In a previous paper on "Cats and Childhood Leukemia" the authors found that children 1-14 years of age have roughly a double relative risk of leukemia if they were reported to have been exposed to a cat which was sick or died. This finding, based on the childhood leukemias of the Tri-State Survey, has now been confirmed by a similar finding in approximately 1,400 adult cases and adult random-sample controls. A somewhat stronger relationship was found for exposure to ill or dead canaries and parakeets. An equivocal relationship was found for pet dogs. Degree of relationship depends on type of leukemia. These findings do not necessarily conflict with the data for earlier negative reports. It is shown that unless the health status of the pet is included in the analysis, the relationships will be missed.

Pets and Adult Leukemia

Irwin D.J. Bross, Ph.D.; Sister Rosalie Bertell, Ph.D.; and Robert Gibson, Ph.D.

1. Introduction

In a previous article on "Cats and Childhood Leukemia"[1] we have reported an interesting relationship between exposure to ill or dead cats and the occurrence of leukemia in children. This finding was based on interviews with the families of 300 children with leukemia and 831 children from a random sample of the same geographic areas. The data was part of the information collected in a Tri-State Survey which was carried out between 1959 and 1962 in designated areas of New York, Maryland, and Minnesota. The positive findings for the children in this sample survey made it incumbent upon us to extend the scope of our analysis in two directions. First, we wanted to see whether the original relationship could be confirmed in the much larger series of 1,400 adults with leukemia and 1,370 random sample controls. Would there be any relationship between exposure to sick cats and the occurrence of leukemia in adults? Second, the findings suggested that exposures to animals other than cats should be analyzed. Would the relative risk of leukemia in adults be increased by exposure to other sick pets? We shall answer the two basic questions that have just been raised as well as some of the collateral questions that they suggest. For example, the possibility of a third factor which affects both humans and pets will be briefly considered.

The organization of this material reflects a secondary purpose of this paper, which is to clarify the role of statistical analysis in etiological investigations that use survey data. Some misconceptions about this role are currently prevalent—not only among clinicians and laboratory investigators but among epidemiologists and statisticians as well. The analysis of any body of complex data, laboratory or clinical or survey data, is likely to present difficulties in the choice of statistical methods, the evaluation of the results, and the presentation of the findings. Although analysis of retrospective interview data may present more than its share of difficulties, there is ample evidence in recent years that the use of modern statistical techniques by experienced and competent investigators has produced objective, reliable, and medically useful information on etiological factors in cancer and other diseases.

In establishing the chain of events which connect an environmental factor—animal viruses, for instance—to human disease, the most difficult link is ordinarily the one involving humans. Whenever a direct approach through experimentation on human subjects is precluded by ethical and legal considerations, the only feasible alternative for study is likely to be an indirect approach which is based on information that has come through a linguistic channel such as an interview. Only on rare occasions are more direct approaches—such as experimentation on preventive methods—a practical alternative. It is therefore unrealistic and unscientific to criticize retrospective interview data as if the choice were between this data and some "ideal scientific data". The choice is between developing and using techniques which can cope with the problems of interview data or relying instead on speculations and "expert opinions" which have no factual basis. The choice is between epidemiology as an empirical science or as an art form.

2. Demographic Characteristics

The procedures used in the sampling and interviewing in the Tri-State Survey have been described elsewhere[2] and will only be sketched here. The basic approach in the Survey was to obtain the cases of adult leukemia from tumor registry listings and to obtain a comparison series by a random sample from the same geographic areas. The stratification procedures and sampling rates in the random sample were chosen to provide approximately equal numbers of cases and controls. The population over 65 was sampled at a higher rate than the population in the 15-64 age range but there was no attempt to match the age distribution of leukemia cases exactly. The distribution of cases and controls in the three states of the survey are shown in Tables 2.01 and 2.02.

As can be seen from Table 2.01, the bulk of the cases and controls came from New York with the remainder divided between Maryland and Minnesota. Table 2.02 shows the distribution of cases and controls by age and sex. In this and subsequent tables, 13 cases from New York with an inadequate report of age have been omitted. For the leukemia

Table 2.01—Distribution of Cases and Controls in the Three States

	Cases		Controls	
	Number	%	Number	%
New York	803	56.8	863	63.0
Maryland	267	18.9	237	17.3
Minnesota	343	24.3	270	19.7
Total	1413	100.0	1370	100.0

cases the male/female sex ratio is 1.5, about what is usually reported. In the controls the sex ratio is closer to unity but there is the expected preponderance of women in the controls over 65. To deal with the differences with respect to age and sex between the leukemia cases and the random sample controls, the statistical techniques used in the analysis always make an adjustment for age and usually will adjust for both age and sex simultaneously.

The proportion of non-whites in this sample was relatively low, both in the cases and in the controls. In the controls the percentage was just under 10% and in the cases about 5%. In view of these small proportions, an adjustment for this was neither necessary nor feasible in the subsequent analysis.

Table 2.03 shows the distribution of the cases by type of leukemia and state. In all states the Acute Lymphatic (AL) leukemias were the least frequently reported while the Chronic Lymphatic (CL) were the most frequent type. The Acute Myeloid (AM) leukemias were somewhat more frequently reported than the Chronic Myeloid (CM) in the overall series. However there were differences between New York and Minnesota in this respect.

The male/female sex ratios are different for the various types of leukemia ($\chi^2 = 6.94$) ranging from high values of 1.67 and 1.79 for Acute and Chronic Lymphatic leukemias respectively to a low of 1.17 for the Chronic Myeloid. The Acute Myeloid shows an intermediate sex/ratio, 1.47, which is similar to that in the overall case series, 1.50.

Table 2.04 shows the distribution of cases by type of leukemia and age. For all but the Acute Lymphatic cases, the largest number of cases is found in the series over 65 and the fewest in the 15-44 age group. However this pattern is reversed in the Acute Lymphatic series. In the children, the overwhelming majority of the cases were Acute Lymphatic leukemias and this suggested that the best chance of finding a relation to pets in the adult data might be in this type of leukemia.

The demographic distributions of the Tri-State Survey impose some restrictions on the analysis—particularly on the extent to which the series can be cross-tabulated by the various factors. Even though the number of interviews in this survey, about 1,400 cases and almost as many controls, is larger than in most of the previous surveys of leukemia that have been attempted, the numbers in specific sub-series may become small. For instance, the number of Acute Lymphatic cases, 115, becomes a limiting factor in any cross-tabulations which involve leukemia type. The age distribution of the Acute Lymphatic series poses a further problem since there are only 29 and 35 cases respectively in the two older age groups. Detailed cross-tabulations soon result in cross-categories where the number in the series is close to zero. This, in turn, tends to produce somewhat erratic estimates of relative risks and other statistical quantities. The analytic problems and complexities mount up as the cross-tabulations become finer and finer.

For this reason, the order of presentation here will be the order of increasing analytic complexity. It will begin

Table 2.02—Distribution of Cases and Controls by Age and Sex

	Cases			Controls		
Age	Male	Female	Total	Male	Female	Total
15-44	131	100	231	236	247	483
45-64	291	182	473	300	250	550
65+	419	277	696	136	201	337
Total	841	559	1400	672	698	1370

Table 2.03—Number of Observations by Type of Leukemia and State

Leukemia	New York		Maryland		Minnesota		Total	
Type	Number	%	Number	%	Number	%	Number	%
AL*	72	9.1	23	8.6	20	5.8	115	8.2
CL	241	30.5	78	29.2	152	44.3	471	33.6
AM	218	27.6	63	23.6	52	15.2	333	23.8
CM	114	14.4	53	19.8	89	25.9	256	18.3
Other	145	18.4	50	18.7	30	8.7	225	16.1
Total	790	100.0	267	100.0	343	100.0	1400	100.0

Note: *AL = Acute Lymphatic Leukemia
CL = Chronic Lymphatic Leukemia
AM = Acute Myelocytic Leukemia
CM = Chronic Myelocytic Leukemia
Other = Leukemia Without Complete Description by Type

with the questions which can be answered with relatively simple statistical procedures and proceed to questions involving further cross-tabulations, where there is greater variability and where more sophisticated techniques have been used. While this order of presentation has the advantage of showing how the results emerge with greater clarity as the analysis goes deeper, it has the disadvantage that the most interesting and important findings do not come until the latter part of the presentation.

3. Exposure to Animals

In this section the evidence on animal exposure is considered without regard for health status of the animal or the type of leukemia. This type of analysis has been used in the previous studies[3,6] and has led to negative results. At this level of analysis, the results for the Tri-State Survey are in agreement with these negative studies. The Tri-State Survey gives essentially negative results for animal exposures when the health status of the animal is not taken into consideration.

There were two separate questions in the Tri-State schedule that involved animal exposure—one concerning general exposure and the other specifically concerned with pets. In Table 3.01 the basic results for both non-pet and pet exposures are presented, but separated. Eleven non-pet animals and a residual "other animal" series are shown in Table 3.01 in order of increasing estimates of relative risk. The first column of numbers shows the total cases and the second column shows the number of cases exposed to a given animal, the third column is the weighted percentage of cases reporting exposures to a given animal. The fourth column is the corresponding number of controls, the fifth column the number of exposed controls and the sixth column a weighted percentage for the controls. The slight variation in numbers in the first and fourth columns represent occasional clerical errors, unclear statements, or omissions which made the exposure status of an individual uncertain. The analysis used here is a standard statistical procedure for adjusting for the differences in age and sex between cases and controls as was noted in the preceding section. In effect, all comparisons are made within an age-sex category and then recombined by a weighting procedure. The relative risks are calculated using the Woolf-Haldane[7] procedure. The probabilities for statistical testing (or P Value) are calculated using the slightly more conservative Cochran Test.[8] We

Table 2.04—Number of Observations by Type of Leukemia and Age

Leukemia Type	15-44		45-64		65+		Total	
	Number	%	Number	%	Number	%	Number	%
AL	51	22.1	29	6.1	35	5.0	115	8.2
CL	20	8.6	161	34.0	290	41.7	471	33.6
AM	71	30.7	127	26.8	135	19.4	333	23.8
CM	53	22.9	88	18.6	115	16.5	256	18.3
Other	36	15.6	68	14.4	121	17.4	225	16.1
Total	231	100.0	473	100.0	696	100.0	1400	100.0

Table 3.01—Age-Sex Adjusted Relative Risks for Pet and Non-Pet Exposure

	Case			Control				
	Total number	Number exposed	%*	Total number	Number exposed	%	P-Value	Adj rel risk
Non-pets								
Sheep	1333	151	10.5	1349	148	12.2	0.18	0.85
Duck	1328	277	20.4	1344	278	21.7	0.45	0.93
Geese	1317	195	14.4	1346	175	14.4	0.98	1.00
Pig	1329	479	34.0	1351	420	33.3	0.73	1.04
Mice	1367	128	9.0	1348	115	8.6	0.73	1.06
Cattle	1348	561	38.8	1350	468	36.9	0.32	1.09
Horse	1356	600	41.1	1352	494	38.9	0.25	1.10
Hamster	1326	37	3.1	1348	44	2.7	0.63	1.12
Goat	1332	123	8.7	1348	98	7.6	0.31	1.16
Guinea Pig	1348	31	2.5	1347	29	2.0	0.44	1.17
Oth. Anim.	1340	200	15.2	1310	170	12.6	0.07	1.19
Chicken	1335	704	50.3	1354	574	44.7	<0.01	1.26
Pets								
Dog	1370	723	55.1	1353	757	53.0	0.28	1.09
Cat	1360	364	27.5	1352	366	25.7	0.33	1.09
Bird Pets	1373	423	32.4	1357	393	27.6	<0.01	1.26

*Weighted Percentages: Exposed/(Exposed + Not Exposed to this type of animal)

took the precaution of using two standard statistical procedures for age-sex adjustment although in almost all cases they lead to essentially the same results.[9]

The weighted percentages of persons reporting exposure to the various animals listed in Table 3.01 are quite similar in the cases and in the controls. Almost all of the relative risks are in the range 0.8 to 1.2 within 20% of unity. Considering the inherent sampling errors in the estimation of a relative risk, there is little need to assume anything more than sampling variation to account for most of these results. A possible exception is the relative risk for chickens. This is the largest risk, 1.26, and the only one for which the probability of the Cochran Test is significant at the 1% level. It might be noted that about half of the cases and controls report exposure to chickens—and this is the highest exposure reported among the non-pet animals. For all of the rest of the animals in the list there is little suggestion of any relationship between exposure to the animal and occurrence of leukemia. There is a faint suggestion of a difference in the "other animal" series but this is a heterogeneous group including turtles, fish, and some unusual animals.

The results for pets are shown in a separate list in Table 3.01. Neither cats nor dogs show very much suggestion of a relationship between exposure and adult leukemia. The relative risks for both are within 10% of unity and the probabilities are nowhere close to the 5% level. This lack of relationship cannot readily be attributed to sample size. About half of the cases and controls report exposure to dogs and about a quarter report exposure to cats, so there are large series involved.

Here and in subsequent analyses, canaries and parakeets have been combined into a single category of "bird pets." The results for birds are somewhat more positive. The relative risk is 1.26—the same as was previously noted for chickens—and the Cochran Test is significant at the 1% level. The results in Table 3.01 might suggest further investigation of the risks for birds but they are negative for dogs and cats.

In addition to the generally negative results of Table 3.01, it might be noted that there is little similarity between the ranking of the animals in this table and the corresponding ranking of the animals in the data on exposure of the children in the Tri-State Survey. In the children, exposure to cats was of borderline significance but the relative risk for chickens was 0.96. The relative risk for sheep was high for children, 1.70, and the lowest for adults. This again suggests that the ranking in Table 3.01 is largely a reflection of sampling variation (with the possible exception of chickens and pet birds).

Some additional analyses of the exposure data might seem called for. The non-pet animals in Table 3.01 can be combined in various ways. For instance, ducks and chickens might be combined. Or mice, hamsters, and guinea pigs combined and classed as rodents. These combinations were made and the analyses run. However there was little change in the relative risks. Another possibility was to consider *any* animal exposure or *any* pet exposure or the combination of these two factors. One rationale for making such an analysis is the time factor. The pet exposures are often reported in adulthood only and the animal exposures are often reported in childhood only. For instance in a random sub-sample of 40 histories, the exposure to chickens occurred at least 20 years prior to the interview in 80% of the cases and 65% of the controls whereas pet exposures occurring more than 20 years prior to the interview were not reported in this survey. This time factor, which is of considerable importance, is discussed in detail in Section 7. The results of the analysis of pet and non-pet exposures, by geographical area, are given in Table 3.02.

Table 3.02—Age-Sex Adjusted Relative Risks for Pet and Non-Pet Exposure by Geographical Area

	Case Number			Control Number				
	Not exp	Exp	%*	Not exp	Exp	%	P-Value	Adj rel risk
Non-Pets								
New York	78	115	57.2	96	120	57.4	0.99	1.00
Maryland + Minnesota	50	150	73.4	56	121	71.6	0.70	1.11
3 States Combined	128	265	65.8	152	241	63.0	0.42	1.13
Pets								
New York	78	208	74.4	96	295	72.8	0.64	1.06
Maryland + Minnesota	50	137	75.1	56	146	71.5	0.45	1.19
3 States Combined	128	345	74.8	152	441	72.3	0.38	1.13
Pets and non-pets								
New York	78	320	80.6	96	327	76.0	0.13	1.29
Maryland + Minnesota	50	208	79.8	56	173	75.7	0.31	1.26
3 States Combined	128	528	80.4	152	500	75.7	0.06	1.31

*Weighted Percentages: Exposed/(Exposed + No Animal Exposure)

As can be seen from Table 3.02 there is some tendency for the relative risks to be higher in the persons exposed to both pets and non-pet animals but this difference does not quite make statistical significance at the 5% level. There is also some indication here that the results in New York are not very different from the results in the other two states. All in all, however, it appears that the combination of animals does not clarify the picture to any great extent.

4. Exposure to Sick* Pets

The results of the analysis of exposure to animals turned out to be essentially negative when the health status of the animals was not taken into account. This agrees with what was previously reported by others and with what was reported for the Tri-State data on childhood leukemia. In the adults there is a borderline relative risk for the chickens and the pet birds and in the children there was a similar situation for the cats. However no clear-cut findings were obtained and, indeed no such findings would be expected even if illness in humans is related to illness in animals. The key point here is that when the sick animals and the well ones are combined, the latter would tend to dilute out any effect from the former and the relationship, if any, could at best be a weak echo of the underlying relationship. There is, in effect, a misclassification of sorts. That is, we would be interested in exposures where, say, there was a risk of transmission of a virus whereas we are actually dealing with all exposures. From this standpoint including the well animals in the "exposed" series is a misclassification of the exposure status and would obscure the actual relationship.[10,11]

On the other hand, bringing the survey information on the health status of pets into the analysis also presents some evident difficulties. The information was obtained as a report of a respondent and was not checked against veterinarian records. At the time of the survey, 1959 to 1962, there were insufficient funds for this purpose and a check of these data now is practically impossible. Another problem with these data is recall, particularly in respondents over 65. This problem is more acute in the cases than in the controls. A further problem is that many cases were deceased at the time of interview and it was necessary to rely on a spouse or other near relative for information. Intercomparison in the cases and in the controls suggest more underreporting of exposure among the cases than among the controls. Moreover the relative risks do not appear to be greatly affected by the factors. The question concerning ill pets was limited to a one-year period, hence minimizing the recall problem. Also the difference between primary and secondary respondents proved to be relatively small. Of the 308 live cases 9.4% reported ill pet exposure, while 9.0% of the 1,105 deceased cases reported similar exposure. This shows a very slight trend toward underreporting among non-primary respondents, which would in turn slightly lower the relative risk estimates.

The relatively small number of controls reporting exposure to ill pets—about 5% of the total number—places sharp restrictions on the extent of cross-tabulation possible. For this reason, some of the tables are combined for sex and weighted only over age. There were no significant differences in the relative risks for the male and female which would have made this combination inadvisable.

It is important for comparison with the data on children previously reported,[1] to keep in mind that a different set of health status categories was used with the children. Recall problems were not a major difficulty in the childhood study due to the relatively short period of time involved. For the children any report of illness *or* death of a cat during the time interval between birth and time of diagnosis (or interview) was classified as a positive exposure. With adults, only exposure to ill, or ill and dead pets reported in the year prior to death or interview are counted as positive exposure. The implications of this classification for various viral hypotheses are again considered in Section 7. This classification is a neutral one in the sense that it deliberately avoids making a choice between the man to pet, pet to man, or common environmental factor hypotheses currently proposed in the literature.[4,12,13,14]

Tables 4.01a, b and c give the basic data on exposure.

It might be well to note also that illness in pets may be undetected by the average person unless the symptoms are quite severe and productive of very clear deviations from the normal behavior pattern. This would again result in underreporting of exposure for both cases and controls, which in turn tends to reduce relative risk estimates. The negative results in the previous section show this dilution effect. The positive results reported in this section would only be strengthened if there was some way to eliminate misclassifications.

In the study of childhood leukemia, the relationship to cats did not come into clear focus until the health status of the cats was taken into account. When this was done the effects were striking. The relative risks jumped from 1.35 for exposure to any cats to 2.24 for exposure to ill or dead cats. The probability levels went from borderline significance at the 5% level to significance well beyond the 1% level. In chronic disease epidemiology, it is relatively rare to have negative and borderline results come into sharp focus in this way when a subclassification of a variable is made. When this happens, it is a good indication that the analysis has hit pay dirt.

As can be seen from Table 4.02, taking into account the health status of the pets had much the same effect in the analysis of the adult data that it had previously produced in the analysis of the childhood leukemia.

The relative risk for adults with an exposure to pet birds jumped from 1.26 for a report of any exposure to 1.99 when there is a report of a sick pet bird. In other words making the sub-classification by health status allows us to go from about a 25% increase in risk to close to a 100% increase in risk. The significance level for the test of this relationship is pushed well beyond the 1% level. Even more interesting is what happens with the cats. From an unimpressive relative risk of 1.09 for all cats, the relative risk jumps to 1.75 for sick cats. At the same time the probability goes from nowhere near statistical significance, to the 5% level. For dogs the picture is less clear but even here there is much more indication of a relationship than was previously found.

One further question may be asked at this point. Namely, does severe illness and recent death in a family make one more aware of illness in pets? A second sample of

* In this report sick pet includes all pets reported as both sick and dead. Dead pet refers only to those animals reported dead but not sick.

Table 4.01a—Number of Observations for the Exposure Status of Cat Pet by Sex and Age

	Age		No cat	Well cat	Dead cat*	Sick Cat*	Total
	15-44	Case	82	32	11	2	127
		Control	158	52	21	1	232
Male	45-64	Case	209	40	28	8	285
		Control	214	45	34	4	297
	65+	Case	306	57	44	5	412
		Control	107	11	13	3	134
	15-44	Case	68	24	4	3	99
		Control	159	58	26	1	244
Female	45-64	Case	137	24	15	2	178
		Control	187	29	29	4	249
	65+	Case	194	40	26	7	267
		Control	161	19	19	0	199

*NOTE: Cats reported both sick and dead are listed as sick. Those reported as dead without report of sickness are listed as dead.

Table 4.01b—Number of Observations for the Exposure Status of Dog Pet by Sex and Age

	Age		No dog	Well dog	Dead dog*	Sick dog*	Total
	15-44	Case	35	58	29	7	129
		Control	76	89	58	9	232
Male	45-64	Case	114	77	77	17	285
		Control	124	73	83	16	296
	65+	Case	227	102	76	9	414
		Control	80	20	30	2	132
	15-44	Case	41	35	17	5	98
		Control	75	102	60	7	244
Female	45-64	Case	84	46	38	9	177
		Control	119	61	63	7	250
	65+	Case	146	51	60	10	267
		Control	122	29	44	4	199

*NOTE: Dogs reported as sick and dead are listed as sick. Those reported dead without report of sickness are listed as dead.

Table 4.01c—Number of Observations for the Exposure Status of Bird Pets by Sex and Age

	Age		No bird	Well bird	Dead bird*	Sick bird*	Total
	15-44	Case	80	18	26	4	128
		Control	175	28	24	6	233
Male	45-64	Case	186	35	49	14	284
		Control	200	55	38	3	296
	65+	Case	308	40	55	9	412
		Control	108	15	10	1	134
	15-44	Case	55	24	12	5	96
		Control	148	47	41	9	245
Female	45-64	Case	129	22	21	5	177
		Control	180	30	33	7	250
	65+	Case	192	30	37	12	271
		Control	153	16	28	2	199

*NOTE: Birds reported sick and dead are listed as sick. Those reported as dead with no report of sickness are listed as dead.

Table 4.02—Age-Sex Adjusted Relative Risks for Sick Pet Exposures

Pets	Case			Control			Adj rel	
	No pet	Sick Pet	%*	No pet	Sick pet	%	P-Value	risk
Dog	647	57	9.3	596	45	6.4	0.07	1.47
Cat	996	27	2.8	986	13	1.4	0.04	1.75
Bird Pets	950	49	5.3	964	28	2.5	<0.01	1.99

*Weighted Percentages: Sick Pet/(Sick Pet + No Pet)

controls, referred to as "ill and dead controls," was matched as closely as possible in age, sex and date of release from (or expiration in) the same hospital, with a random sample of 211 leukemia patients. All the patients in this control group were non-leukemic. Of these controls, 6 persons (3%) reported exposure to ill dogs, none to ill cats or birds. This percentage agrees exactly with the 3% exposure to ill dogs reported by the larger control series. There is no reason to believe that reporting of ill pets is a phenomenon common to many families of individuals affected by a severe disease.

5. Acute vs. Chronic Leukemia

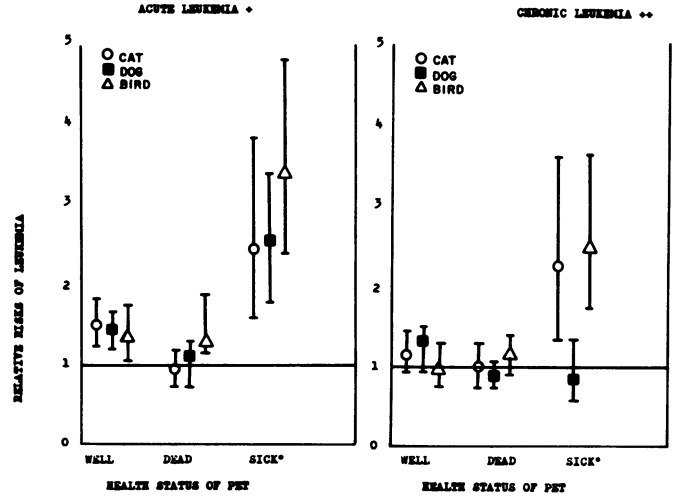
As soon as the type of leukemia is brought in as a factor in addition to the health status of the pet, the problems of the cross-tabulation become increasingly troublesome. The basic analytic tool, the Cochran Test, is fairly stable even when the number of observations in a cell of the cross-tabulation gets down toward zero. However there is some possibility that significance tests and the relative risks are running into technical difficulties and it is desirable to apply alternative methods of analysis to the cross-tabulated data as a check on the results. The alternative procedure used in this section is that of a mathematical model of the data similar to the model used in the analysis of the childhood leukemia material. The model can be used to calculate relative risks but these are not necessarily the same as risks computed by the methods of the previous sections. The model involves a set of parameters for the various health status categories of the pets and these are fitted by a direct minimization of Chi-square in the age-by-health-status contingency table. The technique is too complex to describe in full but a detailed report on the technique will be sent to readers on request.

The alternative procedure for analysis has its assumptions and limitations but they tend to be different from those of the Cochran Test. However empty cells in the contingency tables are troublesome to some extent with any procedure. To avoid such cells, the new analysis adjusts for age but not for sex. Even with this restriction there were still numerous empty cells when the cross-tabulation involved all five types of leukemia. Therefore the data was recombined into just two classes. The first, called "acute" in this section, included the acute lymphatic, acute myelocytic, and the residual "other" type. The second, called "chronic" in this section, included the chronic lymphatic and chronic myelocytic leukemias. The analysis was then carried out for cats, dogs, and pet birds and for the "acute" and "chronic" groupings that have just been described. In each case there were the three age categories used previously and the four categories of health status of pet (no pet, well pet, dead pet, sick pet) so that there were 12 cells in each contingency table. A separate mathematical model was fitted for each of the six contingency tables. The arithmetic labor in this process is very heavy and would not have been attempted if a remote terminal computing system had not been available for this purpose. It is to be noted that although these relative risks were calculated by means of the mathematical model, they are strikingly similar to those calculated by the standard methods.

The results of the analysis are shown in graphical form. Figure 5.01 shows estimates of relative risks and approximate confidence intervals for each estimate.

Perhaps the most striking feature at first glance is the similarity of all three pets insofar as their general pattern is concerned. In all of the pets the estimates of the relative risks in the "well pet" and "dead pet" categories are fairly close to unity. For all except the combined "chronic" series for the dogs, the "sick pet" category has an elevated relative risk which is significantly different from unity. These intervals are not completely independent of each other—the combined "acute" and combined "chronic" risks for each pet involve the same control series but the repetition of the pattern is strong evidence that there is a definite relationship between sick pets and sick humans. The existence of such a relationship does not, of itself, give any clear indication of the underlying mechanism. However it does suggest that this data can give some clues as to mechanism

Figure 5.01—Confidence Intervals (95%) of the Age-Sex Adjusted Relative Risks in Adult Leukemia by Pet, Health Status of Pet and Leukemia Types



*Sickness Occured Within 1 Year Prior to Interview or Death of Subject

+Acute Leukemia- Acute Lymphatic, Acute Myelocytic and Leukemia without complete description by type

++Chronic Leukemia- Chronic Lymphatic and Chronic Myelecytic Leukemia

in the sense that a given hypothesis may be roughly evaluated by how well it explains the pattern.

It can be seen from Figure 5.01 that the effect for sick pets seems to show more strongly among the patients in the acute series than among those in the chronic series. The confidence intervals are, however, too wide to permit this tendency to be statistically demonstrable. We will return to this point in a later discussion of timing considerations for specific viral hypotheses. The pattern of risks is largely (but not entirely) similar in the acute and chronic series. Exposure to sick cats and sick birds have relative risks greater than two in both series. For sick dogs, however, there is a marked difference with the elevated risk in the acute series disappearing completely in the chronic series.

As a further check on the analytic techniques, the same data was analyzed by an unweighted analysis of variance and estimates of the relative risks were obtained from the means. Almost all of the estimates were similar to those shown in the graphs and even when the estimates were different, the pattern was unaffected. For example, the relative risk for the bird pets in the acute series was 2.70 by the analysis of variance approach instead of 3.39 by direct minimization of Chi-Square[15] but in either case it was the largest of the relative risk estimates. The main point of the parallel analyses is to check whether the results are very dependent on the choice of a particular statistical technique. Here the alternative analyses—while very different from a technical standpoint—lead to essentially the same results.

6. Type of Leukemia

In this section each type of leukemia will be considered separately. The analytical procedures will again employ the Woolf-Haldane estimates of relative risk with the Cochran probability values. In order to avoid the problem of zero cells when this degree of cross-tabulation is introduced, pet types have been combined for the preliminary study of exposure to sick pets. Because of the large number of selectors being used, these first two tables were limited to the sub-population of white cases and white controls (about 90% of the total sample).

Table 6.01a shows the age-adjusted relative risks for exposure to any sick pet. All leukemia types show significantly high relative risks except the chronic myeloid group, where the effect is conspicuously absent.

In Table 6.01b the rather consistent increase in relative risk with age becomes apparent. Note also the stronger effect in the acute lymphatic patients, as might be expected from the result for children.

When we consider specifically the exposure to sick cat pets, the degree of cross-tabulation necessary results in some cross-categories in which there are no cases of a given type of leukemia. The Cochran Test is relatively stable even with empty cells, but the Woolf-Haldane procedure for calculating the relative risks runs into more serious difficulties. The procedure involves taking logarithms of the counts in the cells and the arithmetical process cannot handle the logarithm of zero. For this reason it is standard practice to add a small constant which prevents the count from going to zero, but this device puts an upward bias into the resulting estimate of the relative risk.

Table 6.02 shows the age and sex adjusted relative risks and significance tests for the exposure to sick cats. The adjusted ratios here, as in the previous tables of this type,

Table 6.01a—Age Adjusted Relative Risks and Significance Tests for Exposure to Sick Pets by Leukemia Types (White Cases and Controls Only)

Leukemia type	Case Number exposed to			Control Number exposed to			P-Value	Adj rel risk
	No pet	Sick pet	%*	No pet	Sick Pet	%		
AL	24	16	40.03	344	77	17.88	0.01	3.50
CL	161	40	20.97	344	77	13.47	0.03	1.72
AM	83	32	28.23	344	77	16.56	0.01	2.08
CM	77	13	14.66	344	77	14.40	0.95	1.11
Other	56	25	31.81	344	77	14.92	0.01	2.87

*Weighted Percentages: Sick Pet/(Sick Pet + No Pet)

Table 6.01b—Number of Sick Pets and Relative Risks for any Sick Pet Exposure by Age and Type of Leukemia (White Cases and Controls Only)

Leukemia type	15-44		45-64		65+	
	Any sick pets	Rel. risk	Any sick pets	Rel. risk	Any sick pets	Rel. risk
AL	8	2.98	4	3.64	4	4.13
CL	2	1.27	18	1.42	20	2.46
AM	7	1.36	15	1.97	10	3.22
CM	1	0.45	6	0.97	6	1.86
Other	5	1.93	11	3.09	9	3.35
Control	30	1.00	37	1.00	10	1.00

are ratios of the general form: sick pet/(sick pet + no pet). For the cats all of the relative risks are 2.0 or more, but only the chronic lymphatic series makes statistical significance at the 5% level.

The largest relative risk occurs in the acute lymphatic series but it is subject to the bias noted above. The lack of statistical significance in this series is due in part to the smaller numbers in the cells, as can be seen in Table 6.03a. The youngest group shows the highest relative risk to this pet type.

These tables, 6.03b and 6.03c, again have limitations due to very small numbers in the cells, but the high relative risks in the two acute series persist with each type of pet.

The relationship between exposure to sick dogs and the type of leukemia is shown in Table 6.04. Here there is little evidence of any relationship for the chronic leukemias. On the other hand the relationship with the acute leukemias appears to be about as strong as the corresponding relationship to cats. With the dogs the number of sick pets is somewhat larger so that the acute lymphatic, acute myelocytic, and "other" series all attain statistical significance at the 5% level and the "other" series is significant at the 1% level.

The lack of a relationship for exposure to sick dogs in the chronic leukemia series is an exception to the general pattern of relationship between sick pets and leukemias. It represents a piece of evidence which would seem to contradict some hypotheses about underlying mechanisms. For example, one possible hypothesis to explain these relationships would be in terms of transmission of a virus from the human with leukemia to the pet. The definition of sick pet in the adult study involves a time relationship compatible with this hypothesis. However it would be plausible to expect that the opportunities for this kind of reverse transmission would be more favorable in patients with chronic leukemias than in those with acute leukemias. This is contrary to the general pattern of the relative risks and, in particular, conflicts with the apparent absence of a relationship between exposures to sick dogs and the chronic leukemias. The statistical results do not, of course, rule out the possibility that some cases of illnesses in the pets resulted from

Table 6.02—Age-Sex Adjusted Relative Risks and Significance Tests for Exposure to Sick Cats By Leukemia Type

Leukemia type	Case Number exposed to No cat	Sick cat	%*	Control Number exposed to No cat	Sick cat	%	P-Value	Adj rel risk
AL	79	3	3.6	986	13	1.3	0.09	4.94
CL	345	13	3.7	986	13	1.6	0.03	2.62
AM	230	4	1.7	986	13	1.4	0.74	1.97
CM	187	3	1.6	986	13	1.6	0.99	2.06
Other	155	4	2.6	986	13	1.5	0.28	3.12

*Weighted Percentages: Sick Cat/(Sick Cat + No Cat)

Table 6.03a—Number of Sick Cats and Relative Risks for Sick Cat Exposure by Age and Type of Leukemia

Age	AL Sick cats	AL Rel.* risk	CL Sick cats	CL Rel. risk	AM Sick cats	AM Rel. risk	CM Sick cats	CM Rel. risk	Other Sick cats	Other Rel. risk	Control Sick cats
15-44	2	10.57	0	0.00	1	3.30	1	5.29	1	6.09	2
45-64	0	0.00	5	2.13	2	1.10	1	0.72	2	2.39	8
65+	1	3.91	8	3.41	1	0.98	1	1.02	1	1.03	3

* Relative Risk = (Sick Cats in Case/No Cat in Case) / (Sick Cats in Control/No Cat in Control)

Table 6.03b—Number of Sick Dogs and Relative Risk for Sick Dog Exposure by Age and Type of Leukemia

Age	AL Sick dogs	AL Rel.* risk	CL Sick dogs	CL Rel. risk	AM Sick dogs	AM Rel. risk	CM Sick dogs	CM Rel. risk	Other Sick dogs	Other Rel. risk	Control Sick dogs
15-44	4	1.99	1	1.35	4	2.10	0	0.00	3	2.36	16
45-64	3	3.17	8	1.14	8	1.76	2	0.53	5	2.03	23
65+	1	1.77	4	0.80	5	2.47	1	0.53	8	4.81	6

* Relative Risk = (Sick Dogs in Case/No Dog in Case) / (Sick Dogs in Control/No Dog in Control)

viral transmission from leukemic humans. But it does make this an unlikely explanation of the overall relationships that have been found in this study.

Table 6.05 shows the relationships between sick birds and the type of leukemia. The relative risks for the bird pets are numerically similar to those for cats with the exception of the acute myelocytic series. However the level of statistical significance that is reached in the bird pets is higher than that for the cats and only the chronic myelocytic series fails to show significance at the 5% level. The highest relative risk is once again found in the acute lymphatic series.

The estimates and significance tests in Tables 6.01—6.05 are not entirely independent since there is some reuse of the same information but nevertheless the degree of dependence is not high enough to account for much of the similarities over different pets and different types of leukemia. Indeed the exceptional results for sick dogs and chronic leukemias are evidence that the patterns do not have to be the same. Consequently the overall pattern in this data provides mutual reinforcement and confirmation of the individual findings for pets and types of leukemias. The suggestion from the childhood leukemia study that the acute lymphatic leukemias might show the strongest relationship to pet illnesses is borne out by the adult data. While individual estimates or tests might be open to technical objections, it is not possible to explain away any appreciable part of the findings on such grounds.[16]

7. Discussion

The data on the adults in the Tri-State Survey contains so much information—and the information is so complex—that no one analysis or set of analyses can be

Table 6.03c—Number of Sick Bird Pets and Relative Risks for Sick Bird Pets Exposure by Age and Type of Leukemia

	AL		CL		AM		CM		Other		Control
Age	Sick birds	Rel.* risk	Sick birds	Rel. risk	Sick birds	Rel. risk	Sick birds	Rel. risk	Sick birds	Rel. risk	Sick birds
15-44	2	1.54	1	1.66	4	2.10	0	0.00	2	2.27	15
45-64	1	2.24	6	2.05	6	2.88	2	1.17	4	3.54	10
65+	2	7.97	8	3.32	4	3.61	5	5.34	2	1.99	3

* Relative Risk = (Sick Bird Pets in Case/No Bird Pets in Case) / (Sick Bird Pets in Control/No Bird Pets in Control)

Table 6.04—Age-Sex Adjusted Relative Risks and Significance Tests for Exposure to Sick Dogs By Leukemia Type

	Case			Control				
Leukemia	Number Exposed to			Number Exposed to			P-Value	Adj rel
type	No dog	Sick dog	%*	No dog	Sick dog	%		risk
AL	48	8	14.5	596	45	6.7	0.03	3.06
CL	248	13	5.6	596	45	5.6	0.99	1.18
AM	134	17	11.7	596	45	6.2	0.03	2.21
CM	123	3	2.5	596	45	6.1	0.12	0.74
Other	94	16	14.9	596	45	5.8	0.01	3.13

* Weighted Percentages: Sick Dog/(Sick Dog + No Dog)

Table 6.05—Age-Sex Adjusted Relative Risks and Significance Tests for Exposure to Sick Birds by Leukemia Type

	Case			Control				
Leukemia	Number exposed to			Number exposed to			P-Value	Adj rel
type	No bird	Sick bird	%*	No bird	Sick bird	%		risk
AL	67	5	7.0	964	28	2.7	0.04	4.13
CL	335	15	4.6	964	28	1.8	0.01	2.72
AM	217	14	6.2	964	28	2.3	0.01	3.03
CM	181	7	3.6	964	28	2.3	0.31	2.01
Other	150	8	5.6	964	28	2.0	0.01	3.35

* Weighted Percentages: Sick Bird/(Sick Bird + No Bird)

exhaustive or tell the whole story. The aim of the present analysis is to give a general picture of the relationships between exposure to animals and the occurrence of leukemia in adults. With a little ingenuity, the data can also be used to try to confirm or refute various specific etiological hypotheses which might be proposed. In particular it is pertinent to the hypotheses which involve viruses in one role or another—hypotheses which have become increasingly popular in recent years.[6,17,18,19] Some of the difficulties in putting viral hypotheses about human leukemia to an empirical test have been noted in the introduction. The Tri-State Survey data is far from ideal for the purpose of testing hypotheses about underlying mechanisms but it is likely to be about the best data available in the immediate future.

In what follows, no specific etiological hypotheses are proposed and the ones which come up in the discussion are only introduced to illustrate some salient feature of the data. A stringent empirical test of a specific hypothesis requires precise—preferably mathematical—formulation and a corresponding statistical analysis for the purpose. Here we will be concerned only with the broad limitations and restrictions on etiological hypotheses which might be proposed as explanations of the present findings. As for etiological speculations, this is a game which the reader can play at least as well as the authors and it is not our intention to enter this game.

Perhaps the most important point which should be noted—both with respect to etiological hypotheses and for the public health implications—is that most of the cases of human leukemia studied here *cannot* be explained by the relationships between pets and leukemia. That is, while the relationships seem to be real they can only account for a small fraction of the total cases. Using calculations based on the mathematical model mentioned in Section 5, the excess number of leukemia cases for all of the sick pets taken together would be about 55. This is only about 4% of the total cases of adult leukemia.

This estimate of the number of leukemic cases explained by the observed relationship is based on several assumptions and therefore might be rather far from the mark. It might be something of an overestimate if there are extraneous factors that happen to be correlated with sick pets and these factors are contributing to the observed relative risks. The relative risks in Section 5 are large enough to make it unlikely that extraneous factors are playing an important role. But while application of the Size Rule[16,20] makes it unlikely that there is much effect from such extraneous factors, they cannot be entirely ruled out.

On the other hand, the dangers of underestimation of the effect are, perhaps, somewhat greater. For example, pet illnesses are likely to be underreported rather than overreported. Moreover the relationship which here only covers the last year may extend back two or more years. Finally, the non-pet animal exposures have not been brought into the calculations. Hence it is possible that the proportion of cases explainable by animal exposures is substantially larger than the above estimate. Even so, however, it would remain true that the etiology of most of the adult leukemias *cannot* be explained by the relationship to pets.

The possibility of pets being hosts to a virus that can produce human leukemia has been the subject of many recent studies.[1,3,4,6,14,17,18] This hypothesis requires a latent period which according to the report of Norris, Jackson and Aaron[17] may be a three-year interval. The data available from the Tri-State Survey, since it confines reports of ill pets to the period of one year prior to death or interview, is unsuited to furnishing any support for this hypothesis.

Hypotheses based on ill pet exposure as a trigger-experience which either overloads the patient's immunomechanisms in such a way that he can no longer fight off the latent disease, or which introduces something into the patient's system which activates the virus to the point of clinical manifestation, can be more readily tested by the Tri-State data.

A study of the time relations between the illness of the pet, the diagnosis of leukemia, and the date of interview or death revealed that only 45 cases reported exposure during a time interval not consonant with these trigger hypotheses. These exposures were 3 months or more after diagnosis of leukemia. More than half of these reported exposures not consonant with a trigger hypothesis occur in the chronic cases. This was to be expected since 59% of the acute leukemia cases reporting ill pet exposure had a time lapse of less than 3 months between the date of diagnosis and date of interview or death, while among the chronics only 13% reported this short a time interval.

The two acute series involve 2 reports of exposure to ill cat pets, 7 reports of exposure to ill dog pets and 3 reports of exposure to ill bird pets, which could be considered misclassifications under the trigger hypothesis. While their deletion from the study would bias the relative risk downward (since no corresponding reduction of the time interval of reported exposure can be made for the controls), it would not appreciably change the results. The positive findings for the acute leukemias are therefore concordant with a trigger hypothesis.

The number of hypotheses and counterhypotheses which might be formulated has no upper limit. To mention just one more class, there would be counterhypotheses of reverse transmission of illness from human to pet. Some objections to the human-to-pet direction of transmission have been previously noted in Section 6. However it is always possible to add codicils to casually stated etiological hypotheses which seem to explain away any given item of contrary evidence. It might be noted however that some forms of man-to-pet transmission would be evidence for a viral etiology for leukemia. Perhaps experimental canaries should be added to the environment of leukemic patients in a simple controlled study.

The main point of this discussion is that the results have shown the potential utility of the Tri-State Survey data in providing an empirical test for the various speculative hypotheses. It also shows the impossibility, in any single article, of dealing directly with all of the hypotheses about the etiology of leukemia that are currently popular. We would be glad to collaborate with any serious investigators who would like to utilize our informational and methodological resources to try to get at least tentative answers to this etiological question. If the information is in the schedules and has been coded, it would be possible to get such answers in a relatively short time.

ACKNOWLEDGMENTS

We are indebted to the collaborators who participated in the design and conduct of the Tri-State Leukemia Survey and, in particular, Drs. Saxon

Graham, Abraham Lilienfeld, Leonard Schuman, Morton Levin and Mr. John Dowd. The analysis of the data has been carried out with the assistance of Mrs. Alice Chamberlain, Mr. Jack Tidings, Mr. Natarajan, Miss Jeanne Coombs, and others in the RPMI Departments of Biostatistics and Epidemiology.

References

1. Bross, I.D.J. and Gibson, R. Cats and childhood leukemia. J. Medicine 1:180-187, 1970.
2. Graham, S.; Levin, M.L.; and Lilienfeld, A. et al. Methodological problems and design of tristate leukemia survey. Ann. New York Acad. Sc. 107:557-569, 1963.
3. Hanes, B.; Gardner, M.B.; and Loosli, C.G. et al. Pet association with selected human cancers: A household questionnaire survey. J. Nat. Canc. Inst. 45:1155-1162, 1970.
4. Schneider, R.; Dorn, C.R.; and Klauber, M.R. Cancer in households: A human-canine retrospective study. J. Nat. Canc. Inst. 41:1285-1292, 1968.
5. Schneider, R. Unpublished results.
6. Gardner, M.B. Current information on feline and canine cancers and relationship or lack of relationship to human cancer. J. Nat. Canc. Inst. 46, 1971.
7. Haldane, J.B.S. Estimation and significance of location of ratio of frequencies. Ann. Human Genetics 20:309-311, 1955.
8. Cochran, W.G. Some methods for strengthening common χ^2 tests. Biometrics 10:417-451, 1954.
9. Sheehe, P.R. Combination of log relative risk in retrospective studies of disease. AJPH 56:1745-1750, 1966.
10. Bross, I.D.J. Misclassification in 2x2 tables. Biometrics 10:478-486, 1954.
11. Tenenbein, A. Double sampling for estimating from bionomial data with misclassifications. J. Amer. Stat. Assoc. 65:1350-1361, 1970.
12. Brill, B.A.; Tomonaga, M.; and Heyssel, R.M. Leukemia in man following exposure to ionizing radiation. Ann. Intern. Med. 56:590-609, 1962.
13. Miller, R.W. Environmental agents in cancer. Yale J. Biol. Med. 37:487-507, 1965.
14. Van Hoosier, G.L.; Stenback, W.A.; and Mumford, D.M. et al. Epidemiological findings and electron microscopic observations in human leukemia and canine contacts. Int. J. Cancer 3:7-16, 1968.
15. Bross, I.D.J. and Blumenson, L.E. Statistical testing of a deep mathematical model for human breast cancer. J. Chron. Dis. 21:493-506, 1968.
16. Bross, I.D.J. Spurious effects from an extraneous variable. J. Chron. Dis. 19:637-647, 1966.
17. Norris, F.D.; Jackson, E.W.; and Aaron, E. Prospective study of dog bite and childhood cancer. Cancer Research 31:383-386, 1971.
18. Essex, M. Can feline oncornaviruses cause leukemia in man? Unpublished paper.
19. Essex, M.; Klein, G.; Snyder, S.P.; et al. Feline sarcoma virus induces tumours: Correlation between humoral antibody and tumour regression. Unpublished paper.
20. Bross, I.D.J. Pertinency of an extraneous variable. J. Chron. Dis. 20:487-495, 1967.

Dr. Bross, Sister Bertell, and Dr. Gibson are affiliated with the Departments of Epidemiology and Biostatistics, Roswell Park Memorial Institute, Buffalo, New York. This investigation was supported by Public Health Service Research Grant No. CA-11531 from the National Cancer Institute. Reprint requests should be addressed to: Dr. Irwin D.J. Bross, Department of Biostatistics, Rosewell Park Memorial Institute, 666 Elm St., Buffalo, New York 14203. This paper was submitted for publication in July, 1971.