STATISTICAL STUDIES OF THE EFFECT OF LOW LEVEL RADIATION FROM NUCLEAR REACTORS ON HUMAN HEALTH

MORRIS H. DeGROOT
Carnegie-Mellon University

1. Possible effects of nuclear reactors

Government policy with regard to the construction and operation of nuclear power plants is of great public concern, not only because of the possibility of a serious accident at one of these plants, but also because of the possibility that radioactive discharges from these plants during their routine operation may affect the health of nearby populations. In particular, because of the vulnerability of the human fetus, it is possible that exposure of a population to these discharges may be reflected in the infant mortality rate, the fetal death rate, the prematurity rate, and similar health indices of the population.

Since several nuclear reactors have been in operation in the United States for at least five years, and some for more than ten years, the relevant data for a statistical study of this problem are largely available in published records. A study of this type would necessarily be retrospective in nature and confined to short term effects of low level radiation. If these effects are discernible, then they should be reflected in certain relationships between the health indices mentioned above for a given population and various measures of radioactivity in the environment.

2. Populations to be considered

Annual infant and fetal mortality rates, as well as prematurity rates, are typically available on a county by county basis in the published vital statistics of each state. It is suggested for simplicity, therefore, that counties form the basic units of population to be considered. Thus, for a given reactor, annual health indices for the county containing the reactor and for nearby counties would be investigated over a period both before and after the reactor became critical for possible relations with measures of the total annual radioactive

This research was supported in part by the National Science Foundation under grant GP-23708.

discharges from the reactor. Obviously, the counties that might be affected by a given reactor can lie in more than one state.

Furthermore, the health indices for a given county containing or near a reactor can be compared with the corresponding indices in certain "control" counties which are located far from the reactor but which are similar to the given county with regard to other characteristics.

3. Variables to be considered

The basic purpose of the type of study being discussed here is to relate health indices such as the annual infant mortality, fetal mortality, prematurity, and fertility rates for a given population to measures of the annual amounts and compositions of radioactive gaseous and liquid discharges from a given nuclear reactor. It is clear, however, that many other variables besides the radioactive discharges from the reactor can affect these health indices.

Some of the variables which ideally should be included in the study are the distribution of the population by age, sex, and race; meteorological data pertinent to the times of discharge of gaseous effluents and to the geographic distribution of the population; socioeconomic indices such as income, housing, education, and the quality of medical care; the sources of food and water; natural background radiation levels; radioactive fallout from bomb tests; levels of air pollution, both SO_2 and particulate matter; and personal characteristics, such as smoking and dietary habits. Obviously, the list could be extended almost indefinitely and, equally obviously, it will be very difficult to obtain the relevant data for many of them.

In addition, besides simply looking at the overall infant mortality rates for a given population and its various stratifications, it would be valuable to look at these rates for various specific causes of death. Although certain causes of death can more easily be associated with radiation effects than others, this analysis may not be as straightforward as it might at first appear. For example, it is possible that an infectious disease could cause the death of an infant whose susceptibility has been increased by exposure to low level radiation, but not cause the death of an infant who has not been so exposed. Even accidental deaths must be considered. One great hazard of being rushed to the hospital with an acute respiratory ailment is the high speed and often reckless ambulance or automobile ride that one must undergo.

Clearly, when analyzing data of the type being discussed here, one must always interpret changing rates with caution. Although there has been a general decrease in the infant mortality rate in the United States during the period from 1952 to 1967, there has also been a general increase in the prematurity rate over that period. This increase might reflect the increasingly deleterious effects of radiation, air pollution, or other environmental agents, or of changing practices of prenatal care. On the other hand, it might merely reflect changing practices in the reporting of birth weights, or even the beneficial effects of changing medical practices which convert potential fetal deaths into premature live births and consequently also bring about a decrease in the fetal death rate.

4. Results of preliminary regression analyses

In order to get a feeling for the possible magnitudes of the effects of radioactive wastes from nuclear reactors on infant mortality, and for the relative difficulty or ease with which these effects can be identified, some preliminary multiple regression analyses were carried out for the following four reactors: (1) the Dresden reactor in Grundy County, Illinois; (2) the Shippingport reactor in Beaver County, Pennsylvania; (3) the Indian Point reactor in Westchester County, New York; and (4) an experimental reactor at Brookhaven National Laboratory in Suffolk County, New York. In these regression models, the infant mortality rate in a given county containing or near a nuclear reactor, or the logarithm of this rate, was regressed on the amounts of radioactive gaseous and liquid wastes from the reactor and on either the infant mortality rate in a specified reference population or simply on a general linear trend in time. It must be emphasized that none of the multitude of other environmental agents and relevant variables listed earlier in this paper were specifically included in the models.

The general outcome of these preliminary studies is the only one that could have been anticipated, in view of the smallness of the effects and the simplicity of the model. Namely, the studies are inconclusive. They neither establish nor disprove the existence of an effect. They do, however, lead to the inescapable recommendation that more comprehensive and detailed studies of these questions are urgently needed.

The regressions that were carried out will briefly be summarized here. Time series of annual infant mortality rates for the United States as a whole, Illinois, Pennsylvania, New York, and the counties containing or near the four reactors studied are readily available from the published vital statistics of the federal government and the individual states, and will not be reproduced here. Time series of annual gaseous and liquid discharges from Dresden, Shippingport and Indian Point were obtained from the report "Radioactive Waste Discharges to the Environment from Nuclear Power Facilities" by Joe E. Logsdon and Robert I. Chissler, March, 1970, Bureau of Radiological Health, Environmental Health Service, Public Health Service, U.S. Department of Health, Education, and Welfare, Rockville, Maryland 20852. The time series of annual sand filter bed discharges and background radiation levels for the Brookhaven reactor were obtained from a report entitled "Background Radiation Levels in Brookhaven National Laboratory" by Andrew P. Hull, which was presented in March, 1970, at licensing hearings for the Shoreham nuclear power station on Long Island. None of these data are included in this paper because they are available in the sources cited and because the primary purpose of my reporting on these regression studies here is not to convince the reader of the validity and strength of particular conclusions that are reached. Rather, the purpose is to indicate that it is not possible to derive strong conclusions about either the existence or the nonexistence of an effect from the simple regression models used here, and to urge that a full scale statistical study of these problems be carried out.

5. Dresden

The Dresden reactor is located in Grundy County, Illinois, and began emitting radioactive discharges in 1960. Infant mortality rates in Grundy County were studied from 1950 to 1967, the most recent year for which these rates have been published, in order to include a relatively modern time period of reasonable length in which the reactor was inoperative as well as a time period of reasonable length in which it was active.

A relation of the following form was studied:

$$M_t = \beta_0 + \beta_1 t + \beta_2 X_{2t} + \epsilon_t. \tag{1}$$

In this relation, the index t represents the particular year being studied in the period from 1950 to 1967. For simplicity, only the final two digits of the year were used for identification, so that the year 1958, say, would be represented by the value $t = 58$. The interpretation of the other variables is as follows: M_t is the infant mortality rate in Grundy County in the year t (that is, the number of infant deaths per 1000 live births in that year), and X_{2t} is a two year moving average (for the years t and $t - 1$) of the liquid discharge (less tritium) from Dresden measured in curies. It was originally intended also to include the yearly gaseous discharges from Dresden in equation (1), but the gaseous and liquid discharges were highly correlated over the entire period. Hence, only the liquid discharges were used in this model.

The least squares estimates of β_0, β_1, and β_2 turn out to be $\hat{\beta}_0 = 55.4$, $\hat{\beta}_1 = -0.606$, and $\hat{\beta}_2 = 1.59$. Their estimated standard deviations are 30.6, 0.546, and 0.943, respectively, which yield the following t-statistics: 1.81, -1.11, and 1.68. Each of these t-statistics has 15 degrees of freedom. Thus, one might find in these values mild evidence of a positive relationship between liquid discharges and infant mortality superimposed on a general downward trend. The peak liquid discharge from Dresden was more than ten curies in 1966 (two year moving average), and it is seen by using the least squares estimate $\hat{\beta}_2$, that this value corresponds to an infant mortality rate of 15.9 deaths per 1000 live births above the overall linear trend.

It must be emphasized that none of these estimates are very reliable. Grundy County has a population of only 22,000 and the average number of infant deaths per year during the period being studied was only 11.4. Furthermore, it must be kept in mind that even if a definite relationship between infant mortality and radioactive discharges was established by these techniques, one would still be unable to conclude definitely that by actually reducing the discharges in future years, the infant mortality rate would be reduced. In fact, the discharges may simply be surrogates for some other variables which are the actual causative agents. However, it is fair to state that each scientist would regard such statistical evidence as at least favoring to a certain extent the hypothesis that the discharges are affecting the infant mortality rate.

When a similar analysis is carried out for LaSalle County, which is directly to the west of Grundy and has a population of 110,000, no evidence of a relationship is found. The least squares estimates are $\hat{\beta}_0 = 24.3$, $\hat{\beta}_1 = -0.029$, and $\hat{\beta}_2 = -0.222$. The corresponding t-statistics are 1.95, -0.13, and -0.58, respectively.

The next step was to replace M_t in (1) by $\log M_t$, since it is generally believed to be more appropriate to try to fit a linear trend to $\log M_t$ rather than to M_t itself. The results obtained were little changed from before. The t-statistics corresponding to $\hat{\beta}_2$ for Grundy and LaSalle Counties became 1.42 and -0.63, respectively. These values are not much different from their previous values.

Here, the fitted value of M_t is equal to the product of a factor $\exp\{\hat{\beta}_0 + \hat{\beta}_1 t\}$ representing the general trend in time, and a factor $\exp\{\hat{\beta}_2 X_{2t}\}$.

For Grundy County, we now have $\hat{\beta}_0 = 4.26$ and $\hat{\beta}_1 = 0.022$, and the least squares estimate of the general trend factor for 1966 is therefore 16.6. Also, we now have $\hat{\beta}_2 = 0.061$. Thus, the effect of the factor due to the liquid discharge of 10 curies in 1966 (two year moving average) is to multiply the estimated infant mortality rate for that year by $\exp\{0.61\} = 1.84$. This factor therefore corresponds to an infant mortality rate of 13.9 deaths per 1000 live births above the general trend for that year. This estimated increase is not much different from the estimated increase of 15.9 found from the linear model without logarithms.

One important consideration that makes an analysis of this type somewhat questionable, is that although the radioactive liquid discharge from Dresden was 0 for each year in the 1950's and only began to be positive in the 1960's, the populations of Grundy and LaSalle Counties may have been exposed to relatively large levels of radiation during the 1950's from bomb tests that were not present in the 1960's. Thus, in fact, the exposure of the population to radiation may actually have decreased when the bomb tests of the 1950's ceased and the Dresden reactor became active in the 1960's, rather than having increased, as is implicitly assumed in the models being used here.

In order to overcome this difficulty, the linear trend $\beta_0 + \beta_1 t$ in (1) was replaced by $\beta_0 + \beta_1 X_{1t}$, where X_{1t} is the infant mortality rate in the entire United States for the year t. In other words, it was felt that the ups and downs in the infant mortality rate in the United States through the years would reflect the general exposure of the population to radioactive fallout from bomb tests as well as other pollutants and other transient and sporadic effects. Thus, rather than assuming a linear trend, we assume that the expected infant mortality rate in Grundy County is a linear function of the infant mortality rate in the United States plus a multiple of the discharges from the Dresden reactor.

The model used is therefore

$$M_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \epsilon_t. \quad (2)$$

Regressions were also carried out in which X_{1t} was (i) the infant mortality rate in Illinois, rather than in the United States as a whole, (ii) the infant mortality rate in Illinois minus Cook County, since Chicago forms most of Cook County and was thought to have its own special characteristics; and (iii) the total infant mortality rate in Boone, DeWitt, Logan, McDonough, and Warren Counties in Illinois, which were chosen because they matched Grundy County to some extent with regard to their rural nature and their size, and were not near the reactor.

When X_{1t} is the United States infant mortality rate, the least squares estimates of β_0, β_1, and β_2 in (2) are $\hat{\beta}_0 = -28.4$, $\hat{\beta}_1 = 1.85$, and $\hat{\beta}_2 = 1.57$. Their estimated standard deviations are 45.8, 1.70, and 0.938, respectively. The t-statistic calculated from $\hat{\beta}_2$ is therefore 1.67. It is curious to note that this value is almost identical to the value found from equation (1).

When X_{1t} is the Illinois infant mortality rate, we find $\hat{\beta}_0 = -87.1$, $\hat{\beta}_1 = 4.36$, and $\hat{\beta}_2 = 0.937$, with estimated standard deviations 71.1, 2.85, and 0.552, respectively. The t-statistic calculated from $\hat{\beta}_2$ is therefore 1.70, again almost identical to the previous values.

When X_{1t} is the infant mortality rate in Illinois minus Cook County, we find $\hat{\beta}_0 = -51.6$, $\hat{\beta}_1 = 3.04$, and $\hat{\beta}_2 = 1.46$, with estimated standard deviations 39.9, 1.65, and 0.651, respectively. The t-statistic for $\hat{\beta}_2$ is therefore now 2.24.

When X_{1t} is the infant mortality rate in the group of matched counties, we find $\hat{\beta}_0 = 20.6$, $\hat{\beta}_1 = 0.0393$, and $\hat{\beta}_2 = 0.758$, with estimated standard deviations 16.5, 0.644, and 0.627, respectively. The t-statistic for $\hat{\beta}_2$ is now 1.21. It is interesting to note that the infant mortality rate in Grundy is almost totally unrelated to the infant mortality rate in the matching counties.

Finally, when M_t in (2) is replaced by the infant mortality rate in LaSalle County, we find that $\hat{\beta}_2$ is again negative in each of these regressions.

In summary, regardless of which regression model was used to study the infant mortality rate in small Grundy County, where the Dresden reactor is located, the coefficient $\hat{\beta}_2$ of the amount of radioactive liquid discharge was always found to be positive although the corresponding t-statistics were of modest magnitude. In neighboring LaSalle County, $\hat{\beta}_2$ was always found to be negative.

6. Shippingport

The Shippingport reactor, which is located in Beaver County, Pennsylvania, started up and began discharging tritium in 1958. It began emitting measurable radioactive gaseous and other liquid discharges the following year. The infant mortality rate in Beaver County, which has a population of 206,000 was also studied for the period from 1950 to 1967. The following regression model was used:

$$M_t = \beta_0 + \beta_1 t + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \epsilon_t, \tag{3}$$

where X_{2t} is a two year moving average of the gaseous discharges from Shippingport, X_{3t} is a two year moving average of the liquid discharges (less tritium), and X_{4t} is a two year moving average of the tritium discharges. Both X_{2t} and X_{3t} are measured in millicuries and X_{4t} is measured in curies.

No evidence of a positive relationship between the discharges and the infant mortality rate was found, and some of the estimated coefficients are negative. In particular, the least squares estimates are $\hat{\beta}_0 = 55.0$, $\hat{\beta}_1 = -0.569$, $\hat{\beta}_2 = -0.0093$, $\hat{\beta}_3 = -0.0023$, and $\hat{\beta}_4 = 0.032$. The corresponding t-statistics, each with 13 degrees of freedom, are 5.28, -3.02, -0.57, -0.24, and 1.13, respectively.

When equation (3) is applied to Allegheny County, Pennsylvania, which is directly to the southeast of Beaver and has a population of 1,628,000, including Pittsburgh, the estimates are $\hat{\beta}_0 = 38.1$, $\hat{\beta}_1 = -0.239$, $\hat{\beta}_2 = -0.0060$, $\hat{\beta}_3 = 0.0050$, and $\hat{\beta}_4 = -0.021$. The corresponding t-statistics are 12.6, -4.37, -1.26, 1.79, and -2.60. The stability of the infant mortality rates in Allegheny County, because of its large population, is reflected here in the relatively large magnitudes of the t-statistics. Among the coefficients, $\hat{\beta}_2$, $\hat{\beta}_3$, and $\hat{\beta}_4$ of the discharges, however, the t-statistic with the largest magnitude corresponds to the negative coefficient $\hat{\beta}_4$ of the tritium discharge. Thus, evidence of a positive relationship is again lacking.

Similar results were obtained when the linear term in time in (3) was replaced by the infant mortality rate in either the United States or Pennsylvania.

7. Indian Point

The Indian Point reactor, which is located in Westchester County, New York, started up and began emitting radioactive liquid discharges in 1962. The infant mortality rate in Westchester County, which has a population of 853,000, was studied for the period from 1950 to 1967 and regression analyses were carried out similar to those described for the Dresden reactor. Equation (1) was studied first with M_t denoting the infant mortality rate in Westchester and X_{2t} denoting the two year moving average of liquid discharges (less tritium) from Indian Point measured in curies. The gaseous and liquid discharges were again highly correlated (both were 0 over much of the period and then they rose together), so only liquid discharges are included in the regression equation. The least squares estimates of β_0, β_1, and β_2 are $\hat{\beta}_0 = 31.7$, $\hat{\beta}_1 = -0.178$, and $\hat{\beta}_2 = 0.059$, with estimated standard deviations 3.51, 0.062, and 0.028, respectively. Thus, the value of the t-statistic corresponding to $\hat{\beta}_2$, with 15 degrees of freedom, is 2.11.

The liquid discharges from Indian Point were more than 35 curies in both 1966 and 1967, and it is seen by using the least squares estimate $\hat{\beta}_2$ that this value corresponds to an infant mortality rate of 2.03 deaths per 1000 live births above the overall linear trend.

When a similar analysis is carried out for smaller Rockland County, which is across the Hudson River from Westchester and has a population of 192,000, the results are $\hat{\beta}_0 = 25.9$, $\hat{\beta}_1 = -0.064$, and $\hat{\beta}_2 = -0.037$, and the corresponding t-statistics are 1.75, -0.25, and -0.32, respectively.

When M_t is replaced by $\log M_t$ in (1), the results are little changed. The t-statistics corresponding to $\hat{\beta}_2$ for Westchester and Rockland Counties are 2.14 and -0.25, respectively. These values are not much different from their previous values. For Westchester County we now have $\hat{\beta}_0 = 3.54$, $\hat{\beta}_1 = -0.0081$, and $\hat{\beta}_2 = 0.0028$. For 1966, the factor of the fitted infant mortality corresponding to the general trend is therefore $e^{3.00} = 20.1$. The factor corresponding to the liquid discharge of 35 curies in that year (two year moving average) is $\exp\{0.098\} = 1.10$. Thus, the estimated increase in the infant mortality rate corresponding to the liquid discharge for that year, above the general trend, is 2.01 deaths per 1000 live births. Again, this value is in close agreement with the value found from the linear model without logarithms.

Next, equation (2) was studied for models in which M_t is the infant mortality rate in Westchester County and X_{1t} is the infant mortality rate in the United States. The least squares estimates are $\hat{\beta}_0 = 6.395$, $\hat{\beta}_1 = 0.571$, and $\hat{\beta}_2 = 0.065$, and the values of the corresponding t-statistics, again with 15 degrees of freedom, are 1.12, 2.68, and 2.11, respectively.

When X_{1t} is the infant mortality rate in New York State, the estimates are $\hat{\beta}_0 = 2.231$, $\hat{\beta}_1 = 0.802$, and $\hat{\beta}_2 = 0.045$, and the corresponding t-statistics are 0.13, 1.14, and 1.05, respectively.

When M_t is taken to be the infant mortality rate in Rockland County in these two models based on equation (2), the estimates of β_2 are $\hat{\beta}_2 = -0.046$ and $\hat{\beta}_2 = 0.055$. The corresponding t-statistics are -0.37 and 0.37.

Although these data can perhaps be interpreted as evidence at least mildly favoring the existence of a positive relationship between radioactive liquid discharges from Indian Point and the infant mortality rate in Westchester County, it must be emphasized that these discharges were 0 until 1962 and then increased monotonely from 1962 to 1967. Clearly then, this simple pattern might be present in the corresponding time series of many other environmental agents, one or more of which might actually be affecting infant mortality in the magnitude being attributed here to the Indian Point reactor. The fact however that these effects seem to be slightly more established in Westchester County than in Rockland County does provide some evidence, albeit weak, against this possibility.

8. Brookhaven

The final reactor to be studied was an experimental reactor at Brookhaven National Laboratory in Suffolk County, New York, for the period from 1951, the year that radioactivity was first used at the Laboratory, to 1968. The population of Suffolk County is 666,000. The following model was used:

$$M_t = \beta_0 + \beta_1 t + \beta_2 X_{2t} + \beta_3 X_{3t} + \epsilon_t, \tag{4}$$

where M_t was the infant mortality rate in Suffolk County in year t, X_{2t} is the two year moving average of the concentration of the sand filter bed discharge in picocuries per liter, and X_{3t} is an average of two offsite background radiation measurements made in year t and the two measurements made in year $t - 1$, measured in milliroentgens per week.

The least squares estimates turn out to be $\hat{\beta}_0 = 27.6$, $\hat{\beta}_1 = -0.142$, $\hat{\beta}_2 = 0.015$, and $\hat{\beta}_3 = -0.265$. The corresponding t-statistics, with 14 degrees of freedom, are 7.89, -2.19, 4.17, and -0.30. The striking aspect of this relation is the large t-statistic corresponding to the coefficient $\hat{\beta}_2$ of the concentration of the liquid discharge from the sand filter bed. From this relation it is found that an increase in the gross beta concentrations of the liquid releases of 300 pCi/liter, the observed value for 1961 (two year moving average) corresponds to an increase in the infant mortality rate of 4.5 deaths per 1000 live births.

These figures must again be interpreted with the greatest caution since the total amount of radioactivity in the liquid releases from the Brookhaven reactor is small, the maximum value being 219 millicuries in 1961. One interesting possibility suggested by this observation is that the actual composition of these releases may be as important as their total amount in affecting health.

It should also be noted that the background radiation levels bear essentially no relation to the infant mortality rates in Suffolk County.

When M_t is the infant mortality rate in Nassau County, which is to the west of Suffolk County on Long Island and has a population of 1,300,000, the estimates of the regression coefficients are $\hat{\beta}_0 = 24.9$, $\hat{\beta}_1 = -0.148$, $\hat{\beta}_2 = 0$ (to six decimal places), and $\hat{\beta}_3 = 1.66$. Only the years from 1951 to 1967 were included in this analysis, since the infant mortality rate in Nassau County in 1968 was not immediately available. The values of the t-statistics, with 13 degrees of freedom, for these four coefficients are 9.34, -2.79, -0.13, and 2.46, respectively. Thus, there is no evidence whatsoever of a relation between the filter bed discharge and the infant mortality rate in Nassau County, but there is now a relation in the observed data between off site background radiation levels and the infant mortality rate.

As before, when M_t is replaced by $\log M_t$ and the above analyses are carried out, the results are little changed.

9. Summary

It should be emphasized again that the results of these preliminary regression studies are inconclusive. They do not present strong evidence that there is a relationship between the exposure of a population to low level radiation from nuclear reactor discharges and the infant mortality rate in the population, and they do not present strong evidence that there is no such relation. The four reactors studied have different designs, and the inconclusive nature of these studies perhaps suggests that the actual composition of the discharges might be important, as well as whether and how these discharges enter the food chain. Some of the many other variables mentioned earlier in this paper, but not included in the regression models, are likely to be very influential.

The simple studies carried out here and their inconclusive results do lead, therefore, to a very strong and important recommendation. A large scale statistical study is urgently needed to aid in resolving this vital issue. Of course, statistical analysis can neither strictly prove nor disprove the hypothesis that exposure of a population to low level radiation increases the infant mortality rate. However, these analyses can substantially raise or lower the probability that the hypothesis is correct. Indeed, a large scale statistical study, such as the study of the effect of smoking on human health, could go far toward bringing the scientific community into agreement on this question.

In my classes, I usually define a scientist to be a person who can keep clearly in mind the distinction between the subjective utility that he assigns to any specific hypothesis and the subjective probability that he assigns to that hypothesis. In other words, a scientist must never let his hope or desire that there is no relation between low level radiation and infant mortality affect his professional evaluation of the probability that such a relation might exist. Statistical studies performed by interdisciplinary teams of scientists, in this strict sense, could provide information that will be of great help in reaching decisions regarding nuclear reactors that might critically affect large segments of the world's population.

◇ ◇ ◇ ◇ ◇

I am indebted to Dr. Ernest J. Sternglass, who initially stimulated my interest in this topic, for many helpful conversations. I am also indebted to Dr. Lincoln J. Gerende for his kind permission to use freely material he had prepared for a research proposal submitted jointly by him, Dr. Kenneth D. Rogers, and myself to the office of the Attorney General of Pennsylvania. I am further indebted to Dr. Gerende and Dr. Floyd H. Taylor for several valuable discussions of this project. Finally, I am indebted to William J. Franks, Jr., who did most of the groundwork and all of the computations for this paper, and whose assistance has been of great value.

Discussion

Question: P. Armitage, London School of Hygiene and Tropical Medicine, London

Isn't the analysis very sensitive to the true nature of the time trend? If the trend is really quadratic (as might be expected) with the curvature, may not the X factor be taking the place of the quadratic term?

Reply: M. DeGroot

The possible effects of the curvature of the trend were investigated by fitting a linear model to the logarithm of the infant mortality rate, a model for which there is some theoretical justification, as well as to the infant mortality rate itself. As I describe in this paper, the results for the two models were in close agreement and the magnitude of the effect of the radioactive discharge was almost the same for both models. The effects of trend curvature are also greatly reduced in those models where the rate in a given county is regarded as a linear function of the rate in some control population such as the state.

Question: V. L. Sailor, Brookhaven National Laboratory

It should be pointed out that the data used by Dr. DeGroot in his analysis of the Brookhaven Laboratory situation (liquid waste) does not have a plausible connection with infant mortality. The liquid waste flows into a stream which flows to the east through a completely uninhabited area to Peconic Bay away from the high density of population. The magnitude of the emissions are so small that they can no longer be detected a few miles off site, nor do the biota show activity. The total amount released over a period of twenty years was about 1¾ curies. During the same period Suffolk County had more than 100,000 times as much radioactivity deposited on it from weapons tests fallout. Gaseous radioactive release from Brookhaven was far greater (millions of curies per year), but these releases do not correlate with infant mortality since when the gaseous releases were high, the mortality rate was dropping. When gaseous releases were reduced, the mortality rate increased.

Reply: M. DeGroot

It is true that the total amount of liquid waste from Brookhaven National Laboratory was small compared to other contaminants. It is possible, therefore, that this discharge, which was zero until 1951, built up to its peak in 1961, and then steadily diminished, is simply acting as a surrogate for some other factor which seriously affects infant mortality but which was not explicitly identified in the analysis. On the other hand, it may well be true that the important consequences of radioactive discharges are derived not simply from the total amount, but rather from the actual composition of the effluent and the way in which various elements enter the food chain or otherwise reach the embryo.

Furthermore, the effect of radioactive releases on infant mortality cannot be measured simply by noting whether infant mortality went up or down in a given year, since there are obviously many other factors affecting infant mortality. The relevant measure of the effect of radioactive releases must be given in terms of whether or not infant mortality was higher in the given year than *it would have been* if these releases were not present but all the other factors were. It is this type of measure that the statistical methods described in this paper attempt to evaluate.

Question: J. Neyman, Statistical Laboratory, University of California, Berkeley

I am curious about the possible change in the socioeconomic composition of the population in a given county that might have occurred after a nuclear facility went into operation.

Also, how variable were the year to year numbers of live births in a given county. Did these numbers exhibit some temporal trend, and could there be any danger of some spurious correlations?

Reply: M. DeGroot

Dr. Neyman has raised two very interesting questions about my paper. First, he is quite correct that the construction of a nuclear reactor at a given site might well lead to changes in the socioeconomic composition of the population near that site which in turn lead to changes in the infant mortality rates. It is difficult to check this possibility because the relevant census data are published only every ten years. My own guess is that although there might be such changes in the immediate vicinity of the reactor (say within a few blocks), it is less likely that the composition of the county as a whole will shift because of the reactor. Of course, it may shift for other reasons in accordance with certain population trends or patterns, which is equally damaging to the analysis. However, I should think that the particular counties considered in my paper, rural Grundy as well as relatively populated Beaver, Westchester, and Suffolk, retained their same general character over the entire period studied. This question clearly requires further and more careful investigation.

Second, Dr. Neyman is again completely correct that a regression analysis based on rates is a tricky business when both the numerators and denominators are random variables, especially if the distribution of the number of live births in the denominator may be changing with time. Here, however, the yearly time series of the number of births and deaths in the various counties do not reveal any "substantial" changes over the period studied. Perhaps more reassuring, a glance at the graph of the time series of the infant mortality rate for each county seems to indicate that the variability of the annual rate remains roughly the same over the entire period.