ Theory* 1970;2:225-43.

15. Godfrey C. Factors influencing the consumption of alcohol and tobacco: the use end abuse of economic models. *Br J Addict* 1989;10:1123-38.

16. Edwards G, Babor TF, Anderson P. *Alcohol policy and the public good*. Oxford: Oxford University Press, 1994.

17. Atkinson AB, Gomulaka J, Stern NH. Spending on alcohol: evidence from the family expenditure survey. *Econ J* 1990;100:808-27.

18. Robinson WS. Ecological correlations and the behaviour of individuals. *Am Soc Rev* 1950;15:351-7.

19. Beaglehole R, Bonita R, Kjellstrom T. *Basic epidemiology*. Geneva: World Health Organisation, 1993.

20. Deaton A, Muellbauer J. *Economics of consumer behaviour*. Cambridge: Cambridge University Press, 1980.

21. Rodgers GB. Income and inequality as determinants of mortality: an international cross-section analysis. *Popul Stud* 1979;33:343-51.

22. **Gravelle** HSE. Time series analysis of mortality and unemployment. *J Health Econ* 1984;3:297-305.

23. **Gravelle** HSE, Backhouse M. International cross-section analyses of the determinants of mortality. *Soc Sci Med* 1987;25:427-41.

24. Ettner SL. New evidence on the relationship between income and health. *J Health Econ* 1996;15:67-86.

(Accepted 21 July 1997)