American Journal of Epidemiology
Vol. 140, No. 10

Printed in U.S.A.

Use of Poisson Regression and Time Series Analysis for Detecting Changes over Time in Rates of Child Injury following a Prevention Program

Louise Kuhn,[1,2] Leslie L. Davidson,[1,2] and Maureen S. Durkin[1–3]

The use of two statistical methods to quantify time trends (Poisson regression and time series analysis) is illustrated in analyses of changes in child injury incidence after implementation of a community-based injury prevention program in Central Harlem, New York City. The two analytical methods are used to quantify changes in the rate of injury following the program, while taking into account the underlying annual and seasonal trends. Rates of severe injury during the period from 1983 to 1991 among children under the age of 17 years living in Central Harlem and in the neighboring community of Washington Heights are analyzed. The two methods provide similar point estimates of the effect of the intervention and have a good fit to the data. Although time series analysis has been promoted as the method of choice in analysis of sequential observations over long periods of time, this illustration suggests that Poisson regression is an attractive and viable alternative. Poisson regression provides a versatile analytical method for quantifying the time trends of relatively rare discrete outcomes, such as severe injuries, and provides a useful tool for epidemiologists involved with program evaluation. *Am J Epidemiol* 1994;140:943–55.

child; epidemiologic methods; incidence; intervention studies; pediatrics; Poisson distribution; wounds and injuries

A major use of epidemiology is to evaluate the impact of public health interventions on morbidity and mortality rates. While there is little dispute that randomized trials offer the strongest design for inference, evaluation of public health interventions must often rely on observational epidemiologic methods. One nonexperimental evaluation approach is to observe incidence rates over time (before and after an intervention), assess the extent to which changes in incidence followed the intervention, and determine whether or not any changes are greater than those predicted on the basis of trends observed beforehand. Statistical methods to quantify temporal variation are thus important for program evaluation.

This paper illustrates the use of two statistical methods to quantify time trends in incidence rates, Poisson regression and time series analysis, and discusses their relative strengths within the context of a particular program evaluation. The application of the two analytical methods uses data on

Received July 6, 1993, and in final form August 8, 1994.

Abbreviations: ACF, autocorrelation function; AR(1) or ARIMA (1,0,0), autoregressive model of the first order; ARIMA, autoregressive, integrated, and moving average; CI, confidence interval; MA, moving average; MA(1) or ARIMA (0,0,1), moving average parameter, first order; PACF, partial autocorrelation function; SCIPP, Statewide Childhood Injury Prevention Program.

[1] Gertrude H. Sergievsky Center, Columbia University, New York, NY.
[2] Division of Epidemiology, School of Public Health, Columbia University, New York, NY.
[3] New York State Psychiatric Institute, New York, NY.
Reprint requests to Louise Kuhn, Columbia University, Sergievsky Center, P & S, Box 16, 630 W. 168th Street, New York, NY 10032.

This paper is the Abraham M. Lilienfeld Student Prize Paper that was presented at the 26th Annual Meeting of the Society for Epidemiologic Research, Keystone, CO, June 16, 1993.

the incidence of severe injuries to children in Northern Manhattan, New York City (1). These data were collected to evaluate the effect of a community-based injury prevention program implemented in Central Harlem. The program, implemented in 1989, focuses on disadvantaged, inner-city, school-aged children, a group at high risk for unintentional and intentional injuries (2). The program activities included renovation of parks and playgrounds to provide safe play areas for children, youth development projects, organized recreational activities, traffic safety education, and violence prevention initiatives. Substantive information on the program and its evaluation has been published previously (3).

There is precedent in the literature for both Poisson regression and time series analysis, but they have not previously been systematically compared; moreover, Poisson regression has not been used, to our knowledge, in program evaluation directly. Poisson regression has been used in age-period-cohort analyses of mortality rates (4–6), and, in one application, to detect reductions in maternal mortality following legalization of abortion (7). The Poisson distribution has been promoted as the distribution of choice for relatively rare disease rates in a small area analysis of diagnosis-specific discharge rates (8). The ability of Poisson regression to adjust for covariates (9–11) is an important strength.

Time series analysis, on the other hand, has been previously used to evaluate the effect of communitywide injury prevention interventions (in all-cases legislation). These included studies of the effects of a 1974 nationwide speed limit law on motor vehicle deaths in the United States (12), a 1982 mandatory child restraint law on child hospitalizations for injury in Michigan (13), a 1986 firearms control ordinance on homicides in Detroit (14), and 1976 legislation restricting access to handguns on firearms-related homicides and suicides in the District of Columbia (15). Within epidemiology, time series methods have been promoted (16, 17) but are not yet widely used.

MATERIALS AND METHODS

Data sources

The data used in this analysis included injury deaths and hospital admissions for injury during the years 1983–1991 among children less than 17 years, residing in Central Harlem and Washington Heights, New York City. The address of each case was used to establish residence and was coded by census tract as either Central Harlem or Washington Heights. Injury refers to ICD-9 E-codes (18) between 800 and 999, which incorporate all trauma, poisonings, and burns as a result of intentional, unintentional, or undetermined injury causes. Records were abstracted from two major hospitals, which admit 76 percent of all hospitalized injury cases among residents in these two communities. Data on fatal cases not admitted to one of the study hospitals were obtained from death certificates. The data included 4,388 patients who survived and 114 persons who died. The 1980 and 1990 US census provided the population denominators, and a weighted average for each intercensus year was calculated.

Evaluation design

The objective of the evaluation was to assess whether or not monthly injury rates had decreased following the program and whether this decrease was greater than what might be expected on the basis of trends before the program. Injuries were observed to have a seasonal pattern, increasing in the summer months, and there was noticeable annual variation in the years before the program (secular increases in school-aged children and secular decreases in younger children). The specific aims were to assess, first, whether a decrease had occurred; second, whether any decrease was confined to Central Harlem (the area in which the prevention program was operational); and third, whether any decrease was specific to older children (5–16 years, the age group targeted by the prevention program) in Central Harlem. The years 1989, 1990, and 1991 were considered the intervention period during which the program could have shown an effect. The monthly incidence rates for the four area-age group combinations (Harlem, ages 5–16 years; Washington Heights, ages 5–16 years; Harlem, ages 0–4 years; and Washington Heights, ages 0–4 years) are plotted in figures 1–4.

Poisson regression

In the Poisson regression analysis of these data, the dependent or outcome variable (Y) is injury incidence obtained in each of a number of subgroups, described by a set of predictor variables $X_1, X_2, \ldots, X_k$. In our case, Y was a count of the number of injuries occurring in each month, in each age category (0–4 years; 5–16 years) and each geographic area (Central Harlem and Washington Heights). Hence, time is treated as an independent covariate. The method of Poisson regression (19–24) is described below.

In Poisson regression, the underlying probability distribution of the dependent variable Y is assumed to be Poisson, which is a discrete distribution usually used to model counts of rare events occurring during an interval of time. A Poisson random variable can take any nonnegative, integer value. The Poisson distribution has only one parameter μ equal to the mean and variance. The distribution is given by the formula:

$$\Pr(Y = y;\mu) = \mu^y e^{-\mu}/y!$$
$$y = 0, 1, 2, \ldots. \qquad (1)$$

The goal of Poisson regression is to fit to observed data a regression equation that accurately models the expected value of the dependent variable Y, $E(Y)$, as a function $\boldsymbol{\lambda}$ of a set X of independent variables $X_1, X_2, \ldots X_k$ and β regression parameters. If Y is the number of events in a subgroup, and N

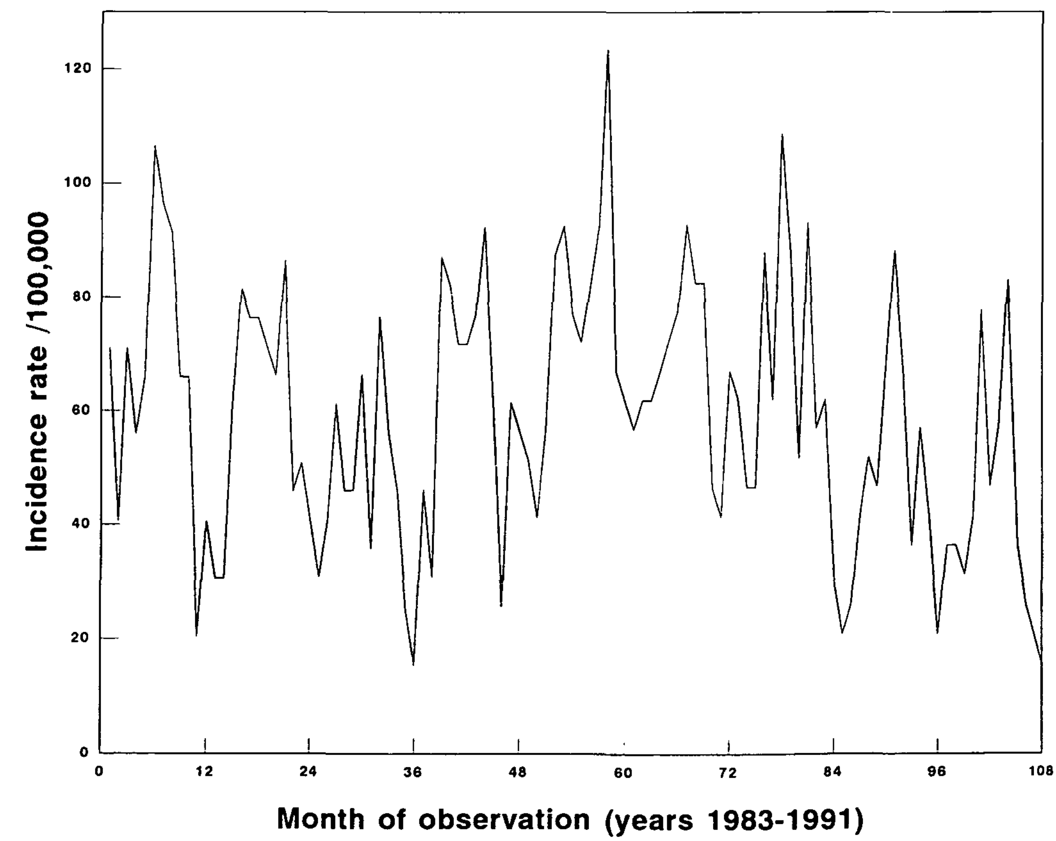

FIGURE 1. Monthly incidence rates of severe injury, ages 5–16 years: Central Harlem, New York City, 1983–1991.

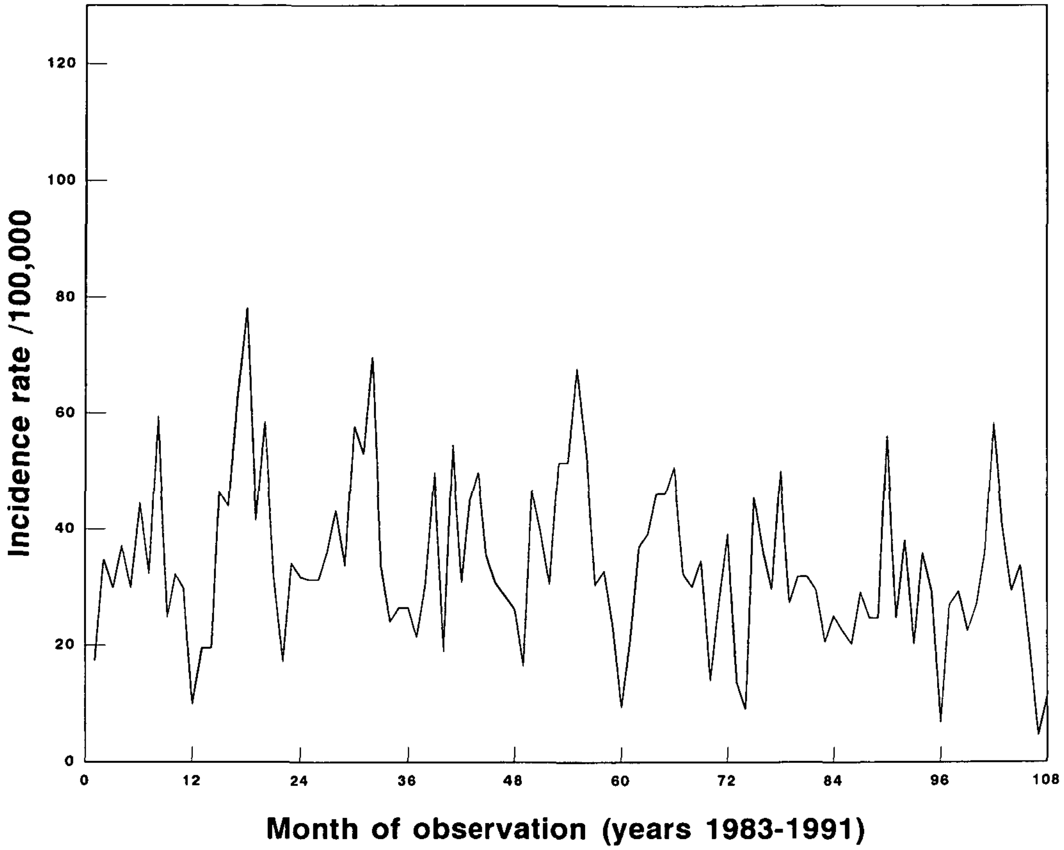

FIGURE 2. Monthly incidence rates of severe injury, ages 5–16 years: Washington Heights, New York City, 1983–1991.

is the population size or total follow-up time in that subgroup, then

$$E(Y) = N\lambda(X, \beta). \quad (2)$$

If this function (equation 2) is substituted into equation 1 above, the general form of the likelihood function for Poisson regression analysis is obtained:

$$L(Y; \beta) = \Pi\,[N\lambda(X, \beta)]^{Y} e^{-N\lambda(X,\beta)}/y!. \quad (3)$$

Estimates of the β parameters are obtained by maximizing this likelihood function. The outcomes (Y) in each subgroup defined by the independent variables (X) are required to be independent for this likelihood to be correct. To use the likelihood function, it is necessary to specify the function λ that is generally assumed to be log-linear. It is the natural log (ln) of the expected rate of events Y that is modeled as a linear function of the X predictor variables and thus

$$\lambda(X,\beta) = \exp(X\beta) =$$

$$\exp(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k).$$

In sum, Poisson regression models can be understood as a subclass of generalized linear models in which the systematic effects are multiplicative, the distribution of the errors is Poisson, and the link function is the natural log.

Four separate regression models were fitted, one to each age and area category (0–4 years, Central Harlem; 0–4 years, Washington Heights; 5–16 years, Central Harlem; 5–16 years, Washington Heights), using maximum likelihood estimation in the statistical package GLIM (25, 26). First,

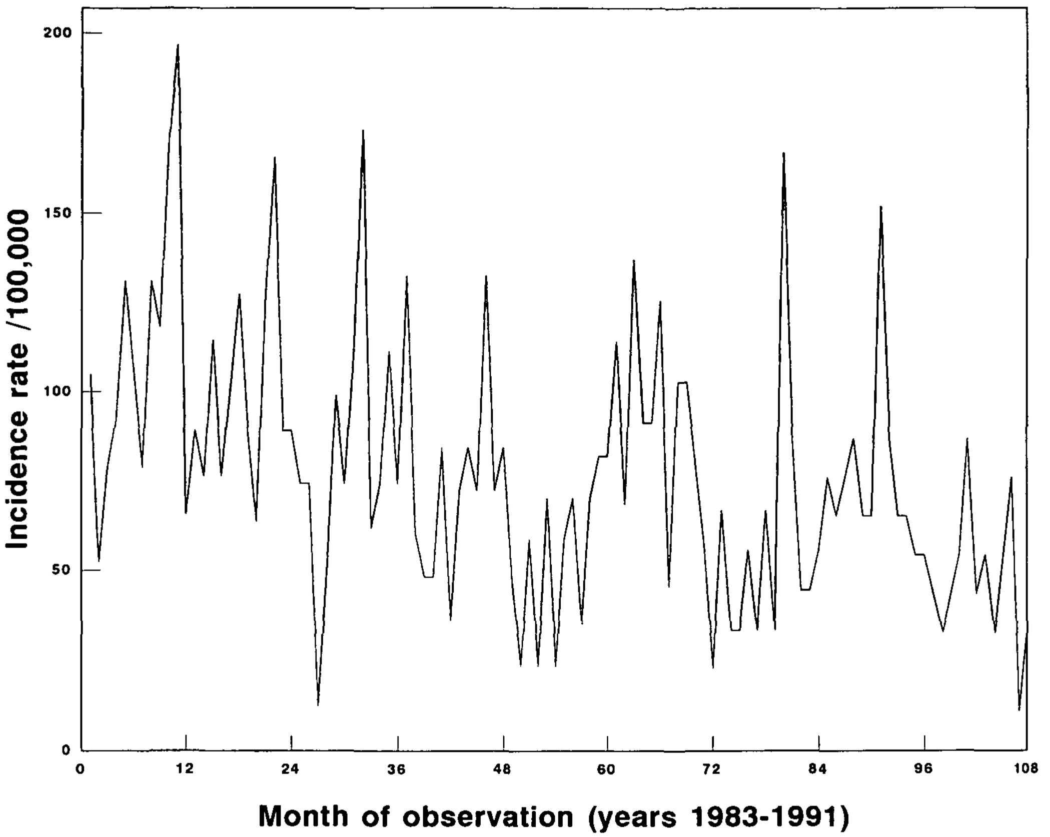

FIGURE 3. Monthly incidence rates of severe injury, ages 0–4 years: Central Harlem, New York City, 1983–1991.

to take seasonality into account, the expected value of the injury rate in each month within each age-area category was modeled as a log-linear function of season. The months January through June were coded increasing 1 through 6, and the months July through December were coded decreasing 6 through 1. Next, an indicator variable with 1989, 1990, and 1991 signifying intervention years was added. A feature of Poisson regression is that the regression coefficients can be interpreted as an estimate of the log of the relative risk, adjusted for the other predictors in the model. In our case, the indicator variable for the intervention provides an estimate of the ratio of the injury rate after the intervention to the injury rate before the intervention, or the change in the injury rate following the onset of the intervention compared with the injury rate before the intervention. Finally, year was entered as a continuous variable to represent the underlying annual trend. This model assumes that the annual trend is uniform over time and that the effect of the intervention is to decrease the level of the injury rate. Confidence intervals and hypothesis tests for the regression coefficients were calculated by assuming asymptotic normality of the maximum likelihood estimates of the regression parameters. The model in each age-area category can be summarized as follows:

$$\ln E(Y) = \alpha + \beta_1(\text{season}) + \beta_2(\text{intervention}) + \beta_3(\text{year})$$

with season as 1 = January, . . ., 6 = June, 6 = July, . . ., 1 = December; intervention as 1 = after (1989–1991) or 0 = before (1983–1988); and year as 83 through 91.

Four separate regression models in each age-area category give equivalent results to a fully specified single model in which age

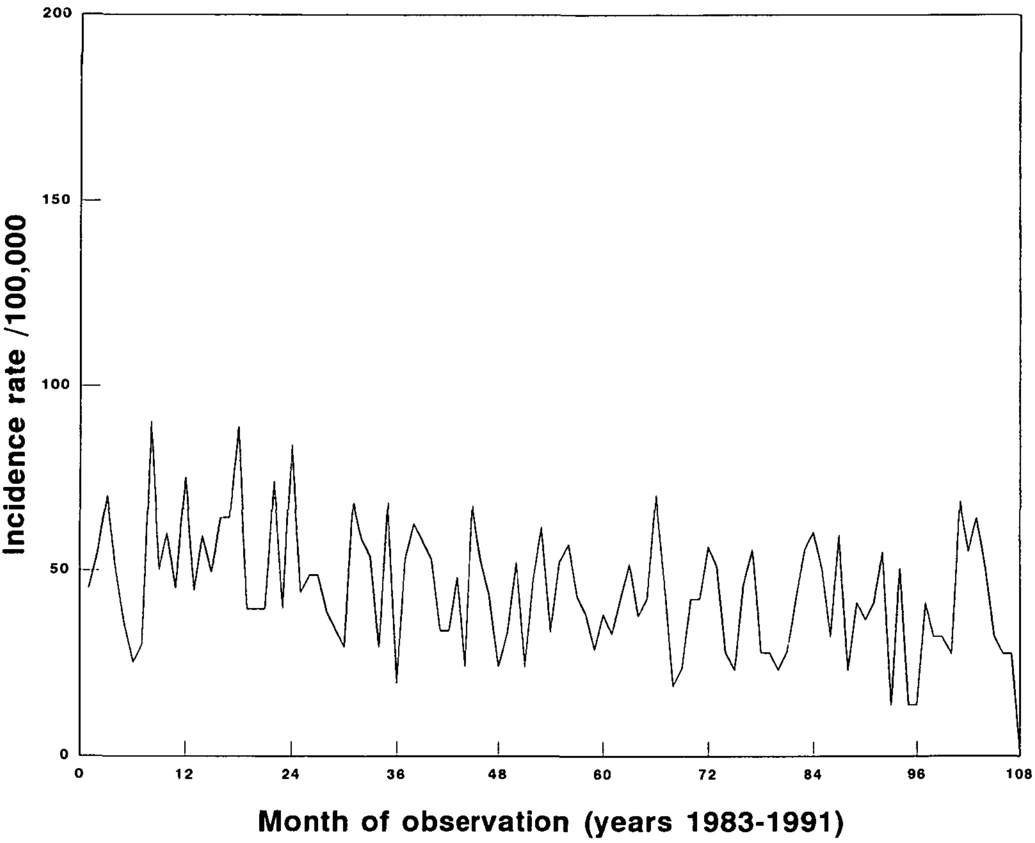

FIGURE 4. Monthly incidence rates of severe injury, ages 0–4 years: Washington Heights, New York City, 1983–1991.

and area are included as covariates. We present the four separate models here to avoid the complexity entailed in the interpretation of three-way interactions, which would be needed in a single model to test the evaluation question. In order to test whether the intervention had a different effect in the intervention area compared with the control area, among the age group targeted by the program, the difference in the regression coefficients for the intervention period from the models in Central Harlem, ages 5–16 years ($\hat{\beta}_{2CH}$), and Washington Heights, ages 5–16 years ($\hat{\beta}_{2WH}$), was considered using the asymptotic normality of the maximum likelihood estimates of the regression parameters. Thus, we test H_0 $\hat{\beta}_{2CH} - \hat{\beta}_{2WH} = 0$, where $\text{var}(\hat{\beta}_{2CH} - \hat{\beta}_{2WH}) = \text{var}(\hat{\beta}_{2CH}) + \text{var}(\hat{\beta}_{2WH})$, with $\hat{\beta}_{2CH}$ and $\hat{\beta}_{2WH}$ independent.

Each model was tested for overdispersion by comparing the deviance with its degrees of freedom. The deviance is a measure of the discrepancy between observed and fitted values, and comparison of it with its degrees of freedom provides a measure of goodness of fit. If the variability observed is greater than that predicted by the Poisson distribution, the data are said to be overdispersed. If the goodness-of-fit test is significant, this may also be an indication of model misspecification. The Pearson residuals were plotted and inspected. A normal quantile plot and the Filliben correlation coefficient were used to check their assumed distributions (26).

Time series analysis

Time series analysis looks at the relation between observations collected at consecutive, regular intervals over time. While it is also a regression method, it predicts the dependent variable not from independent covariates but from values of the outcome at previous points in time. The analytical task, in our case, was to create models in which the injury rate in a particular month was predicted on the basis of injury rates in previous months. Time series models based on autoregressive, integrated, and moving average (ARIMA) parameters were used (27) and will be briefly described based on the presentation in methodological texts (16, 17, 28, 29).

Let $\ldots Y_{t-1}, Y_t, Y_{t+1} \ldots$ denote the injury rate in each month $t-1, t, t+1$; let $\ldots Z_{t-1}, Z_t, Z_{t+1} \ldots$ denote the deviation in the injury rate from the long-term mean at time points $t-1, t, t+1$; and let $\ldots a_{t-1}, a_t, a_{t+1} \ldots$ denote a white noise process or a series of normally and independently distributed random variables with mean zero and variance σ^2_a (random shocks). The autoregressive parameter assumes that the Z_t is linearly dependent on Z_{t-1} and on the random shock a_t:

$$Z_t = \phi_1 Z_{t-1} + a_t$$

where ϕ is a parameter. If only the immediately preceding observation is included, it is referred to as an autoregressive model of the first order (AR(1) or ARIMA (1,0,0)). The moving average parameter considers Z_t as a linear combination of the present (a_t) and previous random shock (a_{t-1}), at the first order MA(1) or ARIMA (0,0,1):

$$Z_t = a_t - \theta_1 a_{t-1}$$

where θ is a parameter. These two models are special cases of a more general model that uses not only the immediately preceding observation but also p and q prior observations: AR(p) and MA(q). These can be combined into the general-case autoregressive moving average model of order p and q, ARIMA (p, 0, q) as follows:

$$Z_t = \phi_1 Z_{t-1} + \ldots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}.$$

The above models assume the series to be stationary; i.e., the probability structure (mean and variance) does not change over time. When this assumption is not met, the series can be differenced, $\nabla Z_t = Z_t - Z_{t-1}$, which usually ensures stationarity. When the analysis is based on the differences rather than the original series, it is called an integrated model, ARIMA (p,d,q). The natural log of the observed injury rate was used as the dependent variable in the time series analysis to reduce the degree to which the variance changes over time.

We developed four separate ARIMA models for the injury rate series in each age and area category (Central Harlem, ages 0–4 years; Washington Heights, ages 0–4 years; Central Harlem, ages 5–16 years; Washington Heights, ages 5–16 years). Based on observation of a seasonal trend, seasonal models were built a priori. The difference of observations a full season apart, 12 months in the case of our data which were collected monthly, was taken, ARIMA $(p,d,q)_{12}$. $\nabla_{12} Z_t = Z_t - Z_{t-12}$.

For each age-area category, the autocorrelation function (ACF) and partial autocorrelation function (PACF) of seasonal differences were calculated. The ACF refers to the correlation between Z_t and Z_{t+k}, where k refers to units of time or the lag. Usually the ACF of an AR(1) follows an exponential curve, whereas the ACF of a MA(1) process shows a single peak at time lag k. The ACF and PACF were used as the main diagnostic tool to identify the parameters in the model following guidelines on their use (28, 29).

The parameters of the provisional model were estimated using the maximum likelihood algorithms implemented by SPSS (30). If the parameters were statistically significant and did not tend toward invertibility (a condition where observations farther away have increasing weight, an indication of model misspecification), the residuals of the model were inspected. If there was no apparent trend in the autocorrelation function of the residuals and if none of the Box-Ljung chi-square statistics at any lag less than 12 was significant, the model was accepted.

Once an ARIMA model that adequately fit the data had been determined, a parameter to estimate the effect of the intervention was included. An indicator variable coded as 0 before the intervention (1983–1988) and 1 after the intervention (1989–1991) was regressed on the series. The regression coefficient for the intervention period indicates the extent to which the series after the intervention date is different from that predicted on the basis of earlier trends defined by the autoregressive, moving average, and integration components of the model (31).

RESULTS

Poisson regression

Adjusting for annual and seasonal trends, there was a decrease in the number of injuries in the 5- to 16-year age group in both areas during the intervention period (table 1). The relative risk (ratio of injury rate after the program to injury rate before the program) was estimated to be 0.79 (95 percent confidence interval (CI) 0.64 to 0.97) in Central Harlem and 0.78 (95 percent CI 0.65 to 0.93) in Washington Heights in the 5- to 16-year age category. This corresponds with point estimates of the reduction in incidence rates of 21 percent in Central Harlem and 22 percent in Washington Heights (5- to 16-year age category). The difference between Central Harlem and Washington Heights in the size of the reduction in the 5- to 16-year age group was not statistically significant (z = 0.115; p = 0.867). In the younger age group in both areas, there were no differences in the injury rate following the program once the annual trend was taken into account.

A strong seasonal effect was seen in the 5- to 16-year age group in both areas, but no seasonal effect was seen among the younger age group. The annual trends were apparently very small in both areas in the 5- to 16-year age group, but in the 0- to 4-year age group, decreases in both areas were statistically significant.

All tests of overdispersion were not significant, suggesting no important extra-Poisson variability. There was no apparent heteroscedasticity in the residuals. Normal quantile plots appeared to fit a straight line well, with high values of the Filliben correlation.

TABLE 1. Poisson regression models in each age-area category, with intervention effect estimates adjusting for seasonal and yearly trends: Northern Manhattan, New York, 1983–1991

Model	$\hat{\beta}$	SE* ($\hat{\beta}$)	p value	RR*	95% CI*
Central Harlem, 5–16 years					
Year	0.01	0.019	0.591		
Season	0.13	0.017	<0.001		
Intervention period	−0.24	0.105	0.024	0.79	0.64 to 0.97
Central Harlem, 0–4 years					
Year	−0.09	0.025	<0.001		
Season	0.03	0.022	0.138		
Intervention period	0.08	0.143	0.582	1.08	0.70 to 1.15
Washington Heights, 5–16 years					
Year	0.0002	0.017	0.991		
Season	0.16	0.015	<0.001		
Intervention period	−0.25	0.094	0.007	0.78	0.65 to 0.93
Washington Heights, 0–4 years					
Year	−0.05	0.021	0.012		
Season	0.02	0.018	0.399		
Intervention period	0.01	0.119	0.942	1.01	0.80 to 1.25

*SE, standard error; RR, relative risk; CI, confidence interval.

Time series

Using seasonal ARIMA time series models of the log-transformed series, we found the intervention period to be associated with a decrease in the injury rate in all models (table 2), but this decrease was almost statistically significant only in Washington Heights (5–16 years). From the time series models, the reduction in the injury rate following the program was estimated to be 21 percent in Central Harlem and Washington Heights (5–16 years). There appeared to be no effect in the younger age group, and confidence intervals around the regression coefficients were very wide.

The specific seasonal ARIMA models selected and their parameters for each age and area category are displayed in table 2. The autocorrelation function of the residuals of these models showed no systematic trend, and none of the Box-Ljung chi-square statistics at any lag less than 12 was significant.

Comparison of Poisson regression and time series analysis results

The magnitude of the intervention effect estimates calculated by the two methods is notably consistent with each other. Both methods show a decrease in the injury rate among children aged 5–16 years in Central Harlem and Washington Heights during the years following the injury prevention program, although the decrease in the time series analysis for Central Harlem is not statistically significant. Both methods suggest that the decrease was not confined to Central Harlem, the area in which the program was implemented. Both methods agree that the decrease was confined to older children, the age group targeted by the program. The intervention effect estimates from the Poisson regression model appeared to be more precise than from the time series analysis models, as confidence intervals around these estimates were narrower in the Poisson regression model.

To illustrate how the two approaches differ in the way they represent or predict variations in monthly incidence rates, we plotted for 5- to 16-year-olds in Central Harlem the rates predicted by the Poisson model (figure 5) and by the ARIMA model (figure 6). The model-predicted rates in both figures are superimposed on the ob-

TABLE 2. Time series analysis (ARIMA* models) in each age-area category of the log-transformed series, with intervention effect estimates: Northern Manhattan, New York, 1983–1991

	$\hat{\beta}$	SE* ($\hat{\beta}$)	*p* value	95% CI*	Incidence†
Central Harlem, 5–16 years					
MA(1)*	−0.25	0.098	0.012		
AR(12)*	−0.46	0.093	<0.001		
Intervention period	−0.23	0.134	0.087	−0.49 to 0.03	63.0
Central Harlem, 0–4 years					
MA(1)	0.11	0.103	0.287		
AR(12)	−0.47	0.097	<0.001		
Intervention period	−0.07	0.172	0.702	−0.41 to 0.27	85.3
Washington Heights, 5–16 years					
MA(1)	0.03	0.103	0.794		
AR(12)	−0.41	0.105	<0.001		
Intervention period	−0.24	0.121	0.054	−0.003 to −0.48	36.5
Washington Heights, 0–4 years					
MA(1)	0.10	0.102	0.349		
AR(12)	−0.54	0.091	<0.001		
Intervention period	−0.16	0.132	0.223	−0.42 to 0.10	47.7

*ARIMA, autoregressive, integrated, and moving average; SE, standard error; CI, confidence interval; MA(1), moving average parameter, first order; AR(12), seasonal autoregressive model using 12-month lag.

†Mean monthly injury rate/100,000 (preintervention).

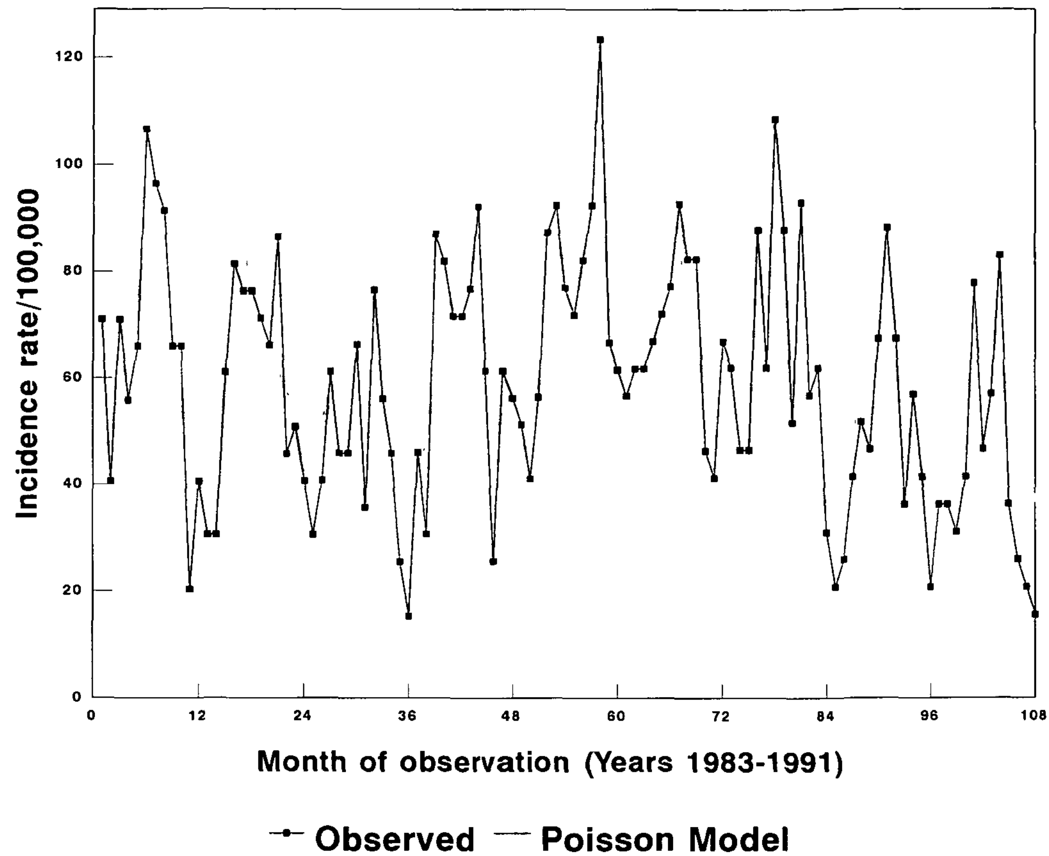

FIGURE 5. Observed and predicted (Poisson model) monthly incidence rates of severe injury, ages 5–16 years: Central Harlem, New York City, 1983–1991.

served rates for the same group (5- to 16-year-old children in Central Harlem).

DISCUSSION

The analysis of temporal variation is a basic aspect of descriptive epidemiology (32) and has relied, to a large extent, simply on graphical display of incidence and prevalence rates. The question thus arises of whether there is interpretational advantage of complex statistical approaches over more simple graphical approaches (4). In the case of program evaluation, there is an explicit need to describe and test the extent of change and hence to use statistical methods to quantify time trends.

A further question that may be asked is whether or not there is an advantage to focusing on time trends in program evaluation over simple before-after comparisons. The evaluation of the Statewide Childhood Injury Prevention Program (SCIPP) in Massachusetts (33, 34) is of interest here as SCIPP was a community-based prevention program targeting a broad range of childhood injury, and evaluation considered the impact on childhood injury incidence, similar to our case. The SCIPP evaluation used the ratio of the injury rate before the program to the injury rate following the program and, as a measure of program effect, compared these ratios across intervention and control communities. In contrast to the SCIPP approach, analyses of time trends, such as Poisson regression and time series analysis, take into account the variability that occurs over the study period apart from the change associated with the intervention. They also avoid the loss of information about variability in incidence over time that

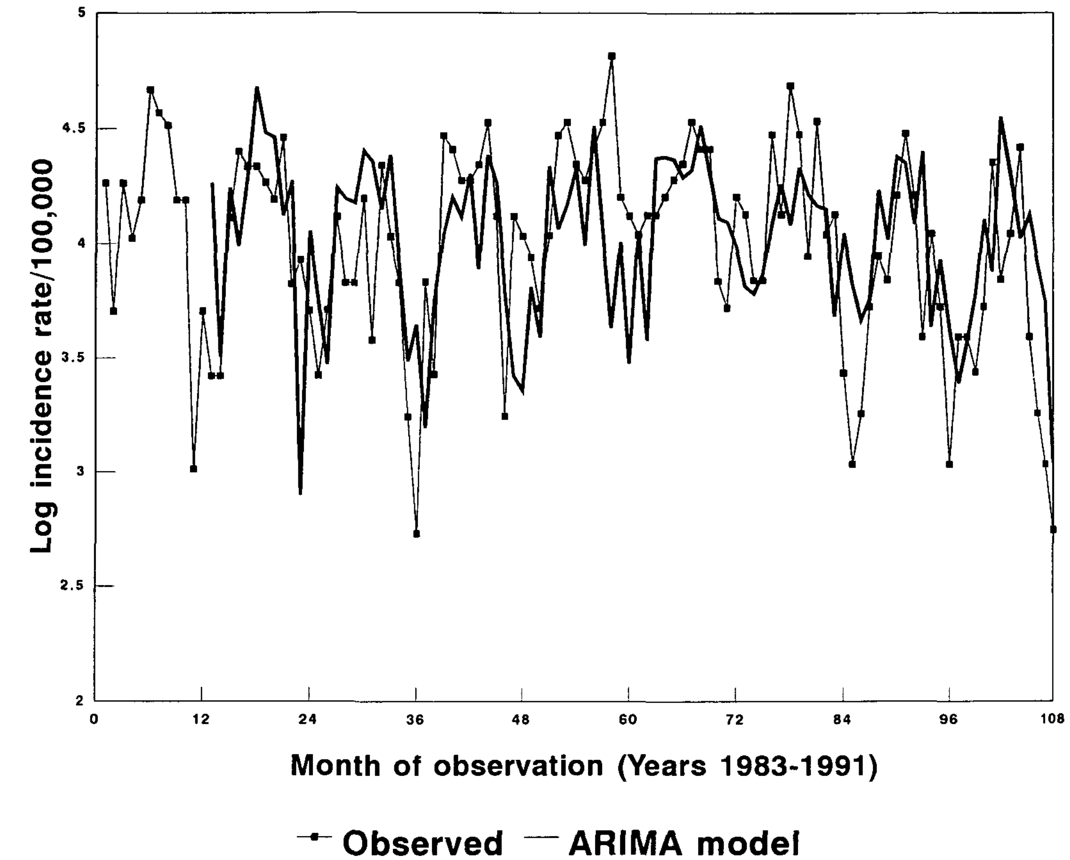

FIGURE 6. Observed and predicted (autoregressive, integrated, and moving average model, ARIMA) monthly log incidence rates of severe injury, ages 5–16 years: Central Harlem, New York City, 1983–1991.

occurs when rates are aggregated into one before and one after rate.

Thus, given the potential use of statistical methods to quantify time trends, let us compare Poisson regression and time series analysis as statistical methods to accomplish this analytical task. Time series analysis has been promoted as the statistical method of choice in the case of data arising from observations collected over long periods of time (35). One reason for this is that time series analysis does not require that observations at sequential time points be independent; in fact, it focuses on the dependence of consecutive observations. Time series analysis has been previously used to evaluate communitywide interventions, including the enactment of injury prevention legislation (12–15).

The illustration presented here suggests that Poisson regression provides an attractive alternative statistical method to time series analysis for epidemiologists wishing to measure an effect of an intervention. Poisson regression provides a means of expressing the effect of the intervention in terms of the rate ratio. It allows the analysis of aggregate data to be congruous with the analytical methods used in case-control and cohort studies (23). This application uses injury rates, but the method is potentially useful to analysis of other vital statistics data. We discuss here the underlying assumptions of Poisson regression, the interpretation of findings from these models, and the ability to include population denominators and covariates.

A strong case can be made for the appropriateness of the Poisson distribution to analyze injury incidence. It is considered the distribution of choice for rare events and has been selected as the appropriate stochastic model for natural variability of vital statistics, such as mortality rates (36). An empirical comparison of different distributions to describe the variability of diagnosis-specific hospital discharge rates concluded that a Poisson probability model with an extrasystematic component of variance was better than other probability models (37).

Some assumptions of Poisson regression may appear, on first impression, to be violated in the present illustration. Poisson regression requires that outcomes are uncorrelated. It seems untenable to assume that injuries occurring over time are independent events. The effects of social conditions and the physical environment are likely to persist over time. Injuries often occur in clusters, such as among motor vehicle occupants and victims of interpersonal violence or house fires. Poisson regression assumes that the population is homogenous with respect to risk of injury. This is another questionable assumption as injuries do not occur at random but have predictable precursors and known patterns of risk. Specific demographic, socioeconomic, and behavioral factors are linked to the occurrence of injuries (2, 38).

Despite these apparent violations of assumptions, tests of overdispersion were not significant, and plots of residuals showed no apparent heteroscedasticity. Other regression diagnostics suggested that the Poisson regression model fit the data well. It may be useful to recall that, more precisely, the Poisson regression model requires that the population meet the above criteria conditional on the independent variables (39). The inclusion of covariates may have resulted in conditional independence and created strata of homogenous risk. This may explain why the method seemed, in this application, to be relatively robust to these apparent violations of underlying assumptions.

The estimates of the intervention effect appeared to be more precise in the Poisson regression models than in the time series analysis models, suggesting increased power of the former method, in this application. One possible explanation for this may be that the time series models required estimation of a larger number of parameters. However, the standard errors may be deceptively smaller than they should be if true autocorrelation in the data is ignored. Furthermore, the apparent efficiency of Poisson regression may not hold in other circumstances and requires further investigation.

Both Poisson regression and time series analysis are statistical methods to quantify time trends, but they differ in their conceptualization of the analytical task and underlying assumptions. This paper has described these methods and illustrated their use to quantify changes in injury rates among children following a prevention program. The two methods provided estimates of the effect of the program that were of a similar order of magnitude, and both had an apparent good fit to the data. While time series analysis has been promoted as the method of choice in the analysis of sequential observations over long periods of time, this illustration suggests that Poisson regression is a viable alternative. The method may be generalizable to other relatively rare discrete outcomes or to the occurrence of severe injury in other contexts and is a useful tool for epidemiologists involved with program evaluation.

ACKNOWLEDGMENTS

This project was supported by the Centers for Disease Control and Prevention (grants R49/CCR 202585 and R48/CCR 205055), the National Institute of Neurological Diseases and Stroke (grants R29 NS27971-03 and R29 NS27971-04), the Mellon Foundation, Health of the Public (grant UCSF 1212SC), the Robert Wood Johnson Foundation, and the New York State Psychiatric Institute.

The authors thank Dr. Bruce Levin for statistical advice; Drs. Barbara Barlow, Pat O'Connor, and Margaret Heagarty who are our coinvestigators in the evaluation of the injury prevention program; and Dr. Jamal Naqvi for assistance with data collection.

REFERENCES

1. Davidson LL, Durkin MS, O'Connor P, et al. The epidemiology of severe injuries to children in Northern Manhattan: methods and incidence rates. Paediatr Perinat Epidemiol 1992;6: 153–65.
2. Durkin MS, Davidson LL, Kuhn L, et al. Low income neighborhoods and the risk of severe pediatric injury: a small area analysis in Northern Manhattan. Am J Public Health 1994;84:587–92.
3. Davidson LL, Durkin MS, Kuhn L, et al. The impact of the Safe Kids/Healthy Neighborhoods Injury Prevention Program. Am J Public Health 1994;84:580–7.
4. Kupper LL, Janis JM, Karmous A, et al. Statistical age-period-cohort analysis: a review and critique. J Chronic Dis 1985;38:811–30.
5. Sverre JM. Trends in suicide mortality among elderly in Norway, 1966–1986. Epidemiology 1991;2:252–6.
6. Woodbury MA, Manton KG, Blazer D. Trends in U.S. suicide mortality rates 1968 to 1982: race and sex differences in age, period, and cohort components. Int J Epidemiol 1988;17:356–62.
7. Parker RA. Analysis of surveillance data with Poisson regression: a case study. Stat Med 1989; 8:285–94.
8. Diabetes Epidemiology Research International Group. Secular trends in the incidence of childhood IDDM in 10 countries. Diabetes 1990;39: 858–64.
9. Nystrom L, Dahlquist G, Wall S. The Swedish childhood diabetes study: an analysis of the temporal variation in diabetes incidence 1978–1987. Int J Epidemiol 1990;19:141–6.
10. Rewers M, Stone RA, LaPorte RE, et al. Poisson regression modeling of temporal variation in incidence of childhood insulin-dependent diabetes mellitus in Allegheny County, Pennsylvania, and Wielkopolska, Poland, 1970–1985. Am J Epidemiol 1989;129:569–81.
11. McMahon LF, McLaughlin CG, Petroni GR, et al. Small area analysis of hospital discharges for musculoskeletal diseases in Michigan: the influence of socioeconomic factors. Am J Med 1991; 91:173–8.
12. Martinez-Schnell B, Zaida A. Time series analysis of injuries. Stat Med 1989;8:1497–1508.
13. Margolis LH, Wagenaar AC, Lui W. The effects of a mandatory child restraint law on injuries requiring hospitalization. Am J Dis Child 1988; 142:1099–1103.
14. O'Carroll PW, Loftin C, Waller JB, et al. Preventing homicide: an evaluation of the efficacy of a Detroit gun ordinance. Am J Public Health 1991;81:576–81.
15. Loftin C, McDowall D, Wiersema B, et al. Effects of restrictive licensing of hand guns on homicide and suicide in the District of Columbia. N Engl J Med 1991;325:1615–20.
16. Catalano R, Serxner S. Time series designs of potential interest to epidemiologists. Am J Epidemiol 1987;126:724–31.
17. Helfenstein U. The use of transfer function models, intervention analysis, and related time series methods in epidemiology. Int J Epidemiol 1991; 20:808–15.
18. Health Care Financing Administration. International classification of diseases. Ninth Revision. Clinical modification. Washington, DC: US GPO, 1980. (DHHS publication no. (PHS) 80-1260).
19. Nelder JA. Log linear models for contingency tables: a generalization of classical least squares. Appl Stat 1974;23:323–9.
20. Kleinbaum DG, Kupper LL, Muller KE. Applied regression analysis and other multivariable techniques. Boston: PWS-Kent, 1988.
21. McCullagh P, Nelder JA. Generalized linear models. New York: Chapman-Hall, 1989.
22. Frome EL. The analysis of rates using Poisson regression models. Biometrics 1983;39:665–74.
23. Clayton D, Schifflers E. Models for temporal variations in cancer rates. I. Age-period and age-cohort models. Stat Med 1987;6:449–67.
24. Clayton D, Schifflers E. Models for temporal variations in cancer rates. II. Age-period-cohort models. Stat Med 1987;6:469–81.
25. Payne CD. The GLIM system release 3.77. Oxford: Numerical Algorithms Group, 1986.
26. Aitkin M, Anderson D, Francis B, et al. Statistical modelling in GLIM. Oxford: Claredon Press, 1989.
27. Box GEP, Jenkins GM. Time series analysis: forecasting and control. Oakland: Holden-Day, 1976.
28. Chatfield C. The analysis of time series: an introduction. New York: Chapman-Hall, 1984.
29. McCleary R, Hay RA. Applied time series analysis for the social sciences. Beverly Hills: Sage, 1980.
30. Statistical Package for the Social Sciences (SPSS). SPSS/PC+ trends: procedures for analyzing and forecasting time series. Chicago: SPSS, Inc, 1988.
31. Box GEP, Tiao GC. Intervention analysis with applications to economic and environmental problems. J Am Stat Assoc 1975;70:70–9.
32. Mauser JS, Kramer S. Epidemiology: an introductory text. Philadelphia: WB Saunders, 1985.
33. Bass JL, Mehta KA, Ostrovsky M. Childhood injury prevention in a suburban Massachusetts population. Public Health Rep 1991;106: 437–42.
34. Guyer B, Gallagher SS, Chang BH, et al. Prevention of childhood injuries: evaluation of the Statewide Childhood Injury Prevention Program (SCIPP). Am J Public Health 1989;79:1521–7.
35. Armitage P, Berry G. Statistical methods in medical research. London: Blackwell, 1987.
36. Brillinger DR. The natural variability of vital rates and associated statistics. Biometrics 1986; 42:693–734.
37. Wolfe RA, Petroni GR, McLaughlin CG, et al. Empirical evaluation of statistical models for counts or rates. Stat Med 1991;10:1405–16.
38. Robertson LS. Injury epidemiology. New York: Oxford University Press, 1992.
39. Schwartz J. Multinational trends in cancer mortality rates: methodological issues and results. Ann N Y Acad Sci 1990;609:136–45.