J. Hyg., Camb. (1977), **78**, 135
Printed in Great Britain

Path analysis: application in an epidemiological study of echinococcosis in New Zealand

BY M. J. BURRIDGE,* C. W. SCHWABE
Department of Epidemiology and Preventive Medicine, School of Veterinary Medicine, University of California, Davis, California 95616, *U.S.A.*

AND T. W. PULLUM
Department of Sociology, University of California, Davis, California 95616, *U.S.A.*

(*Received* 29 *June* 1976)

SUMMARY

The method of path analysis is described in detail. Application of this analytical technique in the interpretation of causal relationships in complex biological systems is demonstrated using data from an epidemiological study of echinococcosis in New Zealand. The results identified the major causal pathways determining *Echinococcus granulosus* prevalence in dogs, based on multiple regression analysis of a linear causal model constructed from prior biological and epidemiological knowledge. Only ethnic and sheep husbandry variables had important direct effects on prevalence in the North Island of New Zealand, with some climatic factors (maximum temperature and relative humidity) and soil porosity acting indirectly through animal husbandry practices. It is suggested that path analysis, by permitting interaction between epidemiological theory and statistical analysis, provides a valuable additional tool to epidemiologists for the study of causal relationships among variables in multivariate systems.

INTRODUCTION

Path analysis is a method for the decomposition and interpretation of relationships among variables in linear causal models using multiple regression procedures. It was first developed over 50 years ago as an aid to the quantitative understanding of population genetics (Wright, 1921). However, little use was made of path analysis until it was introduced to the social sciences by Duncan (1966). Since then it has proved to be a useful approach to quantifying and interpreting causal theory in sociology. Path analysis has received little attention by epidemiologists. Kalimo & Bice (1973) examined the application of path analysis to cross-national epidemiological data on physician utilization, and Goldsmith & Berglund (1974) outlined the potential value of the technique in epidemiological studies of pulmonary diseases.

* Present address: Associate Professor of Epidemiology, College of Veterinary Medicine, Box J-125, JHMHC, University of Florida, Gainesville, Florida 32610, U.S.A. (address for reprints).

Many epidemiological studies, like those in sociology, utilize non-experimental observational data involving many variables. Epidemiologists have relied heavily on conventional multivariate techniques such as multiple regression to analyse complex data. In multiple regression analysis, epidemiological judgement is limited to the selection of 'independent' variables which are believed to influence the dependent variable of interest. Furthermore, using stepwise procedures, the order in which the independent variables enter the regression equation is decided purely on statistical grounds. While these conventional regression techniques yield valuable information, they do not interpret possible causal relationships. In contrast, path analysis does approach the problem of causal interpretation and it provides information about indirect as well as direct effects on a dependent variable. It allows scientists to use their biological knowledge of the system under consideration by sequentially ordering the variables in a linear causal model which represents the causal processes assumed to operate among the variables in nature. It is believed, therefore, that path analysis has an intuitive appeal to epidemiologists and other workers in the health sciences. It permits analysis of complex multivariate systems through manipulation of data in a biologically, as well as statistically, sound manner, enabling quantification of the strength of assumed causal relationships among variables. However, it must be emphasized that path analysis is dependent on the availability of sufficient epidemiological knowledge to construct realistic causal models and that the direction of assumed causal effects is determined entirely from an understanding of the biological processes under study. Furthermore, the method depends on the same statistical assumptions as multiple regression. These assumptions, such as linearity and homogeneity of variance, are difficult to check because of the complexity of most biological systems and, therefore, the results of path analysis should be regarded as exploratory and suggestive.

A national hydatid disease control programme was initiated in New Zealand in 1959, with control measures that have been described by Laing (1961). Data on the hydatid control effort in New Zealand, collected by one of the authors (C.W.S.) during World Health Organization consultantships to the New Zealand Medical Research Council in 1961 and the National Hydatids Council in 1965, formed the basis for a general descriptive account of that programme (Schwabe, 1969). However, since no attempt was made on those two occasions to collect detailed quantitative data, the senior author visited New Zealand in 1975 to carry out an epidemiological analysis of their hydatid disease control programme. One objective of that study was to determine those factors responsible for the great variation in *Echinococcus granulosus* prevalence in dogs in different areas of New Zealand at the beginning of the national control programme. Path analysis was chosen as a potentially useful method of investigating that portion of the study. This paper describes in detail the application of path analysis to epidemiology, using this study of canine echinococcosis in New Zealand as an example, and discusses the results in terms of both the factors affecting the prevalence of *E. granulosus* in dogs and the utility of path analysis for epidemiological purposes. Other aspects of the epidemiological analysis of the New Zealand programme have been published elsewhere (Burridge & Schwabe, 1977*a*, *b*, *c*; Burridge, Schwabe & Fraser, 1977).

MATERIALS AND METHODS

Areas studied

New Zealand was divided initially into 85 areas for hydatid control purposes, each administered by a local authority, and for this study all 72 rural areas on the two main islands were considered, 44 in the North Island and 28 in the South Island. The two major islands of New Zealand were examined separately because of important demographic and agricultural differences: 96 % of all Maoris and 89 % of all dairy farms were found on the North Island.

Final dependent variable

The final dependent variable in this study was the initial prevalence of *E. granulosus* in dogs by control area, as measured by the percentage of canine faecal samples found infected with *E. granulosus* during the first complete testing round of the control programme in 1959–1960. All dogs in New Zealand were tested for infection by dosing with the taeniafuge arecoline hydrobromide and all faecal samples thus collected were examined for cestodes at a central diagnostic laboratory. The mean area prevalence was 3·88 % for both islands, but its variability was greater on the North than the South Island (ranges of 0·88–10·74 % and 1·34–7·03 %, respectively).

Exogenous variables

In path analysis the exogenous variables are those that are predetermined, that is, whose total variation is assumed to be caused by variables outside the set under consideration. No attempt is made to explain their variability.

The exogenous variables used in this study are described in Table 1. The six climatic variables were chosen primarily for their possible effects on the viability of *E. granulosus* eggs which have been shown to be susceptible to exposure to bright sunlight (Ross, 1929), to high temperatures (Colli & Williams, 1972), to desiccation (Ross, 1929; Laws, 1968) and to dispersal by wind (Sweatman & Williams, 1963). Altitude was included for its effect on animal husbandry practices, and soil porosity for its effect on the availability of *E. granulosus* eggs for intermediate hosts, high porosity soils allowing eggs to be carried deep into the ground, especially during periods of heavy rainfall (Sweatman & Williams, 1963). The soils were classified into two broad categories depending on their porosity, with the high porosity group comprising the volcanic soils (pumice soils and loams), coastal sands, and recent alluvial soils. The Maoris were an important ethnic group to consider since they had a race-specific surgical incidence rate of hydatid disease six times that of non-Maoris (Burridge & Schwabe, 1977*a*). Feral pigs and goats were selected as possible alternatives to sheep as food sources for dogs, pig hunting in particular being a favourite sport of New Zealanders. The relative density of these animals by area was ranked from zero for absent to 7 for heaviest density, using the data of Wodzicki (1950).

Table 1. *Variables used in path analysis of canine prevalence of* Echinococcus granulosus *in New Zealand*

Name of variables	Description of variables
Exogenous variables	
Max. Temp./Year	Mean daily maximum temperature in °C for complete year
Max. Temp./Month	Mean daily maximum temperature in °C for hottest month only
Relative Humidity	Mean daily % relative humidity
Wind	Mean daily run of wind in km
Sun	Annual hours of bright sunshine
Rain	Annual mm. of rainfall
Altitude	% of area above 300 m
Soil	% of soils of high porosity
Maori	% Maoris in human population
Feral Pigs	Relative population density of feral pigs
Feral Goats	Relative population density of feral goats
Endogenous variables	
Pasture	% of grazing area composed of unimproved native grasses
Sheep Production	% of farms with sheep production as major source of income
Area per Farm	Mean grazing area per farm in hectares
Sheep per Farm	Mean number of sheep per farm
Sheep Density	Number of ovine livestock equivalents per hectare of grazing land
Rabbits	Relative population density of rabbits
Dogs per Farm	Mean number of working dogs per farm
Dogs per Owner	Mean number of working dogs per working-dog owner
Economics	% of employed rural males earning less than the national median income for that group
Initial *E. granulosus* Prevalence in Dogs*	% of dogs found infected with *E. granulosus* during first testing round in 1959–60

* Final dependent variable.

Endogenous variables

Endogenous variables are those whose variation is assumed to be determined by some linear combination of the variables under consideration; all are ultimately determined by the exogenous variables in the system. The endogenous variables used in this study are described in Table 1. The life-cycle of *E. granulosus* is maintained by a dog–sheep cycle and, therefore, dog, sheep and farm variables were important to consider. The pasture variable was chosen because the dense sward of improved grazing land would provide *E. granulosus* eggs with more protection from climatic factors than the clumps of native tussock grasses. The sheep density variable was estimated on the basis of the relative feed requirements of cattle and sheep, assuming that one unit of grazing land could support alternatively one cow/steer or 8 ewes (i.e. 1 cow or steer = 8 livestock equivalents = 8 ewes). Rabbits were selected as another alternative food source for dogs and their relative density by area was estimated by the ranking procedure described for feral pigs and goats, again using the data of Wodzicki (1950). The economics variable was the best one available by area to reflect the relative economic status of the rural population and was considered for the possible effects of income on ability to alter dog-feeding habits.

Linear causal model

Path analysis requires the construction of a linear causal model, written as a set of structural equations representing the causal processes assumed to operate among the variables in nature. It is assumed that each relationship in the model is linear and that the model is recursive, meaning that there are no reciprocal effects or feedback loops. The model used in this study is represented diagrammatically in Fig. 1 in simplified form. Each exogenous variable enters the causal pathway as indicated by a unidirectional arrow leading from it, whereas each endogenous variable enters the pathway sequentially from left to right. Once entered, each variable is assumed to have an effect on every other endogenous variable added subsequent to it; many of these paths are omitted from Fig. 1 for clarity of presentation. Intercorrelations between exogenous variables (conventionally represented by curved double-headed arrows) and paths from the disturbance variables are not shown for the same reason. Each disturbance variable represents all the unmeasured and residual causes of an individual endogenous variable which are not explicitly identified in the model; the disturbance variables are assumed to be uncorrelated with each other and with the exogenous variables.

Three exogenous variables were not entered at the beginning of the causal pathway. Maoris were considered to have an effect only on the dog, economic and prevalence variables. Feral pig density was known to affect local dog populations as pig hunters own packs of special pig dogs (Sweatman & Williams, 1962). Feral goat density was not thought to affect any variable except the prevalence of canine echinococcosis. Unlike the bush-dwelling feral pigs and goats, rabbits were common on grazing land and their population density might have been influenced by animal husbandry variables; also dogs specifically for rabbitting were distributed throughout New Zealand (Sweatman & Williams, 1962). Consequently, the rabbit variable was considered endogenous, and was placed between the animal husbandry and canine variables in the causal chain.

Procedures of path analysis

The mathematical aspects of path analysis have been described in detail by Land (1969) and will not be repeated herein. However, mention will be made of the derivation of path coefficients. Path coefficients in a recursive model are identical with standardized regression coefficients. They can be generated simply by regressing each endogenous variable on those other variables having direct paths to it. Therefore, a change of one standard deviation in a given independent variable, when all other independent variables in the regression equation are held constant, will alter the dependent variable by an amount equal to the product of its own standard deviation and the path coefficient. In other words, each path coefficient measures the magnitude of the direct effect of one variable upon another.

The computer programme used for this study was the multiple regression procedure of the Statistical Package for the Social Sciences (Nie *et al.* 1975). Initially, the first path in the causal model (Fig. 1) was examined by regressing Pasture on

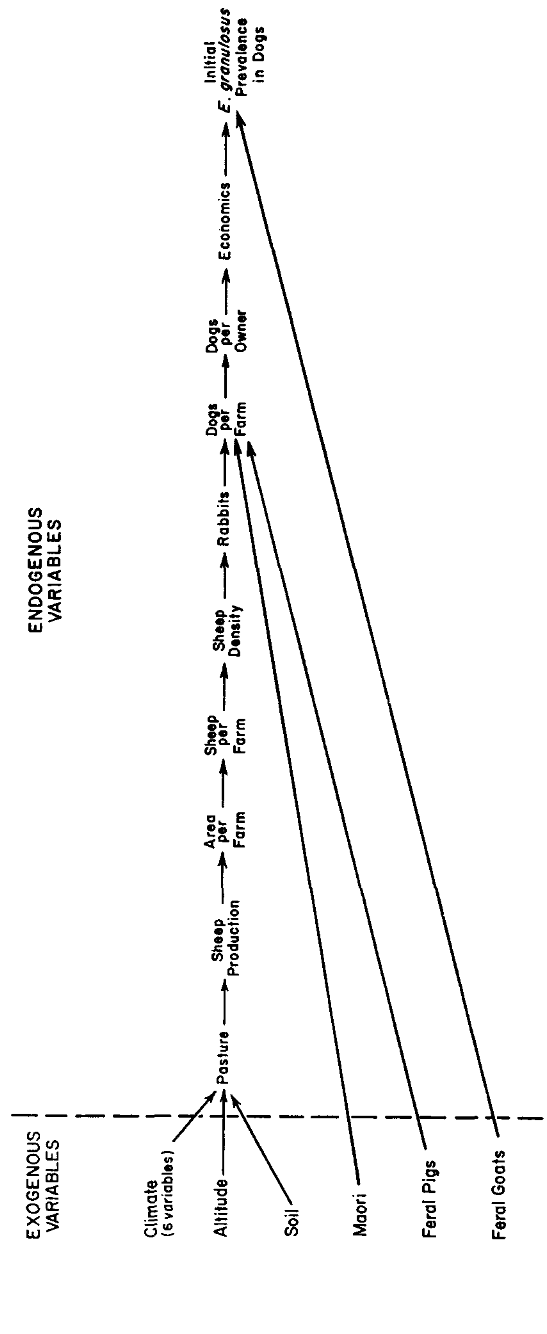

Fig. 1. Basic linear causal model of infection of dogs with *Echinococcus granulosus* in New Zealand.

https://doi.org/10.1017/S0022172400056047 Published online by Cambridge University Press

Altitude, Soil, and the six climatic variables. From that multiple regression, path coefficients were obtained from each independent variable to the dependent variable Pasture. The next intervening variable in the model, Sheep Production, was then added as the dependent variable and was regressed on all antecedent variables (i.e. Altitude, Soil, Climate and Pasture). This procedure was repeated, adding intervening variables individually in sequence, until the final endogenous variable, Initial *E. granulosus* Prevalence in Dogs, was added as the dependent variable. All the variables in the model were present in the final multiple regression equation and from that analysis were obtained additional items of information, the mean and standard deviation of each variable and a correlation matrix showing the zero-order correlation coefficients for every pair of variables.

The complete series of multiple regression analyses gave path coefficients for every possible pathway in the model. Also, the coefficient of multiple determination (or squared multiple correlation coefficient) was computed for each regression equation, giving the proportion of the total variation of each endogenous (dependent) variable explained by the appropriate equation and providing a convenient measure of the relative importance of residual variables not included in each equation.

Interpretation of path analysis

The interpretation of the results of path analysis was based on the methods of Alwin & Hauser (1975) and Duncan (1975).

A zero-order correlation coefficient expresses the degree of linear relationship between two variables and it can be regarded as a measure of their total association, made up of 3 distinct components:

(*a*) *direct effect* being that effect not transmitted via intervening variables but remaining when all other variables have been held constant – measured by the path coefficient;

(*b*) *indirect effects* being those effects mediated through intervening variables – each given by the product of the path coefficients in the appropriate indirect pathways;

(*c*) '*spurious' correlations* being those correlations due to joint dependence on an antecedent variable (i.e. common 'cause') and to correlated exogenous variables.

The sum of direct and indirect effects of one variable on another variable is defined to be the total effect. Total association equals the sum of the total effect and spurious correlations.

The complete analysis for each island produced 134 direct effects (path coefficients) and many more indirect effects. Since this study examined entire populations, a rule-of-thumb method rather than a statistical test was required to identify the more important effects. It was decided to consider only a direct or indirect effect of at least $\pm 0{\cdot}30$ standard units as constituting a major causal effect in the model. Using this arbitrary criterion and the above method of decomposing effects, important causal pathways were identified in the North and South Island models.

Table 2. *Coefficients of multiple determination for regression equations used in path analysis of canine prevalence of* Echinococcus granulosus *in New Zealand*

Dependent variable	Percentage of total variation of variables explained by regression equations: North Island	South Island
Pasture	31·4	78·8
Sheep Production	79·5	78·9
Area per Farm	80·9	74·4
Sheep per Farm	95·1	92·1
Sheep Density	84·7	87·9
Rabbits	65·8	80·6
Dogs per Farm	96·8	93·8
Dogs per Owner	83·1	94·3
Economics	83·9	73·6
Initial *E. granulosus* Prevalence in Dogs	82·9	83·9

RESULTS

A high proportion of the total variation of all endogenous variables, except North Island Pasture, was explained by the regression equations used in the analyses (Table 2), indicating that the most important determinants of all but one of the endogenous variables were indeed measured.

North Island

The major causal pathways determining both initial prevalence and the intervening variables were identified by path analysis (Table 3). This allowed modification of the original model to show the pathways by which different variables acted either directly or indirectly to influence initial *E. granulosus* prevalence in dogs in the North Island (Fig. 2).

The results showed that only the ethnic, altitudinal and sheep husbandry variables had major direct effects on initial prevalence in the North Island, whereas some climatic factors (maximum temperature and relative humidity) and soil porosity acted indirectly through sheep husbandry practices. The canine, economic and feral animal variables played no causal role in the North Island model. All variables causally associated with initial prevalence were interrelated through a pattern of effects with the exception of the Maori factor whose direct causal effect was not influenced by any other variable in the model. A consequence of the interrelationships of variables along the causal pathway was the creation of suppressor effects; for example, the large negative direct effect of Sheep per Farm was counteracted by a number of positive indirect effects, producing a very small total effect for that variable.

Table 3. *Decomposition of effects determining the prevalence of* Echinococcus granulosus *in the rural dogs of the North Island of New Zealand using the method of path analysis*

Endogenous variable	Antecedent variables	Total effect†	Major individual effects† Direct	Indirect
Initial *E. granulosus* Prevalence in Dogs	Max. Temp./Month	+0·66	*	+0·46 via Sheep Production
	Max. Temp./Year	−0·63	*	−0·39 via Sheep Production
	Sheep Production	+0·59	+0·80	−0·37 via Sheep per Farm
	Maori	+0·44	+0·46	*
	Altitude	+0·37	+0·38	*
	Sheep Density	+0·30	+0·33	*
	Area per Farm	−0·18	*	−0·31 via Sheep per Farm
	Sheep per Farm	−0·08	−0·68	*
	Relative Humidity	−0·04	*	−0·34 via Sheep Production
Sheep Density	Altitude	−0·87	*	−0·36 via Area per Farm
	Max. Temp./Year	−0·53	−0·57	*
	Sheep per Farm	+0·50	+0·50	*
	Area per Farm	−0·47	−0·70	*
	Sheep Production	−0·41	−0·44	*
	Pasture	−0·38	−0·39	*
	Max. Temp./Month	+0·09	+0·41	*
Sheep per Farm	Sheep Production	+0·79	+0·55	*
	Max. Temp./Month	+0·63	*	+0·31 via Sheep Production
	Area per Farm	+0·45	+0·45	*
Area per Farm	Altitude	+0·61	+0·52	*
	Sheep Production	+0·53	+0·53	*
	Max. Temp./Month	+0·51	*	+0·30 via Sheep Production
Sheep Production	Max. Temp./Month	+0·57	+0·57	*
	Max. Temp./Year	−0·50	−0·49	*
	Relative Humidity	−0·42	−0·42	*
	Soil	−0·32	−0·34	*
Pasture	Altitude	+0·57	+0·57	*
	Soil	−0·38	−0·38	*

* Individual effects of < ±0·30 standard units.
† Effects measured in standard units.

South Island

The pattern of causal effects was more complex for the South Island with most of the variables having a major direct effect on initial prevalence and many additionally having substantial indirect effects (Table 4). The total effects of Pasture, Economics and Feral Pigs and Goats were due solely to direct causal effects. Also, there were a number of suppressor effects, especially evident with Area per Farm and some climatic factors (Sun and Maximum Temperature). Only five variables (Maori, Sheep Density, Dogs per Farm, Relative Humidity and Wind) had no major effects on the initial *E. granulosus* prevalence in South Island dogs.

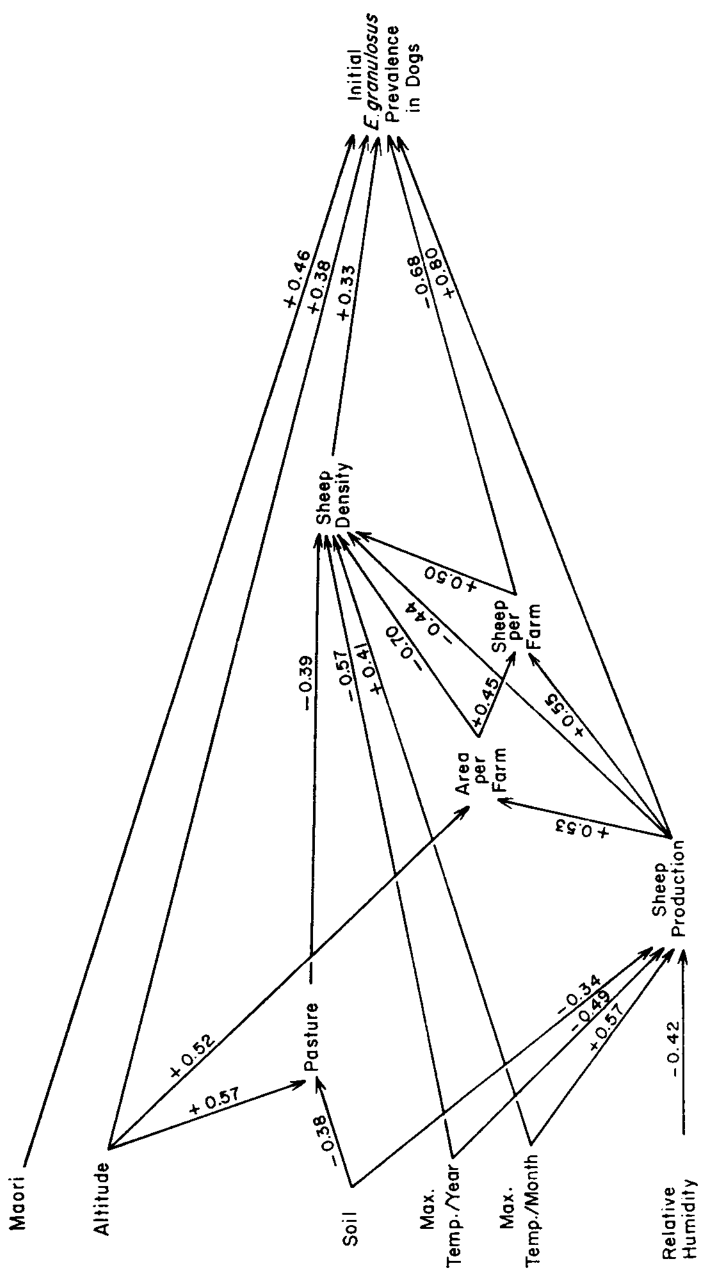

Fig. 2. Modified linear causal model of infection of dogs with *Echinococcus granulosus* in the North Island of New Zealand, with path coefficients shown for each causal pathway.

Table 4. *Decomposition of effects determining the prevalence of* Echinococcus granulosus *in the rural dogs of the South Island of New Zealand using the method of path analysis*

Antecedent variables	Total effect†	Major individual effects† Direct	Indirect
Soil	−0·76	−0·48	*
Pasture	−0·71	−0·68	*
Feral Pigs	−0·71	−0·67	*
Sheep Production	+0·55	+0·81	*
Dogs per Owner	+0·44	+0·31	*
Economics	−0·43	−0·43	*
Sheep per Farm	−0·34	−0·49	−0·33 via Economics
Feral Goats	+0·31	+0·31	*
Relative Humidity	−0·31	*	*
Altitude	−0·29	*	−0·47 via Pasture
Rabbits	+0·27	*	+0·31 via Economics
Sheep Density	+0·26	*	*
Maori	+0·20	*	*
Max. Temp./Year	+0·19	+0·52	−0·40 via Area per Farm
Area per Farm	+0·18	+0·58	−0·32 via Sheep per Farm
Max. Temp./Month	−0·17	−0·93	+0·36 via Sheep Production; +0·34 via Area per Farm
Wind	+0·13	*	*
Rain	−0·12	*	−0·33 via Sheep Production
Sun	+0·08	+0·32	−0·33 via Economics
Dogs per Farm	+0·03	*	*

* Individual effects of < ±0·30 standard units.
† Effects measured in standard units.

DISCUSSION

Conclusions drawn from the results of path analysis, as with all multivariate statistical techniques, should be made with an appreciation of measurement and sampling errors in the data. Sampling errors were of no concern in this study since the data were taken from entire populations. However, measurement errors were present, particularly in the estimation of the final dependent variable. The prevalence of *E. granulosus* in dogs was measured using the taeniafuge arecoline hydrobromide as a diagnostic agent, the best available method of diagnosis of canine echinococcosis although one of poor sensitivity. Under field conditions in New Zealand and Australia, arecoline hydrobromide failed to purge more than a fifth of the dogs dosed with 3·5 mg/kg body weight (Gemmell, 1958, 1968; Jackson & Arundel, 1971; Gregory, 1973). Even when purgation was induced, some infected dogs retained their worm burdens (Gemmell, 1973). Owing to that source of error, only major causal pathways identified by path analysis were considered of significance in determining the *E. granulosus* prevalence in dogs at the beginning of the New Zealand hydatid disease control programme in 1959–60.

There were two major determinants of *E. granulosus* prevalence in the North Island dogs, one ethnic involving the Maori people and the other relating to sheep husbandry. The positive effect of Maoris on prevalence was direct and independent

https://doi.org/10.1017/S0022172400056047 Published online by Cambridge University Press

of the other variables. We have indicated in another study that the high *E. granulosus* prevalence rates in dogs owned by Maoris were due to a combination of poor dog control, cultural factors, and a complex system of land tenure that retarded the development of Maori farming methods (Burridge & Schwabe, 1977*a*). It was to be expected that sheep husbandry factors would influence canine prevalence since the life-cycle of *E. granulosus* was maintained in New Zealand by a dog–sheep cycle. However, the results from the North Island indicated that the specific husbandry factors favouring high prevalence rates were small farms, both in terms of size and numbers of sheep, and dense grazing. The direct effect of altitude on prevalence probably reflected the predominance of sheep production at higher altitudes.

The South Island results were more complex, without a clear picture emerging for the construction of a simplified causal model. Twelve variables had major direct effects on *E. granulosus* prevalence and all substantial indirect effects were mediated through one of those variables. Nevertheless, examination of the direction and magnitude of the largest effects permitted interpretation of the most important relationships determining initial prevalence. The sheep husbandry variables had effects similar to those seen on the North Island. The large negative effects of pasture type and soil porosity were primarily direct and probably due to their influences on the viability and availability of *E. granulosus* eggs. Clumps of native grasses provided eggs with less protection from detrimental climatic factors than the dense sward of improved grasses, and highly porous soils allowed eggs to be carried deep into the ground and away from grazing intermediate hosts. Another variable with a large negative direct effect on canine prevalence was feral pig density. About 30000 feral pigs were killed annually in New Zealand and the livers were often fed raw to pig dogs after the hunt (Sweatman & Williams, 1962). Surveys of wild pigs in New Zealand have shown that they were rarely infected with *E. granulosus*, with only one of 514 animals examined found to harbour a hydatid cyst (Ineson, 1954; Sweatman & Williams, 1962). It would appear possible, therefore, that in areas of heavy feral pig density, those wild animals formed an alternative to sheep as a canine food source, considerably reducing the likelihood of infection of dogs with *E. granulosus*.

One of the main factors limiting the success of path analysis in simplifying the causal pathways in the South Island model was the relative homogeneity of much of the region, especially regarding the Maori and Sheep Production variables that were of major causal significance in the North Island. In addition, initial canine prevalence showed less variability by area in the South Island.

The identification of important pathways in a linear causal model is not the only value of path analysis. This technique also provides a method whereby the consequences of realistic manipulation of variables can be visualized, forming in the context of preventive medicine a rational basis for designing a control programme. With reference to the New Zealand example, measures to control echinococcosis would aim at maximizing negative effects on *E. granulosus* prevalence in dogs and minimizing positive effects. Manipulation of the Sheep Production variable in the North Island model (Fig. 2) – for example, by replacing sheep with beef or dairy cattle – would lead directly to a reduction in canine prevalence, but indirectly to an increase in prevalence through its indirect effects mediated via the intervening variables of Sheep Density and Sheep per Farm. The total effect on prevalence of such a change in the pattern of livestock production would depend on accompanying changes in the intervening variables, which themselves could be estimated from the path coefficients in the model. Whilst this example might not be realistic in the New Zealand situation, it does serve to illustrate how path analysis can help to predict the outcome of various control strategies.

Epidemiological studies in parasitology often require investigation into parasite life-cycles that would be represented in a causal model by a feedback loop. It has been stated that such non-recursive models cannot be examined using path analysis. However, Heise (1969) has pointed out that feedback can be treated in a recursive model under certain specific conditions, namely that the feedback in the system is always a delayed effect, allowing the feedback loop to be represented in the model as a chain of lagged variables. For example, consider the feedback relationship between prevalence of a parasite in definitive hosts ($= x$) and in intermediate hosts ($= y$). Such a relationship could be treated as the temporal sequence $x_{t_1} \rightarrow y_{t_2} \rightarrow x_{t_3}$, where variables x and y are measured at sequential time intervals t_1, t_2 and t_3. To use this procedure, it must be assumed that the residual causes of x (i.e. those causes not identified in the model) do not remain stable over the time interval t_1 to t_3. In other words, the residuals in x_{t_3} must be assumed to be uncorrelated with those in x_{t_1} so that the underlying assumptions concerning residuals are not violated. Using this technique, the factors affecting the prevalence of *E. granulosus* in both the definitive hosts (dogs) and intermediate hosts (primarily sheep) could be examined in the same causal model. That was not possible in this study owing to the lack of prevalence data for intermediate hosts by area. Other situations in which feedback loops and reciprocal effects may be included in causal models, have been described by Duncan (1975).

It is seen that path analysis adds an important dimension to the epidemiological analysis of multivariate systems by approaching the question of causal interpretation in both a biologically and statistically sound manner through construction of realistic causal models of the systems under study. Path analysis should be intuitively appealing to epidemiologists since it allows biological knowledge and epidemiological judgement to be utilized in conjunction with statistical finesse. Areas in which this method would appear to have immediate application are not only studies of systems where the basic underlying causal processes are understood, but also research into conditions of complex aetiology, such as neonatal mortality in various animal species, where a natural temporal sequence of events is evident. It is hoped that in these and other areas of epidemiological endeavour, path analysis will prove a valuable additional tool to health scientists.

The authors are indebted to the National Hydatids Council of the New Zealand Ministry of Agriculture and Fisheries for permission to study their national hydatid disease control programme, and in particular to George A. Thomson, Superintendent (Hydatids Control), for his kind assistance with the canine data.

Thanks are also due to Mr J. D. Coulter of the New Zealand Meteorological Service for help with the climatic data and to the New Zealand Department of Statistics for provision of all other data.

This study was supported in part by research grant 2 RO1 AI07857–10 from the National Institutes of Health, U.S.A., and a grant from the World Health Organization.

REFERENCES

ALWIN, D. F. & HAUSER, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review* **40**, 37–47.

BURRIDGE, M. J. & SCHWABE, C. W. (1977*a*). Hydatid disease in New Zealand: an epidemiological study of transmission among Maoris. *American Journal of Tropical Medicine and Hygiene* **26**, (In the Press).

BURRIDGE, M. J. & SCHWABE, C. W. (1977*b*). Epidemiological analysis of factors influencing rate of progress in *Echinococcus granulosus* control in New Zealand. *Journal of Hygiene* **78, 151.**

BURRIDGE, M. J. & SCHWABE, C. W. (1977*c*). Epidemiological analysis of factors influencing the increase in *Taenia ovis* prevalence during the New Zealand *Echinococcus granulosus* control program. *Australian Veterinary Journal* **53**, (In the Press).

BURRIDGE, M. J., SCHWABE, C. W. & FRASER, J. (1977). Hydatid disease in New Zealand: changing patterns of human infection, 1878–1972. *New Zealand Medical Journal* **85**, (In the Press).

COLLI, C. W. & WILLIAMS, J. F. (1972). Influence of temperature on the infectivity of eggs of *Echinococcus granulosus* in laboratory rodents. *Journal of Parasitology* **58**, 422–6.

DUNCAN, O. D. (1966). Path analysis: sociological examples. *American Journal of Sociology* **72**, 1–16.

DUNCAN, O. D. (1975). *Introduction to Structural Equation Models*. New York: Academic Press.

GEMMELL, M. A. (1958). Arecoline hydrobromide as a taeniafuge in dogs, with special reference to its use in controlling hydatid disease. *Australian Veterinary Journal* **34**, 207–12.

GEMMELL, M. A. (1968). The Styx field-trial: a study on the application of control measures against hydatid disease caused by *Echinococcus granulosus*. *Bulletin of the World Health Organization* **39**, 73–100.

GEMMELL, M. A. (1973). Surveillance of *Echinococcus granulosus* in dogs with arecoline hydrobromide. *Bulletin of the World Health Organization* **48**, 649–52.

GOLDSMITH, J. R. & BERGLUND, K. (1974). Epidemiological approach to multiple factor interactions in pulmonary disease: the potential usefulness of path analysis. *Annals of the New York Academy of Sciences* **221**, 361–75.

GREGORY, G. G. (1973). A survey of tapeworms in rural dogs in Tasmania. *Australian Veterinary Journal* **49**, 273–8.

HEISE, D. R. (1969). Problems in path analysis and causal inference. In *Sociological Methodology* (ed. E. F. Borgatta), pp. 38–73. San Francisco: Jossey-Bass Inc.

INESON, M. J. (1954). A comparison of the parasites of wild and domestic pigs in New Zealand. *Transactions of the Royal Society of New Zealand* **82**, 579–609.

JACKSON, P. J. & ARUNDEL, J. H. (1971). The incidence of tapeworms in rural dogs in Victoria. *Australian Veterinary Journal* **47**, 46–53.

KALIMO, E. & BICE, T. W. (1973). Causal analysis and ecological fallacy in cross-national epidemiological research. *Scandinavian Journal of Social Medicine* **1**, 17–24.

LAING, A. D. M. G. (1961). The implementation of hydatid eradication measures in New Zealand. *Bulletin de l'Office International des Épizooties* **56**, 1030–9.

LAND, K. C. (1969). Principles of path analysis. In *Sociological Methodology* (ed. E. F. Borgatta), pp. 3–37. San Francisco: Jossey-Bass Inc.

LAWS, G. F. (1968). Physical factors influencing survival of taeniid eggs. *Experimental Parasitology* **22**, 227–39.

NIE, N. H., HULL, C. H., JENKINS, J. G., STEINBRENNER, K. & BENT, D. H. (1975). *Statistical Package for the Social Sciences*, 2nd ed. New York: McGraw-Hill.

ROSS, I. C. (1929). Observations on the hydatid parasite (*Echinococcus granulosus*) and the control of hydatid disease in Australia. *Bulletin of the Commonwealth of Australia Council for Scientific and Industrial Research*, no. 40. Melbourne: H. J. Green.

SCHWABE, C. W. (1969). *Veterinary Medicine And Human Health*, 2nd ed. Baltimore: Williams & Wilkins Co.

SWEATMAN, G. K. & WILLIAMS, R. J. (1962). Wild animals in New Zealand as hosts of *Echinococcus granulosus* and other taeniid tapeworms. *Transactions of the Royal Society of New Zealand (Zoology)* **2**, 221–50.

SWEATMAN, G. K. & WILLIAMS, R. J. (1963). Survival of *Echinococcus granulosus* and *Taenia hydatigena* eggs in two extreme climatic regions of New Zealand. *Research in Veterinary Science* **4**, 199–216.

WODZICKI, K. A. (1950). Introduced mammals of New Zealand: an ecological and economic survey. *Bulletin of the New Zealand Department of Scientific and Industrial Research*, no. 98.

WRIGHT, S. (1921). Correlation and causation. *Journal of Agricultural Research* **20**, 557–85.