AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol. 112, No. 6
Printed in U.S.A.

EPIDEMIOLOGIC INVESTIGATION OF OCCUPATIONAL CARCINOGENESIS USING A SERIALLY ADDITIVE EXPECTED DOSE MODEL

A. H. SMITH,[1,2] R. J. WAXWEILER[3] AND H. A. TYROLER[1]

Smith, A. H. (Dept. of Community Health, Wellington Clinical School of Medicine, Wellington Hospital, Wellington 2, New Zealand), R. J. Waxweiler and H. A. Tyroler. Epidemiologic investigation of occupational carcinogenesis using a serially additive expected dose model. *Am J Epidemiol* 112:787–797, 1980.

The epidemiologic identification of occupational carcinogens is complicated by several problems including worker mobility between jobs, variation over time of chemicals and processes used, and the long latency period between exposure and discovery of a tumor. In the light of these problems, a method using the cumulative dose concept has been developed which involves calculating the expected yearly exposure for each case from work histories of all noncases close to the case in year of birth and year of hire. The data required for use of the method include information concerning exposure to the chemicals being studied for each job in each calendar year of the study. Use of the method is illustrated with a study of angiosarcoma of the liver and vinyl chloride exposure in a polymerization plant. The value of the method lies in the wealth of information generated concerning the association between chemical exposures and cancer, including exposure level relationships, latency information, and the possibility that two chemicals might be acting independently or jointly. The serially additive expected dose model is likely to prove particularly useful in the analysis of data collected by occupational health surveillance systems, as well as retrospective studies of the type illustrated.

angiosarcoma; cancer; epidemiologic methods; occupational diseases; vinyl chloride

The epidemiologic identification of occupational carcinogens is generally complicated by several features of the exposure of workers to industrial chemicals, and also by certain aspects of the carcinogenic process itself. The major problems have been presented and discussed elsewhere (1–4) and will only be briefly outlined here. Our main objective is to present a method of data analysis which deals with many of these problems in the identification of likely carcinogenic agents in the work environment. The model used requires exposure information which, although not usually obtained in occupational studies, is generally available with the cooperation of the companies involved.

A major problem in occupational cancer studies is that worker movement from job

Received for publication December 13, 1979, and in final form March 25, 1980.

Abbreviation: SAED; serially additive expected dose.

[1] Occupational Health Studies Group, University of North Carolina at Chapel Hill, NC.

[2] Reprint requests to Dr. A. H. Smith. Current address: Department of Community Health, Wellington Clinical School of Medicine, Wellington Hospital, Wellington 2, New Zealand.

[3] National Institute for Occupational Safety and Health, Cincinnati, OH.

The helpful suggestions by Dr. David Deubner are gratefully acknowledged.

to job within industrial plants is accompanied by variation in exposure to any suspect carcinogen. In some instances movement between different companies is sufficiently frequent to add to this problem. In addition, the chemicals used by workers in individual jobs change over time with the introduction of new production methods and with changes in the products themselves. In many industries, workers are exposed to a large number of different chemicals, further complicating the isolation of carcinogenic agents. Even when the same chemicals are used in the same amounts, changes in work practices and plant ventilation alter exposure levels.

With regard to the carcinogenic process itself, the major complicating factor is the long latency period between exposure to a carcinogen and the diagnosis of cancer. Unless ongoing surveillance systems are operational, this means that exposure information must be sought for 20–30 years prior to conducting a study. The risk of cancer may depend on the level, duration, and continuity of exposure, and accurate information of this kind may be difficult to obtain retrospectively over such a long period of time. An additional complicating feature of the carcinogenic process is that more than one chemical may be involved, either independently, or jointly in an interactive manner.

The method of data analysis which we present here was constructed in an attempt to deal with as many of these problems as possible. Much of the data required is obtained routinely in occupational cancer studies other than those which merely compare disease incidence or mortality with that in the general population. The additional information required is the average exposure level for each job in each calendar year under study, although the method can still be used when classification is limited to exposed and nonexposed jobs. Exposure information allowing a simple ranking of job exposure can be utilized and was made available for this study by the company from interviews of personnel with a knowledge of past and current jobs within the study plant and from records of chemical usage. The method can also incorporate actual chemical exposure measures when they are available.

Method

The basis of the serially additive expected dose (SAED) method in descriptive terms is to compare the observed exposure of each case in a study with the exposures of fellow workers close to the case in year of birth, and in age at commencement of work for the company in the study. If the total work force in a plant is referred to as the cohort, then each case can be thought of as belonging to a subcohort of workers with approximately the same year and age of commencing work in the plant. In each year that a case worked, his exposure can be compared with the other members of his subcohort who were working in that year. If the exposure being studied is causally related to the disease of interest, then the cases should, on average, be exposed for longer periods and/or to higher levels of the causative agent than other workers in their subcohorts who did not get the disease. If the exposure is not causally related to the disease then we would expect case exposures to approximate those of their subcohorts. Thus for each case, we calculate an expected and observed exposure, taking into account both exposure level and duration.

The exposure data required for the analysis consist of some measure of exposure to a given chemical for each job in a plant, for each calendar year involved in the study. The exposure measures should ideally be interval measures, but can also consist of ranked exposure data if this is all that is available. Ranked data will be used to illustrate the application of the method with a 0 to 5 scale of exposure for

TABLE 1

Exposure ratings used to classify jobs

0 = *No exposure*-
1 = *Minimal exposure to low levels:* (Chemical in building—not handled, low vapor pressure and dust level, probably works on different floor)
2 = *Moderate exposure:* (Works around the chemical, but exposure is minimal)
3 = *Works in area subject to occasional high excursions:* (Normally exposure is minimal but occasional spills, leaks, or dust exposure may occur)
4 = *Works in areas where level is high:* (Exposure levels in the area are frequently high. Might consider that some risk is involved if chemical is very toxic)
5 = *Intimate contact—skin or high inhalation:* (Such as poly cleaners in the old days—handling slurry)

TABLE 2

Chemical exposure ratings specific for job identification number and calendar year for a given chemical

Chemical #1 Job identification no.	Calendar year					
	1942	1943	1944	1945	1946 . . .	1973
1	0	0	0	2	2	2
2	5	5	5	5	5	5
3	2	2	1	1	0	0
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
84	4	4	4	1	1	1

each of 12 chemicals examined (table 1). Panels of employees familiar with the past conditions in the plant assigned each job an exposure rank for each calendar year of the study (table 2). These exposure data are then linked with work histories identifying the jobs each worker had in the plant and the calendar time period involved. Exposure dose is estimated by multiplying the exposure level and the duration worked at that level. These "doses" are accumulated over each calendar year. A worker exposed at level 1 for all of a calendar year accumulates 365 dose units; if he spends 200 days at level 2, and the remainder of the year at level 1 he accumulates 565 dose units. (Strictly speaking, the doses should be calculated for work days only, but the accumulation over calendar days is easier.) The fact that both the observed and expected doses are obtained in the same manner prevents bias being introduced.

The calculation of the expected dose for each year a case works is obtained from the cohort dose matrix (figure 1). This is a three-dimensional matrix defined by age of first employment, year of first employment, and calendar year. During employment in the industry, each noncase contributes to the relevant cells of this matrix the number of days spent in the given calendar year at each of the possible exposure levels. For example, an employee aged 22 when first employed in 1944 who spends 200 days of that year at level 2 and 165 days at level 1 would contribute to the corresponding cell shown in figure 1 the number of days at the two levels. The properties of the cohort dose matrix are such that entries into any one cell are based on workers of the same age in the corresponding calendar year, who were first employed at the same age.

Once data from each noncase have been entered into the cohort dose matrix, the

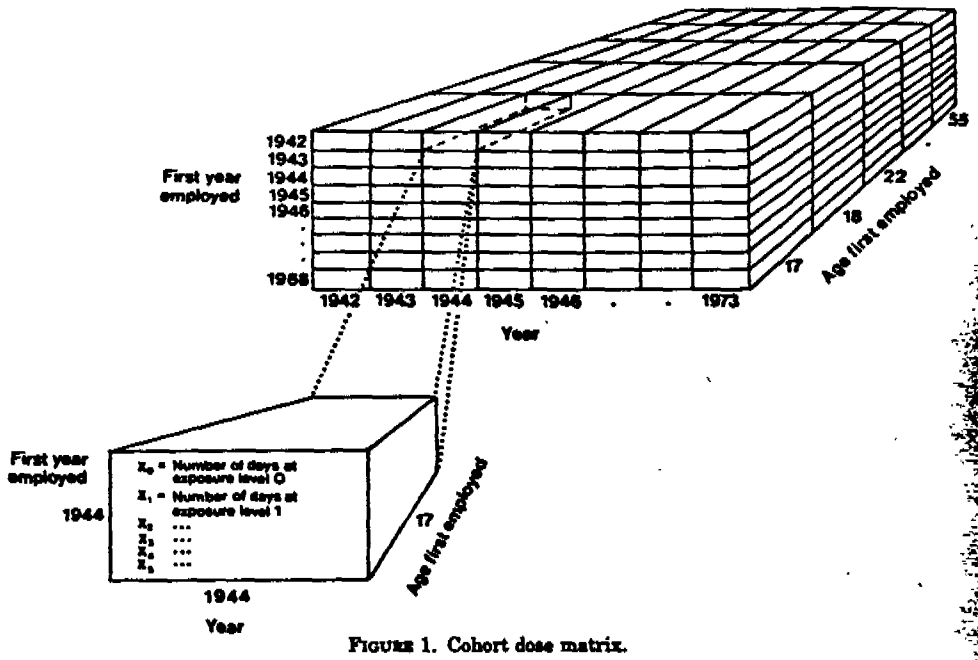

FIGURE 1. Cohort dose matrix.

case exposure dose can be compared with the matrix values. However, this comparison is somewhat restrictive; comparisons might be made with subcohorts having approximately the same year of birth and age at first employment. This is done in the SAED model by using a moving sum accumulation of data in the cohort dose matrix. In the application to be presented we define a subcohort for comparison with a case as consisting of all noncases whose year of first employment and age at first employment are within two years of those of the case. Since in the first year of employment of a case there are no noncases working who were first employed in the two subsequent years, the expected values for this first year are calculated only from noncases who started work in the same year as the case, while in the second year of work the noncases first employed in the year before and the year after the year of hire of the case are included. In subsequent years the estimation is from the whole subcohort within two years of first employm nt and within two years of birth of the case in either direction. When the cohort being studied is small, the moving sum accumulation may need to span more than five years.

The cohort dose matrix is then used to calculate the expected dose for each year that a case works. Table 3 presents the first two years of work for a case who was aged 22 when first employed in 1944. In his first year he worked 115 days, 50 at level 0, followed by 65 at level 5. He thus accumulated a dose of $50 \times 0 + 65 \times 5 = 325$ dose units. The noncase members of his subcohort accumulated 600 days at level 0, 1350 days at level 1, etc. The total number of days of work in 1944 was 3165, and the total subcohort dose was 4830 units, or 1.53 per day. The case worked for 115 days in that year, so his expected dose under the null hypothesis is $115 \times 1.53 = 176$ dose units. His actual exposure was 325 dose units.

The above calculations relate to one

TABLE 3

*Example of data used to calculate the observed and expected dose units for a case in two calendar years**

Year	Exposure level	Case days	Case dose units	Subcohort days	Subcohort dose units
1944	0	50	0	600*	0
	1	0	0	1350	1350
	2	0	0	540	1080
	3	0	0	400	1200
	4	0	0	175	700
	5	65	325	100	500
Total		115	325	3165	4830
1945	0	0	0	1205	0
	1	365	365	800	800
	2	0	0	1500	3000
	3	0	0	340	1020
	4	0	0	110	440
	5	0	0	50	250
Total		365	365	4005	5510

* Because the moving average variable was set at ±2, the value 600 can be interpreted as meaning that the group of persons whose age when first employed was between 20–24 and whose year when first employed was between 1942–1946 worked a total of 600 days at exposure level 0 jobs during 1944.

TABLE 4

The total observed case dose units for each case of angiosarcoma, the expected case dose units, and the difference calculated by the SAED analysis method*

Case	Observed total case dose units	Expected total case dose units	Difference
1	15,728	11,103	+4,625
2	25,765	12,689	+13,076
3	17,145	11,366	+5,779
4	27,232	12,514	+14,718
5	10,264	14,350	−4,086
6	16,413	7,553	+8,860
7	40,412	16,605	+23,807
8	5,375	2,613	+2,762
9	21,820	10,014	+11,806
10	34,372	21,555	+12,817
Average	21,453	12,036	+9,416

* Serially additive expected dose.

year of work. Doses can be summed over a worker's entire work history to give his observed total dose. The expected dose for each year can likewise be summed. Some results are presented in table 4 for the exposure of 10 cases of angiosarcoma of the liver to vinyl chloride monomer. For all but one of the cases the observed total doses are greater than the expected total doses with a mean difference of 9417 (p = 0.004 by paired t-test, p = 0.01 by Wilcoxin signed ranks test).

The SAED analysis also facilitates examination of the relationship between dates of exposure and dates of diagnosis (alternatively, dates of death) of the cases.

TABLE 5

The total case dose units of vinyl chloride for each angiosarcoma case and the expected total case dose units

No. of years prior to death	No. of cases working (total of 10 cases)	Case days of work	Subcohort days of work	Case dose units	Expected case dose units	Difference between observed and expected case dose units
1	2	299	23,623	890	563	327
2	4	1,082	87,533	3,112	1,942	1,170
3	7	2,555	171,572	7,665	4,337	3,328
4	9	3,196	224,096	10,132	5,345	4,787
5	9	3,285	226,932	9,945	5,472	4,473
6	9	3,285	229,741	9,855	5,540	4,315
7	9	3,285	233,626	10,495	5,568	4,927
8	9	3,285	239,066	11,471	5,700	5,771
9	9	3,125	243,553	11,277	5,399	5,878
10	9	3,285	247,948	10,500	5,773	4,727
11	9	3,285	251,083	10,236	5,944	4,292
12	9	3,285	257,741	10,036	5,988	4,048
13	9	3,285	259,848	10,877	6,122	4,755
14	9	3,269	249,430	11,293	6,197	5,096
15	8	2,803	279,338	10,426	5,380	5,046
16	8	2,749	293,751	9,588	5,323	4,265
17	8	2,570	275,032	7,615	4,794	2,821
18	8	2,458	261,655	6,260	4,597	1,663
19	8	2,920	291,071	7,756	5,618	2,138
20	8	2,495	236,498	8,082	4,945	3,137
21	6	1,848	150,202	6,443	3,636	2,807
22	4	1,209	103,944	3,670	2,454	1,216
23	3	1,095	107,942	4,380	2,275	2,105
24	3	1,095	110,643	4,604	2,378	2,226
25	3	1,095	96,261	4,875	2,521	2,354
26	3	753	83,802	3,400	1,646	1,754
27	2	730	79,025	3,285	1,669	1,617
28	2	730	63,288	3,493	1,778	1,715
29	2	399	46,529	1,995	972	1,023
30	1	174	14,119	870	487	383
Totals		64,929	5,438,792	214,526	120,362	94,164

For each case in the study one can calculate for each year prior to death his observed and expected dose. (If the case does not work in a given year, both observed and expected are set to zero.) Table 5 shows the output from the latency analysis program for the 10 angiosarcoma cases. Column seven, the difference between the observed case dose and the expected case dose per case in the study, is plotted in figure 2 after dividing by 10 to obtain the average per case. It can be seen from this plot that the greatest average difference occurred nine years prior to death, and that it was over 400 dose units in the interval 4–16 years prior to diagnosis. Excess exposures to vinyl chloride were mostly accumulated by the cases in this study 4–16 years prior to death, although some cases were exposed 30 years prior to death.

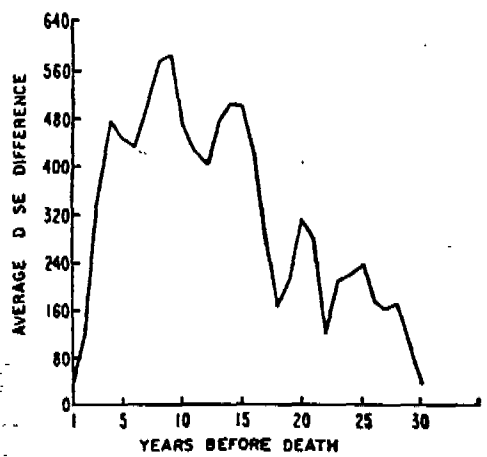

FIGURE 2. Average difference between observed angiosarcoma case dose units and expected dose units of vinyl chloride.

A search can also be made for a dose-response relationship by an exposure level analysis comparing the number of days each case has worked at each level of exposure, with the expected number of days. For example, consider a case who worked for 200 days in 1965 at dose level 2. If his subcohort had a total of 20,000 work days during that calendar year with 2000 at dose level 2, then his expected number of days during 1965 at dose level 2 is $365 \times 2000/20{,}000 = 36.5$. The observed and expected values can be accumulated over the total work history, and the values for each case added. Results for vinyl chloride and 10 cases of angiosarcoma of the liver are shown in table 6. It can be seen that cases worked longer than expected at levels 3, 4 and 5. Any excess at some levels must be accompanied by a deficit at other levels and the deficit here is at levels 0, 1 and 2. An interesting and unexpected finding was the strong association with exposure level 3, which includes jobs involving occasional high excursions from normally minimal levels (table 1).

The model can be adapted to assess whether or not more than one chemical is involved. The angiosarcoma cases were investigated for 19 individual chemicals. While the strongest relationship in terms

TABLE 6

Comparison of observed and expected numbers of days the 10 angiosarcoma cases worked at each exposure level of vinyl chloride

	Dose level						Total
	0	1	2	3	4	5	
Case days	7,467	369	6,539	14,249	23,193	13,112	64,929
Expected case days	12,831	15,633	21,879	1,933	8,091	4,562	64,929
Ratio	0.58	0.02	0.30	7.37	2.87	2.87	
Difference	−5,364	−15,264	−15,340	+12,316	+15,102	+8,550	

of statistical significance was with vinyl chloride, an association was also noted with caprylyl chloride exposure ($p = 0.005$, t-test). This raises a question as to which is more likely to be the carcinogen; or if both are carcinogens, whether they are acting independently or jointly. To answer this question, an additional dimension was added to the model so that each cell from the model shown in figure 1 becomes a two-dimensional matrix (table 7) giving the number of days of exposure for the cohort members at each combination of levels of the two chemicals. This information enables calculation of the conditional expectation of exposure. If a case worked in a job with exposure 2 to vinyl chloride, then the expected exposure to caprylyl chloride was calculated from data in the corresponding column vector (see table 7). Observed and expected doses were accumulated with the results shown in table 8. It can be seen that, given the exposure to vinyl chloride, the observed case exposure to caprylyl chloride was very close to that expected. In contrast, given the exposure levels to caprylyl chloride, there remained some difference between observed exposure to vinyl chloride and that expected. Thus, in spite of the fact that these exposures are correlated in that jobs with exposure to vinyl chloride tend also to involve exposure to caprylyl chloride, the model suggests that caprylyl chloride is not involved. If a difference had remained between observed and expected exposure doses to caprylyl chloride given vinyl chloride exposures, then differences could have been calculated for each level of vinyl chloride exposure to attempt to differentiate between independent and joint action.

TABLE 7

Representation of a cell used in the cohort dose matrix to assess two chemicals, vinyl chloride and caprylyl chloride. X_{ij} is the number of days for the corresponding calendar year, age first employed, and year first employed, that were spent at level i of caprylyl and level j of vinyl chloride exposure

	Vinyl chloride level					
	0	1	2	3	4	5
Caprylyl chloride level						
0	X_{00}	X_{01}	X_{02}	X_{03}	X_{04}	X_{05}
1	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
2	X_{20}	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}
3	X_{30}	X_{31}	X_{32}	X_{33}	X_{34}	X_{35}
4	X_{40}	X_{41}	X_{42}	X_{43}	X_{44}	X_{45}
5	X_{50}	X_{51}	X_{52}	X_{53}	X_{54}	X_{55}

DISCUSSION

The method presented deals with several problems of occupational cancer studies. Worker movement between jobs and variation in job exposure levels according to calendar years are incorporated into the estimation of exposure dose. Latency intervals and exposure levels associated with the cancer can be estimated. Several different chemical associations can be sought for and the possibility that two chemicals might be acting jointly to induce cancer can be investigated. This information is obtained in a manner which avoids bias from potential confounding due to differences between

TABLE 8

Observed, and conditional and unconditional expected case dose units of exposure to vinyl chloride and caprylyl chloride for angiosarcoma cases. The p values for differences between observed and expected values are in parentheses

	Observed dose units	Unconditional expected dose units	Conditional expected dose units given exposure levels to the other chemical
Vinyl chloride	21,433	12,036 (0.004)	19,450 (0.055)
Caprylyl chloride	11,493	3,730 (0.005)	11,579 (0.9)

cases and noncases in age, calendar year of exposure, and age when first employed. Other variables such as race and smoking can be included when appropriate by deriving cohort dose matrices for each race group and for smokers and nonsmokers separately.

The method shares certain features of cohort studies and certain features of case-control studies. Cases are ascertained for a defined work population or cohort, and all work histories are utilized. However, the analysis follows the case-control study pattern. The advantages of using case-control methods of analysis of cohort data include the fact that bias due to potential confounding variables can be avoided without the assumptions required for multivariate analyses (5). A particular advantage with occupational studies is that observed and expected exposures can be derived for each calendar year, a variable which is difficult to incorporate into a multivariate model.

Exposure variables currently employed in occupational health studies include the simple binary categorization of workers according to whether they worked in a given job or plant area for any time at all, or for a minimum period such as five years (6–8). Occasionally estimates of exposure to a postulated causal agent are made for each job in an industry and disease incidence is compared between those exposed at different levels (7, 9). Cumulative dose as in the method presented has been used in studies of the effects of radiation (10) and asbestos exposure (11) and gives a more precise index of exposure resulting in greater power in testing hypotheses concerning occupational carcinogenesis. Collection of additional information is required, but the costs of collecting this information are relatively small compared with the total costs of occupational cancer studies. In the future, health surveillance systems may be implemented in many industrial plants and the type of information required will become more readily available (12, 13). In the meantime, retrospective assessment of exposure by means of company records of chemicals and processes used and interviews with personnel familiar with past conditions can generate data enhancing the power of occupational cancer studies.

Cancer latency is usually defined starting from the point in time of first exposure, although attention may also be given to the midpoint or end of the exposure period (14–16). Rather than fixing attention to these points in exposure time, the SAED model allows one to examine the period of time during which the exposure of the cancer cases is greater than that of comparable noncases. The advantages and disadvantages of such information concerning cancer latency have been discussed in greater detail elsewhere (17). Calculation of the average and range of the times between first exposure or last exposure and diagnosis in occupational cancer studies is influenced by the fact that, unlike angiosarcoma of the liver, most cancers have relatively low occupational attributable risks and some of the occupationally exposed cases will usually have had their cancer induced by nonoccupational causes leading to possible bias in measures of latency. Estimation of latency using the SAED model outlined deals with this situation since, on average, the occupational exposure of cases whose cancer was nonoccupationally induced should be the same as that of comparable noncase workers. If the occupational exposure is indeed causal, the years during which the average case exposure is greater than that expected identifies the period during which causal exposures were accumulated.

The use of the SAED model with ranked exposure data can be criticized since the dose accumulation assumes a continuous scale. However, the SAED model itself is not dependent on having ranked data. When exposure measures are available

it is suggested that they be placed in rank order and categorized into approximately five exposure steps. Jobs for which measures are not available could be assigned to an exposure level in each calendar year on the basis of information concerning manufacturing processes and chemicals used and the accumulation of doses would involve the average of available exposure measures pertaining to each level. When no exposure measures are available, the justification for using subjectively determined ranked data as if it were on a continuous scale is that this should generally lead to increased statistical power over the use of binary exposure data. Increased statistical power will result if those jobs assigned the higher exposure rankings do indeed involve noticeably more exposure than those assigned the lower rankings. The exposure measure becomes virtually continuous once accumulated over calendar time and the paired *t*-test, or Wilcoxin's test, can be used to test differences between observed and expected values.

A second statistical criticism of the method relates to the fact that the same noncase can contribute in at least some years to the estimation of the expected exposure doses of more than one case. Each expected exposure dose is calculated from many noncases. The predominant source of the variance of the difference between observed and expected case doses is the variance of the observed case doses. The contribution of some noncases to the estimation of expected doses for more than one case can therefore have very little effect on the *t*-test or Wilcoxin's test used to measure statistical significance.

We claim that the wealth of information generated by use of the SAED model outweighs these statistical problems. The model allows one to test for the association of individual chemicals with cancer, examine dose response relationships and latency phenomena, and assess if more than one chemical is involved. We have illustrated its use with a known association between vinyl chloride and angiosarcoma of the liver, and have also used it in a study of lung cancer in the same plant (18) in conjunction with methods generating risk estimates. The SAED model, or adaptations of it, may have widespread application in future occupational health surveillance systems, in addition to its use for retrospective studies.

References

1. Enterline PE: Pitfalls in epidemiological research. An examination of the asbestos literature. J Occup Med 18:150–156, 1976
2. Gamble J, Spirtas R: Job classification and utilization of complete work histories in occupational epidemiology. J Occup Med 18:299–404, 1976
3. Mancuso TF, Coulter EJ: Methodology in industrial health studies. Arch Environ Health 6:210–226, 1963
4. Redmond CK, Breslin PP: Comparison of methods for assessing occupational hazards. J Occup Med 17:313–317, 1975
5. Smith AH, Kark JD, Cassel JC, et al: Analysis of prospective epidemiologic studies by minimum distance case-control matching. Am J Epidemiol 105:567–574, 1977
6. Lloyd JW, Lundin FE, Redmond CK, et al: Long-term mortality study of steel workers. IV. Mortality by work area. J Occup Med 12:151–157, 1970
7. McMichael AJ, Spirtas R, Kupper LL, et al: Solvent exposure and leukemia among rubber workers: An epidemiologic study. J Occup Med 17:234–239, 1975
8. Waxweiler RJ, Stringer W, Wagoner JK, et al: Neoplastic risk among workers exposed to vinyl chloride. Annals NY Acad Sci 271:40–48, 1976
9. Tabershaw IR, Gaffey WR: Mortality study of workers in the manufacture of vinyl chloride and its polymers. J Occup Med 16:509–518, 1974
10. Lundin FE, Wagoner JK, Archer VE: Radon Daughter Exposure and Respiratory Cancer. Quantitative and Temporal Aspects. US Dept HEW, 1971
11. Enterline P, DeCoufle P, Henderson V: Mortality in relation to occupational exposure in the asbestos industry. J Occup Med 12:897–903, 1972
12. Barrett CD, Belk HD: A computerized occupational medical surveillance program. J Occup Med 19:732–736, 1977
13. Kerr PS: Recording occupational health data for future analysis. J Occup Med 20:197–203, 1978
14. Armenian HK, Lilienfeld AM: The distribution

of incubation periods of neoplastic diseases. Am J Epidemiol 99:92–100, 1974
15. Cobb S, Miller M, Wald N: On the estimation of the incubation period in malignant disease. J Chronic Dis 9:385–393, 1959
16. Polednak AP: Latency periods in neoplastic diseases. Am J Epidemiol 100:354–356, 1974
17. Smith AH, Checkoway H, Goldsmith DF, et al: An analytic procedure for investigation of cancer latency in matched case control studies illustrated with occupational data. Paper presented at Society for Epidemiologic Research meeting, June 1978. Am J Epidemiol 108:226 (abstract), 1978
18. Waxweiler RJ: An epidemiologic investigation of lung cancer in a multixenobiotic occupational environment. PhD Dissertation, University of North Carolina, 1978

AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol. 112, No. 6
Printed in U.S.A.

EPIDEMIOLOGIC INVESTIGATION OF OCCUPATIONAL CARCINOGENESIS USING A SERIALLY ADDITIVE EXPECTED DOSE MODEL

A. H. SMITH,[1,2] R. J. WAXWEILER[3] AND H. A. TYROLER[1]

Smith, A. H. (Dept. of Community Health, Wellington Clinical School of Medicine, Wellington Hospital, Wellington 2, New Zealand), R. J. Waxweiler and H. A. Tyroler. Epidemiologic investigation of occupational carcinogenesis using a serially additive expected dose model. *Am J Epidemiol* 112:787–797, 1980.

The epidemiologic identification of occupational carcinogens is complicated by several problems including worker mobility between jobs, variation over time of chemicals and processes used, and the long latency period between exposure and discovery of a tumor. In the light of these problems, a method using the cumulative dose concept has been developed which involves calculating the expected yearly exposure for each case from work histories of all noncases close to the case in year of birth and year of hire. The data required for use of the method include information concerning exposure to the chemicals being studied for each job in each calendar year of the study. Use of the method is illustrated with a study of angiosarcoma of the liver and vinyl chloride exposure in a polymerization plant. The value of the method lies in the wealth of information generated concerning the association between chemical exposures and cancer, including exposure level relationships, latency information, and the possibility that two chemicals might be acting independently or jointly. The serially additive expected dose model is likely to prove particularly useful in the analysis of data collected by occupational health surveillance systems, as well as retrospective studies of the type illustrated.

angiosarcoma; cancer; epidemiologic methods; occupational diseases; vinyl chloride

The epidemiologic identification of occupational carcinogens is generally complicated by several features of the exposure of workers to industrial chemicals, and also by certain aspects of the carcinogenic process itself. The major problems have been presented and discussed elsewhere (1–4) and will only be briefly outlined here. Our main objective is to present a method of data analysis which deals with many of these problems in the identification of likely carcinogenic agents in the work environment. The model used requires exposure information which, although not usually obtained in occupational studies, is generally available with the cooperation of the companies involved.

A major problem in occupational cancer studies is that worker movement from job

Received for publication December 13, 1979, and in final form March 25, 1980.

Abbreviation: SAED; serially additive expected dose.

[1] Occupational Health Studies Group, University of North Carolina at Chapel Hill, NC.

[2] Reprint requests to Dr. A. H. Smith. Current address: Department of Community Health, Wellington Clinical School of Medicine, Wellington Hospital, Wellington 2, New Zealand.

[3] National Institute for Occupational Safety and Health, Cincinnati, OH.

The helpful suggestions by Dr. David Deubner are gratefully acknowledged.

to job within industrial plants is accompanied by variation in exposure to any suspect carcinogen. In some instances movement between different companies is sufficiently frequent to add to this problem. In addition, the chemicals used by workers in individual jobs change over time with the introduction of new production methods and with changes in the products themselves. In many industries, workers are exposed to a large number of different chemicals, further complicating the isolation of carcinogenic agents. Even when the same chemicals are used in the same amounts, changes in work practices and plant ventilation alter exposure levels.

With regard to the carcinogenic process itself, the major complicating factor is the long latency period between exposure to a carcinogen and the diagnosis of cancer. Unless ongoing surveillance systems are operational, this means that exposure information must be sought for 20–30 years prior to conducting a study. The risk of cancer may depend on the level, duration, and continuity of exposure, and accurate information of this kind may be difficult to obtain retrospectively over such a long period of time. An additional complicating feature of the carcinogenic process is that more than one chemical may be involved, either independently, or jointly in an interactive manner.

The method of data analysis which we present here was constructed in an attempt to deal with as many of these problems as possible. Much of the data required is obtained routinely in occupational cancer studies other than those which merely compare disease incidence or mortality with that in the general population. The additional information required is the average exposure level for each job in each calendar year under study, although the method can still be used when classification is limited to exposed and nonexposed jobs. Exposure information allowing a simple ranking of job exposure can be utilized and was made available for this study by the company from interviews of personnel with a knowledge of past and current jobs within the study plant and from records of chemical usage. The method can also incorporate actual chemical exposure measures when they are available.

Method

The basis of the serially additive expected dose (SAED) method in descriptive terms is to compare the observed exposure of each case in a study with the exposures of fellow workers close to the case in year of birth, and in age at commencement of work for the company in the study. If the total work force in a plant is referred to as the cohort, then each case can be thought of as belonging to a subcohort of workers with approximately the same year and age of commencing work in the plant. In each year that a case worked, his exposure can be compared with the other members of his subcohort who were working in that year. If the exposure being studied is causally related to the disease of interest, then the cases should, on average, be exposed for longer periods and/or to higher levels of the causative agent than other workers in their subcohorts who did not get the disease. If the exposure is not causally related to the disease then we would expect case exposures to approximate those of their subcohorts. Thus for each case, we calculate an expected and observed exposure, taking into account both exposure level and duration.

The exposure data required for the analysis consist of some measure of exposure to a given chemical for each job in a plant, for each calendar year involved in the study. The exposure measures should ideally be interval measures, but can also consist of ranked exposure data if this is all that is available. Ranked data will be used to illustrate the application of the method with a 0 to 5 scale of exposure for

TABLE 1

Exposure ratings used to classify jobs

0 = *No exposure*
1 = *Minimal exposure to low levels:* (Chemical in building—not handled, low vapor pressure and dust level, probably works on different floor)
2 = *Moderate exposure:* (Works around the chemical, but exposure is minimal)
3 = *Works in area subject to occasional high excursions:* (Normally exposure is minimal but occasional spills, leaks, or dust exposure may occur)
4 = *Works in areas where level is high:* (Exposure levels in the area are frequently high. Might consider that some risk is involved if chemical is very toxic)
5 = *Intimate contact—skin or high inhalation:* (Such as poly cleaners in the old days—handling slurry)

TABLE 2

Chemical exposure ratings specific for job identification number and calendar year for a given chemical

Chemical #1 Job identification no.	Calendar year					
	1942	1943	1944	1945	1946 . . .	1973
1	0	0	0	2	2	2
2	5	5	5	5	5	5
3	2	2	1	1	0	0
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
84	4	4	4	1	1	1

each of 12 chemicals examined (table 1). Panels of employees familiar with the past conditions in the plant assigned each job an exposure rank for each calendar year of the study (table 2). These exposure data are then linked with work histories identifying the jobs each worker had in the plant and the calendar time period involved. Exposure dose is estimated by multiplying the exposure level and the duration worked at that level. These "doses" are accumulated over each calendar year. A worker exposed at level 1 for all of a calendar year accumulates 365 dose units; if he spends 200 days at level 2, and the remainder of the year at level 1 he accumulates 565 dose units. (Strictly speaking, the doses should be calculated for work days only, but the accumulation over calendar days is easier.) The fact that both the observed and expected doses are obtained in the same manner prevents bias being introduced.

The calculation of the expected dose for each year a case works is obtained from the cohort dose matrix (figure 1). This is a three-dimensional matrix defined by age of first employment, year of first employment, and calendar year. During employment in the industry, each noncase contributes to the relevant cells of this matrix the number of days spent in the given calendar year at each of the possible exposure levels. For example, an employee aged 22 when first employed in 1944 who spends 200 days of that year at level 2 and 165 days at level 1 would contribute to the corresponding cell shown in figure 1 the number of days at the two levels. The properties of the cohort dose matrix are such that entries into any one cell are based on workers of the same age in the corresponding calendar year, who were first employed at the same age.

Once data from each noncase have been entered into the cohort dose matrix, the

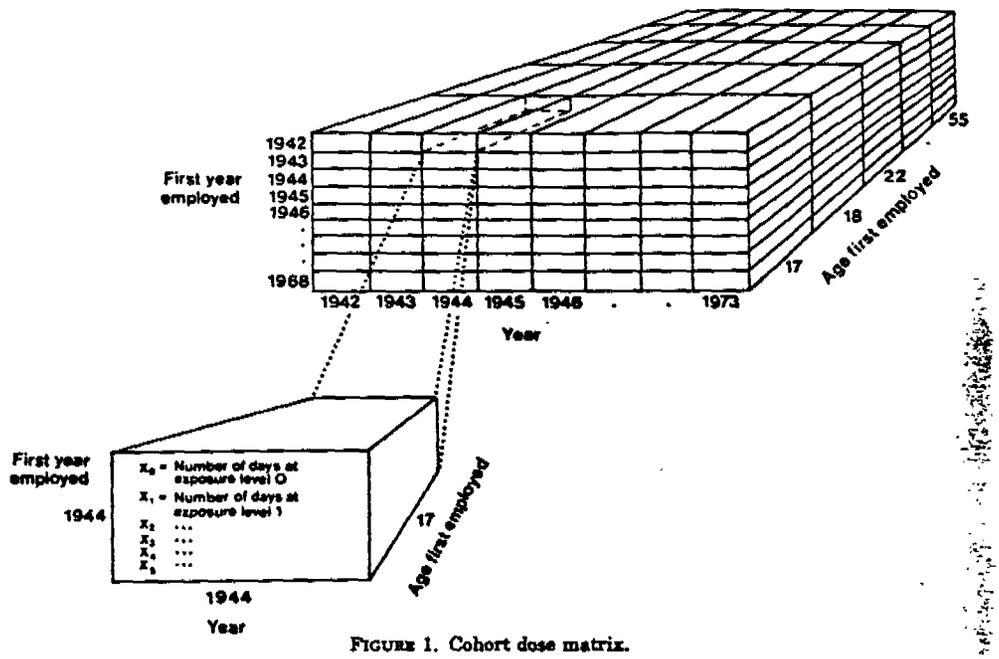

FIGURE 1. Cohort dose matrix.

case exposure dose can be compared with the matrix values. However, this comparison is somewhat restrictive; comparisons might be made with subcohorts having approximately the same year of birth and age at first employment. This is done in the SAED model by using a moving sum accumulation of data in the cohort dose matrix. In the application to be presented we define a subcohort for comparison with a case as consisting of all noncases whose year of first employment and age at first employment are within two years of those of the case. Since in the first year of employment of a case there are no noncases working who were first employed in the two subsequent years, the expected values for this first year are calculated only from noncases who started work in the same year as the case, while in the second year of work the noncases first employed in the year before and the year after the year of hire of the case are included. In subsequent years the estimation is from the whole subcohort within two years of first employment and within two years of birth of the case in either direction. When the cohort being studied is small, the moving sum accumulation may need to span more than five years.

The cohort dose matrix is then used to calculate the expected dose for each year that a case works. Table 3 presents the first two years of work for a case who was aged 22 when first employed in 1944. In his first year he worked 115 days, 50 at level 0, followed by 65 at level 5. He thus accumulated a dose of $50 \times 0 + 65 \times 5 = 325$ dose units. The noncase members of his subcohort accumulated 600 days at level 0, 1350 days at level 1, etc. The total number of days of work in 1944 was 3165, and the total subcohort dose was 4830 units, or 1.53 per day. The case worked for 115 days in that year, so his expected dose under the null hypothesis is $115 \times 1.53 = 176$ dose units. His actual exposure was 325 dose units.

The above calculations relate to one

TABLE 3

*Example of data used to calculate the observed and expected dose units for a case in two calendar years**

Year	Exposure level	Case days	Case dose units	Subcohort days	Subcohort dose units
1944	0	50	0	600*	0
	1	0	0	1350	1350
	2	0	0	540	1080
	3	0	0	400	1200
	4	0	0	175	700
	5	65	325	100	500
Total		115	325	3165	4830
1945	0	0	0	1205	0
	1	365	365	800	800
	2	0	0	1500	3000
	3	0	0	340	1020
	4	0	0	110	440
	5	0	0	50	250
Total		365	365	4005	5510

* Because the moving average variable was set at ±2, the value 600 can be interpreted as meaning that the group of persons whose age when first employed was between 20–24 and whose year when first employed was between 1942–1946 worked a total of 600 days at exposure level 0 jobs during 1944.

TABLE 4

The total observed case dose units for each case of angiosarcoma, the expected case dose units, and the difference calculated by the SAED analysis method*

Case	Observed total case dose units	Expected total case dose units	Difference
1	15,728	11,103	+4,625
2	25,765	12,689	+13,076
3	17,145	11,366	+5,779
4	27,232	12,514	+14,718
5	10,264	14,350	−4,086
6	16,413	7,553	+8,860
7	40,412	16,605	+23,807
8	5,375	2,613	+2,762
9	21,820	10,014	+11,806
10	34,372	21,555	+12,817
Average	21,453	12,036	+9,416

* Serially additive expected dose.

year of work. Doses can be summed over a worker's entire work history to give his observed total dose. The expected dose for each year can likewise be summed. Some results are presented in table 4 for the exposure of 10 cases of angiosarcoma of the liver to vinyl chloride monomer. For all but one of the cases the observed total doses are greater than the expected total doses with a mean difference of 9417 (p = 0.004 by paired t-test, p = 0.01 by Wilcoxin signed ranks test).

The SAED analysis also facilitates examination of the relationship between dates of exposure and dates of diagnosis (alternatively, dates of death) of the cases.

TABLE 5

The total case dose units of vinyl chloride for each angiosarcoma case and the expected total case dose units

No. of years prior to death	No. of cases working (total of 10 cases)	Case days of work	Subcohort days of work	Case dose units	Expected case dose units	Difference between observed and expected case dose units
1	2	299	23,523	890	563	327
2	4	1,082	87,533	3,112	1,942	1,170
3	7	2,555	171,572	7,665	4,337	3,328
4	9	3,196	224,096	10,132	5,345	4,787
5	9	3,285	226,932	9,945	5,472	4,473
6	9	3,285	229,741	9,855	5,540	4,315
7	9	3,285	233,626	10,495	5,568	4,927
8	9	3,285	239,066	11,471	5,700	5,771
9	9	3,125	243,553	11,277	5,399	5,878
10	9	3,285	247,948	10,500	5,773	4,727
11	9	3,285	251,083	10,236	5,944	4,292
12	9	3,285	257,741	10,036	5,988	4,048
13	9	3,285	259,848	10,877	6,122	4,755
14	9	3,269	249,430	11,293	6,197	5,096
15	8	2,803	279,338	10,426	5,380	5,046
16	8	2,749	293,751	9,588	5,323	4,265
17	8	2,570	275,032	7,615	4,794	2,821
18	8	2,458	261,655	6,260	4,597	1,663
19	8	2,920	291,071	7,756	5,618	2,138
20	8	2,495	236,498	8,062	4,945	3,137
21	6	1,848	150,202	6,443	3,636	2,807
22	4	1,209	103,944	3,670	2,454	1,216
23	3	1,095	107,942	4,380	2,275	2,105
24	3	1,095	110,643	4,604	2,378	2,226
25	3	1,095	96,261	4,875	2,521	2,354
26	3	753	83,802	3,400	1,646	1,754
27	2	730	79,025	3,285	1,668	1,617
28	2	730	63,288	3,493	1,778	1,715
29	2	399	46,529	1,995	972	1.023
30	1	174	14,119	870	487	383
Totals		64,929	5,436,792	214,526	120,362	94,164

For each case in the study one can calculate for each year prior to death his observed and expected dose. (If the case does not work in a given year, both observed and expected are set to zero.) Table 5 shows the output from the latency analysis program for the 10 angiosarcoma cases. Column seven, the difference between the observed case dose and the expected case dose per case in the study, is plotted in figure 2 after dividing by 10 to obtain the average per case. It can be seen from this plot that the greatest average difference occurred nine years prior to death, and that it was over 400 dose units in the interval 4–16 years prior to diagnosis. Excess exposures to vinyl chloride were mostly accumulated by the cases in this study 4–16 years prior to death, although some cases were exposed 30 years prior to death.

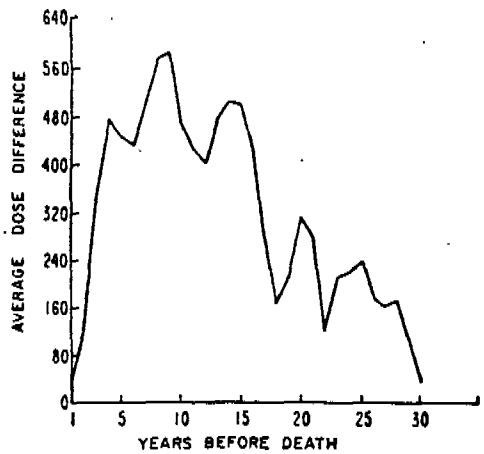

FIGURE 2. Average difference between observed angiosarcoma case dose units and expected dose units of vinyl chloride.

A search can also be made for a dose-response relationship by an exposure level analysis comparing the number of days each case has worked at each level of exposure, with the expected number of days. For example, consider a case who worked for 200 days in 1965 at dose level 2. If his subcohort had a total of 20,000 work days during that calendar year with 2000 at dose level 2, then his expected number of days during 1965 at dose level 2 is $365 \times 2000/20{,}000 = 36.5$. The observed and expected values can be accumulated over the total work history, and the values for each case added. Results for vinyl chloride and 10 cases of angiosarcoma of the liver are shown in table 6. It can be seen that cases worked longer than expected at levels 3, 4 and 5. Any excess at some levels must be accompanied by a deficit at other levels and the deficit here is at levels 0, 1 and 2. An interesting and unexpected finding was the strong association with exposure level 3, which includes jobs involving occasional high excursions from normally minimal levels (table 1).

The model can be adapted to assess whether or not more than one chemical is involved. The angiosarcoma cases were investigated for 19 individual chemicals. While the strongest relationship in terms

TABLE 6

Comparison of observed and expected numbers of days the 10 angiosarcoma cases worked at each exposure level of vinyl chloride

	Dose level						Total
	0	1	2	3	4	5	
Case days	7,467	369	6,539	14,249	23,193	13,112	64,929
Expected case days	12,831	15,633	21,879	1,933	8,091	4,562	64,929
Ratio	0.58	0.02	0.30	7.37	2.87	2.87	
Difference	−5,364	−15,264	−15,340	+12,316	+15,102	+8,550	

of statistical significance was with vinyl chloride, an association was also noted with caprylyl chloride exposure ($p = 0.005$, t-test). This raises a question as to which is more likely to be the carcinogen; or if both are carcinogens, whether they are acting independently or jointly. To answer this question, an additional dimension was added to the model so that each cell from the model shown in figure 1 becomes a two-dimensional matrix (table 7) giving the number of days of exposure for the cohort members at each combination of levels of the two chemicals. This information enables calculation of the conditional expectation of exposure. If a case worked in a job with exposure 2 to vinyl chloride, then the expected exposure to caprylyl chloride was calculated from data in the corresponding column vector (see table 7). Observed and expected doses were accumulated with the results shown in table 8. It can be seen that, given the exposure to vinyl chloride, the observed case exposure to caprylyl chloride was very close to that expected. In contrast, given the exposure levels to caprylyl chloride, there remained some difference between observed exposure to vinyl chloride and that expected. Thus, in spite of the fact that these exposures are correlated in that jobs with exposure to vinyl chloride tend also to involve exposure to caprylyl chloride, the model suggests that caprylyl chloride is not involved. If a difference had remained between observed and expected exposure doses to caprylyl chloride given vinyl chloride exposures, then differences could have been calculated for each level of vinyl chloride exposure to attempt to differentiate between independent and joint action.

TABLE 7

Representation of a cell used in the cohort dose matrix to assess two chemicals, vinyl chloride and caprylyl chloride. X_{ij} is the number of days for the corresponding calendar year, age first employed, and year first employed, that were spent at level i of caprylyl and level j of vinyl chloride exposure

	Vinyl chloride level					
	0	1	2	3	4	5
Caprylyl chloride level						
0	X_{00}	X_{01}	X_{02}	X_{03}	X_{04}	X_{05}
1	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
2	X_{20}	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}
3	X_{30}	X_{31}	X_{32}	X_{33}	X_{34}	X_{35}
4	X_{40}	X_{41}	X_{42}	X_{43}	X_{44}	X_{45}
5	X_{50}	X_{51}	X_{52}	X_{53}	X_{54}	X_{55}

DISCUSSION

The method presented deals with several problems of occupational cancer studies. Worker movement between jobs and variation in job exposure levels according to calendar years are incorporated into the estimation of exposure dose. Latency intervals and exposure levels associated with the cancer can be estimated. Several different chemical associations can be sought for and the possibility that two chemicals might be acting jointly to induce cancer can be investigated. This information is obtained in a manner which avoids bias from potential confounding due to differences between

TABLE 8

Observed, and conditional and unconditional expected case dose units of exposure to vinyl chloride and caprylyl chloride for angiosarcoma cases. The p values for differences between observed and expected values are in parentheses

	Observed dose units	Unconditional expected dose units	Conditional expected dose units given exposure levels to the other chemical
Vinyl chloride	21,433	12,036 (0.004)	19,450 (0.055)
Caprylyl chloride	11,493	3,730 (0.005)	11,579 (0.9)

cases and noncases in age, calendar year of exposure, and age when first employed. Other variables such as race and smoking can be included when appropriate by deriving cohort dose matrices for each race group and for smokers and nonsmokers separately.

The method shares certain features of cohort studies and certain features of case-control studies. Cases are ascertained for a defined work population or cohort, and all work histories are utilized. However, the analysis follows the case-control study pattern. The advantages of using case-control methods of analysis of cohort data include the fact that bias due to potential confounding variables can be avoided without the assumptions required for multivariate analyses (5). A particular advantage with occupational studies is that observed and expected exposures can be derived for each calendar year, a variable which is difficult to incorporate into a multivariate model.

Exposure variables currently employed in occupational health studies include the simple binary categorization of workers according to whether they worked in a given job or plant area for any time at all, or for a minimum period such as five years (6–8). Occasionally estimates of exposure to a postulated causal agent are made for each job in an industry and disease incidence is compared between those exposed at different levels (7, 9). Cumulative dose as in the method presented has been used in studies of the effects of radiation (10) and asbestos exposure (11) and gives a more precise index of exposure resulting in greater power in testing hypotheses concerning occupational carcinogenesis. Collection of additional information is required, but the costs of collecting this information are relatively small compared with the total costs of occupational cancer studies. In the future, health surveillance systems may be implemented in many industrial plants and the type of information required will become more readily available (12, 13). In the meantime, retrospective assessment of exposure by means of company records of chemicals and processes used and interviews with personnel familiar with past conditions can generate data enhancing the power of occupational cancer studies.

Cancer latency is usually defined starting from the point in time of first exposure, although attention may also be given to the midpoint or end of the exposure period (14–16). Rather than fixing attention to these points in exposure time, the SAED model allows one to examine the period of time during which the exposure of the cancer cases is greater than that of comparable noncases. The advantages and disadvantages of such information concerning cancer latency have been discussed in greater detail elsewhere (17). Calculation of the average and range of the times between first exposure or last exposure and diagnosis in occupational cancer studies is influenced by the fact that, unlike angiosarcoma of the liver, most cancers have relatively low occupational attributable risks and some of the occupationally exposed cases will usually have had their cancer induced by nonoccupational causes leading to possible bias in measures of latency. Estimation of latency using the SAED model outlined deals with this situation since, on average, the occupational exposure of cases whose cancer was nonoccupationally induced should be the same as that of comparable noncase workers. If the occupational exposure is indeed causal, the years during which the average case exposure is greater than that expected identifies the period during which causal exposures were accumulated.

The use of the SAED model with ranked exposure data can be criticized since the dose accumulation assumes a continuous scale. However, the SAED model itself is not dependent on having ranked data. When exposure measures are available

it is suggested that they be placed in rank order and categorized into approximately five exposure steps. Jobs for which measures are not available could be assigned to an exposure level in each calendar year on the basis of information concerning manufacturing processes and chemicals used and the accumulation of doses would involve the average of available exposure measures pertaining to each level. When no exposure measures are available, the justification for using subjectively determined ranked data as if it were on a continuous scale is that this should generally lead to increased statistical power over the use of binary exposure data. Increased statistical power will result if those jobs assigned the higher exposure rankings do indeed involve noticeably more exposure than those assigned the lower rankings. The exposure measure becomes virtually continuous once accumulated over calendar time and the paired *t*-test, or Wilcoxin's test, can be used to test differences between observed and expected values.

A second statistical criticism of the method relates to the fact that the same noncase can contribute in at least some years to the estimation of the expected exposure doses of more than one case. Each expected exposure dose is calculated from many noncases. The predominant source of the variance of the difference between observed and expected case doses is the variance of the observed case doses. The contribution of some noncases to the estimation of expected doses for more than one case can therefore have very little effect on the *t*-test or Wilcoxin's test used to measure statistical significance.

We claim that the wealth of information generated by use of the SAED model outweighs these statistical problems. The model allows one to test for the association of individual chemicals with cancer, examine dose response relationships and latency phenomena, and assess if more than one chemical is involved. We have illustrated its use with a known association between vinyl chloride and angiosarcoma of the liver, and have also used it in a study of lung cancer in the same plant (18) in conjunction with methods generating risk estimates. The SAED model, or adaptations of it, may have widespread application in future occupational health surveillance systems, in addition to its use for retrospective studies.

References

1. Enterline PE: Pitfalls in epidemiological research. An examination of the asbestos literature. J Occup Med 18:150–156, 1976
2. Gamble J, Spirtas R: Job classification and utilization of complete work histories in occupational epidemiology. J Occup Med 18:299–404, 1976
3. Mancuso TF, Coulter EJ: Methodology in industrial health studies. Arch Environ Health 6:210–226, 1963
4. Redmond CK, Breslin PP: Comparison of methods for assessing occupational hazards. J Occup Med 17:313–317, 1975
5. Smith AH, Kark JD, Cassel JC, et al: Analysis of prospective epidemiologic studies by minimum distance case-control matching. Am J Epidemiol 105:567–574, 1977
6. Lloyd JW, Lundin FE, Redmond CK, et al: Long-term mortality study of steel workers. IV. Mortality by work area. J Occup Med 12:151–157, 1970
7. McMichael AJ, Spirtas R, Kupper LL, et al: Solvent exposure and leukemia among rubber workers: An epidemiologic study. J Occup Med 17:234–239, 1975
8. Waxweiler RJ, Stringer W, Wagoner JK, et al: Neoplastic risk among workers exposed to vinyl chloride. Annals NY Acad Sci 271:40–48, 1976
9. Tabershaw IR, Gaffey WR: Mortality study of workers in the manufacture of vinyl chloride and its polymers. J Occup Med 16:509–518, 1974
10. Lundin FE, Wagoner JK, Archer VE: Radon Daughter Exposure and Respiratory Cancer. Quantitative and Temporal Aspects. US Dept HEW, 1971
11. Enterline P, DeCoufle P, Henderson V: Mortality in relation to occupational exposure in the asbestos industry. J Occup Med 12:897–903, 1972
12. Barrett CD, Belk HD: A computerized occupational medical surveillance program. J Occup Med 19:732–736, 1977
13. Kerr PS: Recording occupational health data for future analysis. J Occup Med 20:197–203, 1978
14. Armenian HK, Lilienfeld AM: The distribution

of incubation periods of neoplastic diseases. Am J Epidemiol 99:92–100, 1974
15. Cobb S, Miller M, Wald N: On the estimation of the incubation period in malignant disease. J Chronic Dis 9:385–393, 1959
16. Polednak AP: Latency periods in neoplastic diseases. Am J Epidemiol 100:354–356, 1974
17. Smith AH, Checkoway H, Goldsmith DF, et al: An analytic procedure for investigation of cancer latency in matched case control studies illustrated with occupational data. Paper presented at Society for Epidemiologic Research meeting, June 1978. Am J Epidemiol 108:226 (abstract), 1978
18. Waxweiler RJ: An epidemiologic investigation of lung cancer in a multixenobiotic occupational environment. PhD Dissertation, University of North Carolina, 1978