Structural Equation Analysis: The Decomposition of Effects in a Nonadditive Causal Model of Canine Echinococcosis in Cyprus

Marguerite Pappaioanou*, Kyprianos Polydorou#, Calvin W. Schwabe**

The methods of structural equation analysis were used to evaluate a causal model of canine echinococcosis in Cyprus. To account for biologically important interactions, a nonadditive model of parasite transmission for the dog-sheep strain of *Echinococcus granulosus* in Cyprus was constructed.

Causal models have commonly been constructed such that the relationships between variables are assumed to be additive, i.e. the effect of one variable on a second variable is assumed to be independent of the values of all other variables in the model. This assumption is often made to simplify the interpretation of hypothesized causal relationships. However, it is recognized that host, agent, and environmental determinants of disease do not always act independently, but in fact interact with each other.

Procedures for calculating direct and indirect effects while accounting for interactions between variables have been described by Stolzenberg (1979). The use of these methods to analyze a specified model of canine echinococcosis in Cyprus is presented in order to demonstrate how more realistic causal models can be utilized in veterinary epidemiologic research.

MATERIALS AND METHODS

Variables reflecting the canine, livestock, demographic, and environmental aspects of the epidemiology of canine echinococcosis were chosen for use in the analysis and are described in Table 1. A causal ordering of the variables was determined by consideration of the parasite's life cycle and a time sequence of factors. Hypothesized causal relationships were expressed as a series of structural equations. In a structural equation, a dependent (effect) variable is written as the sum of independent (causal) variables multiplied by their respective structural parameters.

*Assistant Professor of Epidemiology & Statistics, and **Professor of Epidemiology, Department of Epidemiology & Preventive Medicine, University of California, Davis, CA 95616; #Director of the Department of Veterinary Services, Republic of Cyprus

When the assumptions of multiple regression have not been severely violated, and if the error terms can be assumed to be uncorrelated with the causal variables in the equation, ordinary least squares procedures used in multiple regression analysis can be used to estimate the structural parameters (Kenny 1979; Duncan 1975). The system of 17 structural equations specified in the model is presented in Table 2. The Biomedical Computer Programs P-Series 1R Multiple Linear Regression Program (Dixon & Brown 1979) was used to estimate each of the equations in the model.

In nonadditive models, interactions between variables are commonly represented by the inclusion of product terms in the appropriate structural equations. For example, in Structural Eq. (1) (See Table 2) product terms were used to indicate interactions between the following variables (See Table 1 for definitions): 1) SHEEP, GOATS, CATTLE, DAIRY and PIGS, respectively with DOGS72; 2) SHEEPDEN, GOATDEN and CATTLEDEN, respectively, with SOIL and RAIN72A (interactions used to account for infected livestock status); and 3) K9PHUM72 with ROOMS (ROOMS was used as an indicator of relative income status). A detailed description of the methods of data collection, variables used and model specification is presented by Pappaioanou et al. (1982).

Methods for Calculating Total, Direct and Indirect Effects

Illustration: The procedures used to calculate the total, direct and indirect effects of DOGS72 (X_1) on the dependent variable EGPREV72 (Y) are described below.

Structural Eq. (1) was written as (See Tables 1 and 2):

$$Y = b_0+b_1X_1+\sum_{i=2}^{7} b_iX_i+b_8(X_1X_3)+b_9(X_1X_4)+b_{10}(X_1X_5)+b_{11}(X_1X_6)+ b_{12}(X_1X_7)+b_{13}X_8+b_{14}X_9+b_{15}X_{10}+b_{16}X_{11}+b_{17}X_{12}+b_{18}(X_8X_{11}X_{12})+ b_{19}(X_9X_{11}X_{12})+b_{20}(X_{10}X_{11}X_{12})+b_{21}X_{13}+b_{22}X_{14}+b_{23}X_{13}X_{14}+E_1$$

where b_i, i = 1,2,3,.....,23 are the structural parameters; and E_1 is the error term.

The direct effect in nonstandardized units of DOGS72 on EGPREV72 is obtained by ignoring the indirect pathways of causation in the model [i.e. ignoring Structural Eqs. (2)-(17)] and calculating the partial derivative [dY/dX_1].

$$dY/dX_1 = b_1+b_8X_3+b_9X_4+b_{10}X_5+b_{11}X_6+b_{12}X_7$$

The direct effect is then calculated by multiplying the respective structural parameters (regression coefficients) by a specified level of the interacting causal variables. For this analysis, the mean levels of interacting variables were used in the calculations, although values at one standard deviation above or below the mean could also be used.

Example: Results of the multiple regression analysis of Structural Eq. (1) are shown in Table 3. The direct effect of DOGS72 on EGPREV72 is calculated in the following manner:

$$dY/dX_1 = -0.00093+(-2.24 \times 10^{-7})(591.84)+(4.06 \times 10^{-7})(389.41)$$
$$+(7.99 \times 10^{-7})(19.97)+(-3.50 \times 10^{-6})(19.27)$$
$$+(1.57 \times 10^{-6})(36.86) = -0.000898 \simeq -0.0009$$

The interpretation of this value can be stated as, all other variables held constant, at the average number of village sheep, goats, cattle, dairy cows and pigs, increasing the dog population by one dog causes a decrease in the prevalence of *E. granulosus* in dogs on the average by 0.0009 units.

The indirect effect of X_1 on Y through intermediate variables (X_i) can be obtained via the chain rule of calculus and calculating the sum of products $\sum_{i \neq 1} [dY/dX_i][dX_i/dX_1]$. Since DOGS72 indirectly affects EGPREV72 only through K9PHUM72, this expression simplifies to $[dY/dX_{14}][dX_{14}/dX_1]$. Structural Eq. (3) was written as (see Table 2):

$$X_{14} = a_0+a_1X_1+a_2X_{15}+a_3X_{13}+a_4X_{16}+E_{14}$$

where a_i, i = 1,2,3,4 are the structural parameters; and E_{14} is the error term.

Example: Using the chain rule, the indirect effect of DOGS72 on EGPREV72 via K9PHUM72 is calculated in the following manner [see Table 3 for the results of the multiple regression analysis of Structural Eq. (3)]:

$$[dY/dX_{14}][dX_{14}/dX_1] = [(b_{22}+b_{23}X_{13})a_1]$$
$$= [(-0.90861+1.78508(0.824))0.00096]$$
$$= 0.00054$$

The total effect of DOGS72 on EGPREV72 was calculated by adding the direct and indirect effects (-0.0009+0.00054 = -0.00036).

RESULTS AND DISCUSSION

The calculated total, direct and indirect effects of each causal variable on EGPREV72 for Greek Cypriot villages are presented in Table 4. The average number of dogs per person (K9PHUM72) had the largest direct and total effect on the dependent variable. The implication of this and other findings for the Cyprus Anti-echinococcosis Campaign are discussed by Pappaioanou et al. (1982).

Variables which have only minimal causal effects at the average value of interacting factors could have major effects at high or low values of co-determinants. Therefore, the causal relationships between variables in a nonadditive model cannot be totally evaluated until the effects of variables, calculated at low and high values of interacting variables, are compared. A computer program is being written (Pappaioanou and Dilgard, in preparation) to facilitate the calculations described in this paper

REFERENCES

1. Dixon WJ. and Brown MB. 1979. BMDP-79: Biomedical Computer Programs--P-Series. Berkeley: University of California Press.

2. Duncan OD. 1975. Introduction to Structural Models. New York: Academic Press.

3. Kenny DA. 1979. Correlation and Causality. New York: John Wiley & Sons.

4. Pappaioanou M., Polydorou K., Schwabe CW. 1982. The effects of canine, livestock, environmental and demographic factors on the prevalence of Echinococcus graunlosus in Cypriot Dogs. (Submitted to the International Journal of Epidemiology).

5. Pappaioanou M. and Dilgard P. A computer program for the decomposition of causal effects in nonadditive structural equation models. (In preparation).

6. Stolzenberg RM. 1979. The measurement and decomposition of causal effects in nonlinear and nonadditive models, in Sociological Methodology 1980 (KF Schuessler, ed.). San Francisco: Jossey-Bass.

Table 1. Description of Variables Used in the Structural Analysis of Canine Echinococcosis in Cyprus

Dependent (Effect) Variable	Description
EGPREV 72(Y)	Proportion of dogs infected with *E. granulosus*, June-December, 1972[a]
Independent (Causal) Variables	
DOGS 72 (X_1)	Number of dogs tested for *E. granulosus*, June-December, 1972
SLTHSE 72(X_2)	Presence/Absence of approved slaughtering facilities June-December, 1972
SHEEP (X_3)	Mean number of sheep, autumn 1970-1973[b]
GOATS (X_4)	Mean number of goats, autumn 1970-1973[b]
CATTLE (X_5)	Mean number of native cattle, autumn 1970-1973[c]
DAIRY (X_6)	Mean number of dairy cows, autumn 1970-1973[c]
PIGS (X_7)	Pigs for fattening and boars & sows, 1973
SHEEPDEN (X_8)	Mean number of sheep per square mile grazing area
GOATDEN (X_9)	Mean number of goats per square mile grazing area
CATTLEDEN(X_{10})	Mean number of cattle per square mile grazing area
SOIL (X_{11})	Proportion of village land area not composed of Kafkala and Terra Rosa on Kafkala (nonporous) Soils
RAIN72A (X_{12})	Total rainfall (mm.) January-May, 1972
ROOMS (X_{13})	Proportion of village dwellings having one to three rooms
K9PHUM 72(X_{14})	Mean number of dogs tested for *E. granulosus* per village resident, June-December, 1972
FLOCKPHUM(X_{15})	Number of sheep & goat flocks with more than 50 animals per flock per village resident
HUMANPOP(X_{16})	Number of village residents, 1960[d]
FLOCKS	Number of sheep & goat flocks with more than 50 animals per flock
HUMANDEN	Mean human population per village square mile
GRAZE	Square miles of village land area used for grazing
SCRUB	Proportion of grazing area classified as scrubland
RAINAVG	Average village rainfall (mm.), 1941-1970
AREA	Village land area (square miles)

[a]Examination of feces by the pan flotation method following the administration of arecoline hydrobromide.

[b]Anthrax vaccinations, Department of Veterinary Services, Republic of Cyprus.

[c]Foot and Mouth Disease Vaccination, Department of Veterinary Services, Republic of Cyprus.

[d]1960 Census of Population and Agriculture, Republic of Cyprus.

Table 2. System of Structural Equations for the Analysis of Canine Echinococcosis in Cyprus, January-April, 1973

(1) ECPREV72 = a_1+b_1(DOGS72)+b_2(SLTHSE72)+b_3(SHEEP)+b_4(GOATS)+ b_5(CATTLE)+b_6(DAIRY)+b_7(PIGS)+b_8(SHEEP*DOGS)+ b_9(GOATS*DOGS)+b_{10}(CATTLE*DOGS)+b_{11}(DAIRY*DOGS)+ b_{12}(PIGS*DOGS)+b_{13}(SHEEPDEN)+b_{14}(GOATDEN)+ b_{15}(CATTLEDEN)+b_{16}(SOIL)+b_{17}(RAIN2A)+ b_{18}(SHEEPDEN*RAIN72A*SOIL)+b_{19}(GOATDEN*RAIN72A*SOIL)+ b_{20}(CATTLEDEN*RAIN72A*SOIL)+b_{21}(ROOMS)+b_{22}(K9PHUM72)+ b_{23}(K9PHUM*ROOMS)+E_1

(2) SLTHSE72 = a_2+b_{24}(HUMANPOP)+b_{25}+(ROOMS)+b_{26}(SHEEP)+b_{27}(GOATS)+ b_{28}(CATTLE)+b_{29}(DAIRY)+b_{30}(PIGS)+E_2

(3) K9PHUM72 = a_3+b_{31}(DOGS72)+b_{32}(FLOCKPHUM)+b_{33}(ROOMS)+b_{34}(HUMANPOP)+E_3

(4) DOGS72 = a_4+b_{35}(FLOCK)+b_{36}(HUMANPOP)+b_{37}(ROOMS)+E_4

(5) SHEEPDEN = a_5+b_{38}(SHEEP)+b_{39}(GRAZE)+b_{40}(SCRUB)+b_{41}(HUMANDEN)+E_5

(6) GOATDEN = a_6+b_{42}(GOATS)+b_{43}(GRAZE)+b_{44}(SCRUB)+b_{45}(HUMANDEN)+E_6

(7) CATTLEDEN= a_7+b_{46}(CATTLE)+b_{47}(GRAZE)+b_{48}(SCRUB)+b_{49}(HUMANDEN)+E_7

(8) FLOCKPHUM= a_8+b_{50}(SHEEP)+b_{51}(GOATS)+b_{52}(HUMANPOP)+b_{53}(ROOMS)+E_8

(9) FLOCK = a_9+b_{54}(SHEEP)+b_{55}(GOATS)+b_{56}(HUMANPOP)+b_{57}(HUMANDEN)+ b_{58}(ROOMS)+E_9

(10) SHEEP = a_{10}+b_{59}(HUMANPOP)+b_{60}(HUMANDEN)+b_{61}(GRAZE)+b_{62}(SCRUB)+ b_{63}(RAINAVE)+b_{64}(ROOMS)+E_{10}

(11) GOATS = a_{11}+b_{65}(HUMANPOP)+b_{66}(HUMANDEN)+b_{67}(GRAZE)+b_{68}(SCRUB)+ b_{69}(RAINAVE)+b_{70}(ROOMS)+E_{11}

(12) CATTLE = a_{12}+b_{71}(HUMANPOP)+b_{72}(HUMANDEN)+b_{73}(GRAZE)+b_{74}(SCRUB)+ b_{75}(RAINAVG)+b_{76}(ROOMS)+E_{12}

(13) DAIRY = a_{13}+b_{77}(HUMANPOP)+b_{78}(HUMANDEN)+b_{79}(GRAZE)+b_{80}(SCRUB)+ b_{81}(RAINAVG)+b_{82}(ROOMS)+E_{13}

(14) PIGS = a_{14}+b_{83}(HUMANPOP)+b_{84}(HUMANDEN)+b_{85}(RAINAVG)+b_{86}(ROOMS)+E_{14}

(15) HUMANDEN = a_{15}+b_{87}(AREA)+b_{88}(HUMANPOP)+E_{15}

(16) SCRUB = a_{16}+b_{89}(GRAZE)+b_{90}(RAINAVG)+E_{16}

(17) GRAZE = a_{17}+b_{91}(AREA)+b_{92}(RAINAVG)+E_{17}

Table 3. Results of Multiple Regression Analysis for Structural Equations (1) and (3)[a]

Dependent Variable	Independent Variable	Mean[b]	Standard Deviation[b]	Regression Coefficient[c]
EGPREV72(Y)		0.060	0.097	--
	DOGS72 (X_1)	17.852	20.866	-0.00093 (b_1)
	SLTHSE72 (X_2)	0.640	0.481	0.01577 (b_2)
	SHEEP (X_3)	591.841	918.767	0.00003 (b_3)
	GOATS (X_4)	389.413	464.606	-0.00003 (b_4)
	CATTLE (X_5)	19.969	28.142	0.00008 (b_5)
	DAIRY (X_6)	19.268	76.537	0.00017 (b_6)
	PIGS (X_7)	36.863	149.072	-0.00009 (b_7)
	SHEEP*DOGS72	NR	NR	$-2.24X10^{-7}$ (b_8)
	GOATS*DOGS72	NR	NR	$4.06X10^{-7}$ (b_9)
	CATTLE*DOGS72	NR	NR	$7.99X10^{-7}$ (b_{10})
	DAIRY*DOGS72	NR	NR	$-3.50X10^{-6}$ (b_{11})
	PIGS*DOGS72	NR	NR	$1.57X10^{-6}$ (b_{12})
	SHEEPDEN (X_8)	103.573	127.859	0.00024 (b_{13})
	GOATDEN (X_9)	98.674	112.896	0.00003 (b_{14})
	CATTLEDEN(X_{10})	5.569	9.339	-0.00028 (b_{15})
	SOIL (X_{11})	0.866	0.256	0.09662 (b_{16})
	RAIN72A (X_{12})	227.566	73.084	0.00019 (b_{17})
	SHEEPDEN*RAIN72A*SOIL	NR	NR	$-4.76X10^{5}$ (b_{18})
	GOATDEN*RAIN72A*SOIL	NR	NR	-9.16X10 (b_{19})
	CATTLEDEN*RAIN72A*SOIL	NR	NR	0.00169 (b_{20})
	ROOMS (X_{13})	0.824	0.142	-0.02655 (b_{21})
	K9PHUM72 (X_{14})	0.030	0.022	-0.90861 (b_{22})
	K9PHUM72*ROOMS			1.78508 (b_{23})
K9PHUM72 (X_{14})				
	DOGS72 (X_1)	17.852	20.866	0.00096 (a_1)
	FLOCKPHUM(X_{15})	0.011	0.014	0.38372 (a_2)
	ROOMS (X_{13})	0.824	0.142	-0.01007 (a_3)
	HUMANPOP (X_{16})	664.577	615.302	-0.00003 (a_4)

[a] 363 Greek Cypriot Villages used in analysis; [b] Figures rounded to nearest thousandth; [c] Because total population of Greek Cypriot villages were used in the analysis, tests of significance were not carried out. NR: Mean and Standard deviation not required for calculations.

Table 4. Decomposition of Effects[a] Determining the Prevalence of *Echinococcus granulosus* in the Dog Populations of Rural Greek Cypriot Villages, June-December, 1972

Predetermined[b] Variable	Total Effect	Direct[c] Effect	Indirect Effects
DOGS72	-0.00036	-0.00090	0.00054 via K9PHUM72
SLTHSE72	0.01577	0.01577	
K9PHUM72	0.56289	0.56289	
SHEEP	0.00001	0.00003	-0.00001 (**)
GOATS	-0.00003	-0.00002	* (**)
CATTLE	0.00053	0.00009	0.00044 (**)
DAIRY	0.00010	0.00011	-0.00001 (**)
PIGS	-0.00006	0.00006	* (**)
SHEEPDEN	-0.00005	-0.00005	
GOATDEN	*	*	
CATTLEDEN	0.00169	0.00169	
RAIN72A	-0.00027	-0.00027	
SOIL	-0.02418	-0.02418	
ROOMS	0.04073	0.02672	0.01402 (-0.011 via GOATS)
FLOCK	-0.00046		-0.00046 (**)
HUMANPOP	*		* (**)
HUMANDEN	0.00003		0.00003 (**)
FLOCKPHUM	0.2096		0.2096 via K9PHUM72
SCRUB	-0.01048		-0.01048 (**)
RAINAVG	-0.00007		-0.00007 (**)
GRAZE	-0.00030		-0.00030 (**)

[a]Effects measured in natural units; [b]See Table 1 for a description of variables; [c]See Table 2 (Structural Eq. (1)) for interacting variables used in calculations; *Effects within the limits ± 0.00001; **Individual effects within the limits ± 0.01