# An explanation of the persistent doctor-mortality association

F W Young

**Abstract**
*Study objective*—The aim of the study is to explain the persistent but puzzling positive correlation of physicians per capita and mortality rates, when income is controlled, which has been reported many times since it was first observed in 1978. The explanation that is proposed and tested is that expanding urban-industrial regions attract an oversupply of doctors. Also, but independently, rural people migrate to urban-industrial areas where they suffer from the stress of adapting to urban-industrial life. Consequently, their death rates rise.
*Source material*—Using data from the 47 Japanese prefectures, the 3000+ counties of the USA and a set of 29 mostly European countries, the explanation was examined by adding the appropriate test variable to a basic equation linking physicians per capita to mortality, net of income.
*Results*—The test variables dissolved or reduced the original correlation in two of the three samples, but the signs did not change from positive to negative, as would be expected on the basis of conventional biomedical theory. The available test variable (refugees) did not reduce the correlation for the 29 countries but a particular subset of countries was identified that did.
*Conclusion*—The conceptual and empirical analysis exposed the positive correlation as spurious, but the availability of medical specialists had little impact on mortality rates in competition with the social and economic variables that were used as controls.

(*J Epidemiol Community Health* 2001;55:80–84)

In 1978, Cochrane *et al*[1] called our attention to what has turned out to be a persistent positive association between physicians per capita and death rates. Their article also examined environmental and dietary factors and these, together with GNP per capita, served as controls for the 18 countries in their sample. Although they did not doubt that a causal interpretation was unjustified, their tests demonstrated a robust correlation. Accordingly, they called it "the anomaly which will not go away".[2]

Since then, this unexpected correlation has been replicated and expanded in a number of studies. These have distinguished the adult mortality rate from the infant mortality rate that Cochrane *et al* emphasised and they have

also looked at a wider range of medical facilities, including hospitals.

Several studies analysed data for the less developed countries circa 1970–80. Wimberly[3] found that change in health expenditures and in physicians per 100 000 population made no significant contribution to the prediction of life expectancy. In the equation for infant mortality, physicians made a significant but ("unexpected") positive contribution, net of GNP per capita, school enrolment, and two economic dependency measures.

A review of studies of "avoidable" mortality and health services by Mackenbach *et al*[4] concluded that the associations of health care variables and mortality were "weak and inconsistent" (page 110). The three studies that disaggregated the correlations found that half of the 16 significant associations were positive—that is, the more health care, the higher the mortality rates. Similar findings have been reported by Matteson *et al*,[5] McKinlay and McKinlay[6] and Jayachandran and Jarvis.[7] Going back in time, Williams[8] has shown that the migration of peasants to the industrialising Glasgow region from about 1820 to 1870 had a marked negative impact on death rates attributable to urban poverty and the destruction of the agrarian economy that they left. Given his purposes, he did not report the availability of physicians.

Clarke *et al*[9] found no significant association in US counties between an index of primary care (physicians in general practice, family physicians, internists, paediatricians, obstetricians/gynaecologists and nurse practitioners) and infant mortality when they controlled on a range of social structural factors, including average income. Chen and Lowenstein[10] found the expected negative correlation between physicians and infant mortality in their sample of 60 developing countries, but they did not control on GDP, so the association may not be final.

Sankrithi *et al*[11] have shown that for a sample of 129 countries, using an exponential multivariate model and controlling on "spurious correlation" reverses the positive physician-infant mortality correlation and brings it into line with expectations. In other words, the original positive correlation was a statistical artefact and there is no spurious correlation to explain.

I have not followed their implicit advice for a number of reasons. Firstly, they do not give a substantive rationale for preferring their exponential regression format. As for bringing the correlation into line with expectations, these are theoretical expectations on which

**Department of Rural Sociology, Cornell University, Ithaca, NY 14853, USA**

Correspondence to:
Dr Young
(Fwy1@cornell.edu)

reasonable people differ. For some, it is just as reasonable to expect no correlation, one way or the other. Meanwhile, the reason everyone uses linear regression applies in this case. It is the simplest model and therefore the best starting point.

Cochrane *et al* looked for but found no convincing explanation. They left it to others "to extricate doctors from this unhappy position" (page 204). This article takes up their challenge. Even if no one believes that the positive correlation reflects causality, its existence is an embarrassment for the role of medical interventions in the biomedical explanation of health. Also, the exercise of demonstrating spuriousness should throw light on the role of medical facilities in comparison with the social and economic variables.

### A new explanation
The explanation proposed here is that, as compared with other regions, the expectation of opportunities in the growing industrial cities initially attracts an over supply of doctors. Once in practice, doctors in new regions enjoy fewer economies of scale, which means that they are more numerous as compared with the mature regions. These same industrialising cities attract rural immigrants whose health habits and supports break down in the context of city life. Thus, the places with the most doctors also have the highest death rates, but the two variables are associated only by common location.

The "industrialising context" explanation for the doctor-mortality correlation is general because such regions can be found in all parts of the world at one time or another. Still, it is epoch specific, because it assumes the existence of medical practices, particularly obstetrics and accident therapies, which are capable of decreasing the death rate. What it lacks, and what is necessary for an adequate research design, is a specification of the types of rural immigrants. These may be classified as follows:

*Accepted immigration*—from the point of view of the host city or country, some types of immigration are accepted and even considered desirable. Rural people who are of the same race or ethnic group as those in the cities tend to be welcomed, even if they arrive from other countries. The term "accepted" acknowledges this positive response but also implies that the newcomers will be greeted with a certain amount of resentment and resistance.

*Limited immigration*—the best example of the limited type is the "guest worker" category in European cities where the newcomers have certain rights, but citizenship is not one of them. This category also includes de facto guest workers such as the Mexican migrants who cross the southern US border illegally and then manage to stay and settle. Also included are refugees from other countries, many of whom function as guest workers until, or if, they acquire citizenship.

*Segregated immigration*—the principal example here is slavery in the American South and the continued segregation of black people even after the abolition of slavery. The mechanism of ethnic segregation reflects the ambivalence of the host country in that the newcomers have residence and even citizenship rights, but their presence is openly resented and actively circumscribed. People of African descent seem to be special targets of this type of discrimination, although the process is not unknown for other ethnic groups.

### Design of research and tests of hypotheses
The typology of industrial worker migration should improve the precision of tests of the industrial context hypothesis, but it also complicates the design because it is unlikely that one set of countries or provinces is adequate for examining all three types. As a consequence, the tests presented here used three samples (actually, universes).

The first consists of the 47 Japanese prefectures as of 1990, when the cross sectional analysis captures the impact of accepted immigration during the postwar period.

At the other end of the assimilation continuum, ethnic segregation is explored with the data from the 3000+ counties of the US as of 1980–1990. The middle type of limited assimilation is represented by a sample of mostly European countries as of 1990. The sample includes all but one of those used by Cochrane *et al* but, of course, the date is 20 years later. Although guest workers and refugees are still arriving in European and other advanced industrial countries, it is reasonable to suppose that the process has improved since their 1970 study. On the other hand, the well known hostility and anti-immigrant legislation are evidence that the process is far from the "accepted" category.

All of these datasets are cross sectional and they consist of countries and administrative units within countries. Happily, the ecological fallacy is not a threat in this case because the aim of the research is to dissolve an ecological correlation between two rates and explain why a causal interpretation is spurious. The ecological fallacy becomes a threat only if the analysis turns up a negative doctor-mortality correlation, which then tempts us to claim that doctors caused the reduced mortality. As we shall see, such a correlation occasionally appears, but then we must rely on strong theory to overcome the threat.

The analysis strategy should now be apparent. The first step is to replicate the positive doctor-mortality correlation, controlling on income per capita and any other relevant variables. Then a test variable—a measure of migration to industrialising regions—is added to the equation with the aim of dissolving or at least reducing the original correlation. With small samples, statistical significance may not be attainable, but the pattern of changes should conform to the accepted format for unmasking a spurious correlation. Regression analysis shows this dissolution process and it also shows the contribution of the control variables, the understanding of which is important in the regional industrialising process.

*Table 1   Regression analysis of age and sex standardised mortality in Japanese prefectures (n=47)*

| Predictors | *Male* | *Mortality* | *Female* | *Mortality* |
|---|---|---|---|---|
|  | 1 | 2 | 1 | 2 |
| Income p/c | −0.16 | 0.12 | 0.26 | 0.52★ |
| Physicians per 100 000 | 0.14 | 0.00 | −0.11 | −0.24 |
| Industrial area |  | 0.51★ |  | 0.49★ |
| $r^2$ | 0.02 | 0.14 | 0.06 | 0.17 |

Numbers are standardised regression coefficients. ★ = Significant at the 0.05 level or better.
Standard errors are dropped because this is the universe and because it is the shifts in the strength of the coefficients that are paramount.
Mortality rate, age adjusted male, 1995 (mean: 713; min–max: 618–774).
Mortality rate, age adjusted female, 1995 (mean: 377; min–max: 323–439).
Income p/c, Income per capita in 000s. (2863: 2149–4255).
Physicians per 100 000 population. (185; 109–350).
Industrial area. The seven prefectures in the far north east and all prefectures south west of Osaka prefecture. (27 prefectures, or 57 per cent).
Sources: Japan Statistical Yearbook 1998 and Kensei 1999 (income per capita); Ministry of Health and Welfare, 1998 (age standardised mortality).

### Accepted immigration in Japan

Japan since 1950 may be classified as a case of "Accepted" rural-urban migration from the countryside to the growing industrial cities in the south west and north east. The sample is limited to the 47 Japanese prefectures, but the quality of the data is high. Table 1 cuts directly to the question at hand, postponing a detailed description of the variables. Thus, male age adjusted mortality is negatively related to income per capita but positively related to physicians per 100 000, although neither is significant. However, when the dichotomy of mature versus growing regions is introduced in the second column, the regression coefficient for physicians drops to zero and the coefficient for the test region is significantly positive. In other words, the peripheral urban context accounts for the initial positive regression coefficient. In the context of the universe of prefectures, this shift is substantively significant.

Equations 3 and 4, where age adjusted mortality for women is the criterion, pose a new problem, because the signs for these regression coefficients are the reverse of those for men. Although non-significant at the 0.05 level, the coefficient for income per capita is strong, while the coefficient for physicians is negative. When the industrialising context variable is added, it is significantly positive as before and the negative coefficient for physicians per capita increases. That is, the greater the number of physicians, the lower the female mortality rate.

What about infant mortality? When this criterion was inserted in the equations in table 1 no significant coefficients or shifts in their size appeared. The probable reason for this is that in Japan the mean for infant mortality is quite low (4.4) with a range in 1995 from 3.3 to 5.9. In short, this criterion manifests so little variation that statistical analysis is undercut.

The descriptive statistics for the five variables are presented at the foot of the table. The age adjusted mortality rate for men is almost twice as high as that for women and it is higher in the industrialising region (725) than in the heartland (697). In contrast, the rates for women in the two regions are almost identical (378 versus 376). Income per capita is straightforward although the averages for men and women, if we had them, are probably quite disparate. Physicians per capita averages are 168 in the mature region and 199 in the growth area.

The classification of the expanding industrial regions is based on descriptions in Kornhauser[12] (page 125) but the fact that the proportion of large cities (100 000 and over) is greater in the mature region than in the south (35% versus 24%) also supports the interpretation.

### Ethnic segregation in US counties

The data on the US counties permit the analysis of an extreme type of "internal exclusion", all the more pernicious because it applies to long term residents. Although the argument is basically the same, it assumes a new form. Beginning about 1940, African Americans migrated to the cities in the South and elsewhere in the US, but now, instead of a temporary handicap until they could adapt to urban life, they had to contend with permanent ethnic discrimination.

The sample consists of all the counties in the contiguous US with the exception of the Virginia cities. These and the counties of Hawaii and Alaska were omitted in order to reduce the heterogeneity of the sample. Although the dataset was quite complete,[13] a number of data deficient counties were also omitted.

There are many ways to measure ethnic discrimination, but with the data available in this study, a factor analysis of the organisational characteristics of the counties seemed to be the best way to proceed. The analysis (not shown) generated four factors but it is the Ethnic Dualisation factor that concerns us here. The variables and their loadings are as follows: per cent of population that is black (0.77), county located in South (0.70), murders per 100 000 (0.66), violent crimes per 100 000 (0.63), per cent voting in the 1988 election (−0.55) and per cent change in proportion of high school graduates (0.49).

The per cent black in 1990 and location in the South defines the region, and the rates for murder and violent crimes reflect a polarised social organisation. The per cent that vote loads negatively, reflecting the political apathy in segregated counties. Finally, the per cent change in the proportion of the eligible population with a high school education shows a positive loading, but it is unclear whether it is the whites or blacks or both who are contributing most strongly to the rise in the educational level. The factor would be improved if we had an indicator of the white oligarchy that probably controls most of these counties.

This factor is weak in comparison with the others because it explains only 8.2% of the variance (out of a total of 65%), but that is probably because of its restriction to one region of the country. That fact reminds us that there are many other types of ethnic discrimination in other parts of the US, but black-white polarisation is surely the most notorious.

We are now in a position to examine the regression analyses of infant mortality as shown in table 2. The variables used in the

*Table 2 Regression analysis of infant mortality in US counties, 1990*

| Predictor | 1 | 2 | 3 |
| --- | --- | --- | --- |
| Median income | −0.20★ | −0.10★ | −0.09★ |
| Physicians per 1000 (logged) | 0.11★ | 0.04 | 0.02 |
| Hospital beds per 100 000 | −0.07★ | −0.05★ | |
| Health expenditures (logged) | 0.08★ | 0.09★ | |
| Ethnic dualisation | | 0.26★ | 0.26★ |
| $r^2$ | 0.04 | 0.10 | 0.08 |

All numbers are standardised regression. ★ = Coefficients significant at the 0.05 level or better.
Infant mortality per 1000 live births, averaged over 1991–93 (mean: 7; range: 4–21).
Median family income (28337; 10903–65201).
Physicians per 1000 people (1.1; 0–21.1).
Hospital beds per 100 000, (380.8; 0–4945).
Health expenditures as a per cent of total budget (1.7; 0–9.3).
Ethnic dualisation factor score of ethnic dualisation. See text. (0; −2.2–5.5).
Sources: US Bureau of the Census. 1991 (Census); 1992a (CO-STAT4) and 1992b (County Business Patterns).

*Table 3 Regression analysis of age adjusted mortality in US counties, 1990*

| Predictor | 1 | 2 | 3 |
| --- | --- | --- | --- |
| Median income | −0.36★ | −0.17★ | −0.16★ |
| Physicians per 1000 (logged) | 0.22★ | 0.07★ | 0.04★ |
| Hospital beds per 100 000 | −0.07★ | −0.05★ | |
| Health expenditures (logged) | 0.03 | 0.04★ | |
| Ethnic dualisation | | 0.54★ | 0.54★ |
| $r^2$ | 0.11 | 0.36 | 0.36 |

All numbers are standardised regression. ★ = Coefficients significant at the 0.05 level or better. Age standardised morality, deaths per 100 000, averaged over three years (1989–91), and standardised on the base of the national age structure.

KEY POINTS
● The positive association of doctors per capita and age standardised mortality has been found repeatedly since this "anomalous" correlation was first noticed in 1978.
● The correlation is spurious and arises from the dynamics of newly industrial regions. Mortality increases among rural people who take the new jobs, while an oversupply of doctors migrates to the region in search of opportunities. The two groups are connected only by common location.
● Variants of this process can be found in regions all over the world.
● Controlling for type of region reduces the doctor-mortality association for men, often to zero; but it does not reverse as medical advocates might expect. This outcome raises the question of the role of modern medicine in industrialising regions.

tables are all well known and are described at the foot of the table. Equation 1 shows the relation of median family income in 1989 and three measures of medical facilities to infant mortality averaged over three years, 1990–1993. Income is negatively related, as always, and physicians per capita is positively correlated, as we have come to expect. Health expenditures per capita also shows a positive relation, but hospital beds is negatively related. The working hypothesis that covers these facts is that the two indicators that are easily concentrated, physicians and health expenditures, respond to market forces in the cities, in contrast with hospital beds that are usually administratively allocated.

Equation 2 introduces the test variable, Ethnic Dualisation, and it is strong, positive and significant. Its introduction reduces the physician coefficient to non-significance. In contrast with the working hypothesis, it does not reduce the coefficient for health expenditures.

Equation 3 shows only the three basic variables in order to eliminate any collinearity that the other variables might have introduced. Evidently, their presence did not affect the contribution of physicians, because that coefficient remains non-significant.

Table 3 shows the same three equations, but this time the dependent variable is age adjusted mortality. Physicians per capita is strong, significant and positive. In contrast, health expenditures is non-significant, again contradicting the working hypothesis that they tend to concentrate.

The addition of Ethnic Dualisation in equation 3 has the hypothesised effect. The coefficient for Dualisation is quite strong and it sharply reduces the coefficient for physicians, even though the latter variable remains significant. When the two other measures of medical facilities are deleted, as shown in equation 3, the coefficient for physicians further decreases, but it is still marginally significant.

## Allowed immigration in high income countries

Although the "guest worker" type falls between the two extremes, it is considered last in order to focus the whole analysis on the findings first generated by Cochrane *et al*. These authors drew their sample of 18 countries according to three criteria: a per capita income of $2000 per capita in 1970, population or more than two million, and a heterogeneous population. The last criterion eliminated Japan and the grounds of genetic homogeneity. Using 17 of their 18 countries (Scotland omitted), I replicated their result (not shown) for 1990. Controlling for GNP per capita, the regression coefficient for men was 0.18 while that for women was 0.05. Neither of these associations was significant, and this fact underlined the need for a larger sample.

As it turned out, the Human Development Report[14] contained most of the data necessary for this test, and the countries in the top section of the first table, which ranks countries according to their Human Development Index score that included GNP, were used for this study. Of the 45 countries in the top category, 29 had more than two million population and these included the 17 countries in the original sample. The criterion of genetic heterogeneity was dropped as irrelevant to the urban context hypothesis.

Unfortunately, the HDR did not report data for the one variable that was crucial, the number of guest workers per capita. An attempt to substitute the number of economic and political refugees for the crucial test variable was unsuccessful (in fact, it increased the positive doctor-mortality correlation.)

This unsatisfactory outcome prompted a series of exploratory probes, one of which uncovered a lower correlation between doctors

and mortality for the less affluent half of the OECD countries. When this fact was converted to a dummy variable and entered into the equation, the coefficient for physicians decreased (not shown). With only 29 countries, the original correlation (0.13) was not significant, but it did reduce to 0.02 when the dichotomous variable was added.

What is it about these particular OECD countries that reduced the positive association for physicians? The countries in question are Canada, Australia, United Kingdom, Finland, Italy, Ireland, Spain and Portugal. By definition they are the less affluent OECD countries, but they are still well off. Other than that, however, a search (among the available variables) uncovered no distinguishing attribute. What we are left with then is the fact that a particular block of countries contains a variable that dissolves the doctor-mortality association. It is possible that these countries contain expanding industrial regions, but that could not be demonstrated from the data.

### Conclusion and Discussion

The research reported here confirmed the existence of the doctor-mortality correlation in all three samples and it was possible to dissolve the initial correlation in the Japanese prefectures and the US counties by introducing an appropriate test variable. The mostly European cross national sample did not respond to the proxy test variable but a category of countries was identified that reduced the initial correlation.

Thus, the doctor-death correlation is shown to be spurious, a conclusion that was never in doubt, but in the process of doing so the regional social context for both death rates and the medical personnel who practise in the urban-industrial regions was clarified. Although both types of newcomers were responding to market forces in the growing cities, their activities were not articulated.

It is always possible that someone will find a causal explanation for the doctor-mortality association—germs in doctor's offices, for example—but the evidence presented here argues for a regional context even for this type of connection. A more reasonable expectation is a better understanding of the "economic" aspects of the process. For example, the presence of diseases that are not curable yet require multiple interventions could account for the oversupply of doctors.

A larger question that is posed here has been raised earlier by Cochrane[15] and McKeown.[16] See also Alvarez-Dardet and Ruiz.[17] That question pertains to the efficiency of modern medicine. What are we to make of the fact that in all three samples controlling on the test variable reduced the positive correlations for men, but did not reverse them? If doctors were in fact curing the workers, a lowered death rate should reflect it. Or should it? Medical interventions are weak causes compared with the general level of development. They are also weak, it seems, as compared with the stress of working in an industrialising region. True, the death rate of women in Japan seems to respond to medical interventions (in relation to childbirth?) but the two structural variables were almost twice as strong, and both of them increased the female death rate.

Thus, the explanation of a spurious correlation raises deeper questions. Unlike the initial question, which was minor and almost amusing, the new questions are fundamental. They call attention to the effect of market forces in the delivery of medical care and the powerful effects of structure. Now, in addition to economic development (Japanese women the exception) on health, we must take the negative impact of industrialising regions into account.

1 Cochrane AI, St Leger AS, Moore FJ. Health service "input" and mortality "output" in developed countries. *J Epidemiol Community Health* 1978;**32**:200–5.
2 St Leger AS, Cochrane AL, Moore F. The anomaly that wouldn't go away. *Lancet* 1978;ii::1153.
3 Wimberly DW. Investment dependence and alternative explanations of third world mortality: a cross-national study. *American Sociological Review* 1990;**55**:75–91.
4 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.
5 Matteson DW, Burr JA, Marshall JR. Infant mortality: a multi-level analysis of individual and community risk factors. *Soc Sci Med* 1998 **47**:1841–54.
6 McKinlay JB, McKinlay SM. The questionable contribution of medical measures to decline of mortality in the United States in the twentieth century. *Milbank Memorial Fund Quarterly* 1977; **55**:405–28.
7 Jayachandran J, Jarvis GK. Socioeconomic development, medical care and nutrition as determinants of infant mortality in less developed countries. *Soc Biol* 1986;**33**:301–15.
8 Williams R. The health costs of Britain's industrialization: a perspective from the Celtic periphery. In Platt S, Thomas H, Scott S, *et al*, eds. *Locating health.* Brookfield, VT: Ashgate Publishing, 1993:179–204.
9 Clarke LL, Farmer FL, Miller MK Structural determinants of infant mortality in metropolitan and nonmetropolitan America. *Rural Sociology* 1994;**59**:84–99.
10 Chen MK, Lowenstein F. The physician/population ratio as a proxy measure of the adequacy of health care. *Int J Epidemiol* 1985;**14**:300–3.
11 Sankrithi U, Emanuel I, Van Belle G. Comparison of linear and exponential multivariate models for explaining national infant and child mortality. *Int J Epidemiol* 1991;**20**: 565–70.
12 Kornhauser DH. *Japan, geographic background to urban-industrial development.* New York: Longman, 1982.
13 Young FW, Lyson TA. Structural pluralism and all-cause mortality. *Am J Public Health* (in press).
14 United Nations Development Programme. *Human development report 1999.* New York: Oxford University Press. 1999.
15 Cochrane AL. Effectiveness and efficiency: random reflections on health services. London: Nuffield Provincial Hospital Trust, 1972.
16 McKeown T. The role of medicine: dream, mirage or nemesis. London: Nuffield Provincial Hospital Trust, 1976.
17 Alvarez-Dardet C, Ruiz, MT. Thomas McKeown and Archibald Cochrane: a journey through the diffusion of their ideas. *BMJ* 1993;**306**:1252–5.