RECORD OF SOCIETY OF ACTUARIES
1980 VOL. 6 NO. 4

BIOMETRIC METHODS FOR THE ANALYSIS OF TIME TO VITAL EVENT DATA AND THE ASSESSMENT OF RISK

Moderator: ROBERT J. JOHANSEN. Panelists: GARY KOCH, NORMAN BRESLOW**, REGINA ELANDT-JOHNSON***, NORMAN JOHNSON#, C. RALPH BUNCHER##*

With the increasing expense of intercompany studies on the one hand, and the increasing need to justify underwriting on the other hand, the actuary may need to make greater use of noninsurance studies such as medical follow-ups and clinical trial analyses. Specialized techniques developed for such studies will be described in this concurrent session.

Statisticians have developed specialized tools, some of which are especially adapted to obtain the maximum amount of information from medical follow-up studies which may, because of their nature, involve relatively small groups. Dr. Breslow's paper and the jointly authored paper by Dr. Elandt-Johnson and Dr. Johnson will describe methodology. The latter paper in particular will describe the use of models. It will also compare and contrast the statistician's approaches with the more traditional actuarial methodology.

The actuarial concept of select mortality rates has been noted by epidemiologists in occupational mortality studies of work cohorts. Dr. Buncher will examine these concepts as well as the possibility that work itself is healthful.

Following the presentation by participants from the American Statistical Association, the Discussant will comment from the actuary's viewpoint. Discussion from the floor will follow.

*Dr. Koch, not a member of the Society, is Professor, Department of Biostatistics, University of North Carolina, Chapel Hill.

**Dr. Breslow, not a member of the Society, is Professor, Department of Biostatistics, University of Washington, Seattle, Washington.

***Dr. Elandt-Johnson, not a member of the Society, is Professor, Department of Biostatistics, University of North Carolina, Chapel Hill.

#Dr. Johnson, not a member of the Society, is the Alumni Distinguished Professor, Department of Statistics, University of North Carolina, Chapel Hill.

##Dr. Buncher, not a member of the Society, is Director, Division of Epidemiology and Biostatistics, College of Medicine, University of Cincinnati, Cincinnati, Ohio.

MR. ROBERT J. JOHANSEN: This is the first of two sessions arranged by the American Statistical Assocation. This session and the one on Social Indicators mark the continuation of an interchange of programs between the Society of Actuaries and the American Statistical Association which I have been promoting for a couple of years. In October 1978, the statisticians presented three sessions in the general area of economics at the Society's annual meeting. In August of this year, the Society sponsored two sessions at the Statistical meeting in Houston. The actuaries discussed the new Build and Blood Pressure Study and some recent developments in actuarial methodology relating to analysis of survival studies.

Actuaries and statisticians have a great deal in common in the studies they do and the approaches they take. However, most of the individuals in one sphere are probably unaware of much of the work being done in the other sphere. The purpose of these interchanges is to provide each group with some examples of the work of the other group, and to encourage further inquiry and interchange.

Dr. Gary Koch, the Discussion Leader, will introduce the speakers and provide some comments. Dr. Koch was the principal organizer of this session for the American Statistical Association.

Dr. Koch has been at the University of North Carolina since 1963. Since 1976 he is Professor in the Department of Biostatistics jointly with the North Carolina Highway Safety Research Center. He has been a Visiting Professor at a number of universities and other organizations. He is a Fellow of the American Statistical Association and Chairman of its Biometric Section; Past President of the Biometric Society, Eastern North American Region; a member of the American Public Health Association, and recipient of their 1974 Mortimer Spiegelman Award. Actuaries are well acquainted, of course, with Mort Spiegelman through his textbook.

Before turning the meeting over to Dr. Koch, I would like to clarify one item. The term "censored survival data" which will be used in this afternoon's session may not be familiar to some of you. You may recall from studying Construction of Mortality Tables that individual records exposed to risk formulas usually provide for a partial year exposure for withdrawals during the year and a deduction for those existing at the end of the investigation period. The term "censored" would be applicable to these cases. In a broad sense it applies to termination from any cause other than the cause being studied.

DR. GARY KOCH: I would like to formally thank Bob Johansen and the Society of Actuaries for inviting the Biometric Section of the American Statistical Association to participate in this session.

Biometricians and statisticians in general have substantial interest in the area of time to event data, i.e., the length of time subjects who are exposed to risk of a particular type of condition are free of that condition up until the time that they have it. They see these time to vital event studies in a number of different contexts. Some settings are

clinical trials where subjects are followed until an event like death occurs. The subjects are then evaluated as to comparative survival rates. In other settings, individuals are attempting to maintain a favorable state like remission from a particular disease or actually being in a disease free state until the occurrence of a particular disease.

As a consequence of analyzing many situations where time is the predominant factor, i.e., the statistic studied, the biometricians and statisticians have developed a variety of methods for dealing with statistical questions concerning these situations. In many cases, these methods have been borrowed from actuarial science because many of the concepts in statistics share similar orientations and concerns with those of actuarial science.

The purpose of today's session is to review some of the work which has been going on in this particular area and also to attempt to identify various relationships between the concerns and methods that arose in biomedical research where certain analogies are seen with respect to actuarial science. Some of the dimensions of these concerns refer to the nature of the population that is under study and how that population is dealt with. In some statistical studies, one sees the population as its own phenomenon, and proceeds to do various kinds of descriptive analyses with respect to that population. These are sometimes called data bank type analyses. Other statistical studies involve experimental randomization of subjects to one or two groups. Another type of study is one in which a sample is drawn from some larger population and then an inference is made with respect to that population.

In all of these cases, a particular framework for analysis involves the computation of estimates and then the derivation of standard errors, or other tests of the sampling or measurement error, for the estimates that are obtained. This may involve the production of confidence intervals, significance tests and the like. In actuarial science, these kinds of concepts may have limited application because the data are perceived essentially as a population even though that population is used to make statements of inference of some type about some other population.

In these presentations, we will try to indicate the similarities in the biomedical approach, which involves certain kinds of sampling processes and inference, and the traditional methods and concepts that exist in actuarial studies. The four speakers and the three presentations cut across these particular dimensions. The first presentation will be given by Professor Norman Breslow from the University of Washington. His presentation will review briefly some of the concepts of analysis for clinical trial data, although it has implications for other situations as well. In particular, he will talk about a concept of covariate analysis that has to do with the select concept in actuarial science. Professor Elandt-Johnson and Professor Johnson from the University of North Carolina will try to do some bridging of the gaps between biomedical research and actuarial science. Their presentation will discuss common terms for both frameworks and will indicate some of the phenomena that

are studied in each. In particular, their paper will address competing risks and the multiple decrement problem and the relationships that exist between them. They will also try to draw relationships between what biometricians call the covariance or regression problem and the select problem as it exists in actuarial science.

The third presentation by Dr. Buncher from the University of Cincinnati will focus on an observational population concept known as the Healthy Worker Effect. He will try to interpret it to the extent that it relates to both actuarial and epidemiological concerns.

Professor Norman Breslow will give the first presentation. The title of his discussion is "Statistical Analysis of Medical Follow-up Studies." This is an area in which he has made a number of contributions to statistical literature beginning with work emanating from his graduate training at Stanford University and proceeding on through his work on the faculty at the University of Washington where he has been since 1968. Among his other activities, he was a consultant to the World Health Organization's International Agency for Research on Cancer where he developed a monograph entitled Statistical Methods of Cancer Research - The Analysis of Case Control Studies. In 1978, Dr. Breslow received the Mortimer Spiegelman Award from the American Public Health Association.

DR. NORMAN BRESLOW: Censored survival data arise in many areas of medical research. In clinical trials patients are followed from diagnosis or definitive treatment until relapse or death. The clinical investigator is interested in the statistical distribution of the duration of response or survival, and how it is influenced by treatment and by prognostic factors like age and extent of disease. Data analysis is complicated by the fact that some patients are still alive and well at the study's end, so that the observations of their response times are incomplete or "censored." This means that the well known statistical methods of analysis of variance and linear regression, which have proven so valuable in other areas, cannot be applied without substantial modification.

Similar problems arise with epidemiological studies, for example in occupational health, where the endpoint is the diagnosis of or death from a particular illness and explanatory variables include accumulating industrial exposures. Limitations are imposed on the observation period by death from competing causes. Censorship is encountered also with data from the bioassay of toxic agents in experimental test systems, wherein animals are observed from birth or weaning for the occurrence of cancer or other chronic ailments.

The analysis of such censored survival data has the same goals as in other areas of statistics. First, one wants to characterize the distribution of survival duration in one or more subgroups of the population. Since the numbers of individuals in particular subgroups are rarely large, inferential statistics are needed to decide whether the observed differences are real or are simply due to sampling fluctuations. Mathematical models are used to describe the effects on the survival distribution of different treatment and baseline factors.

For ease of exposition it is convenient to suppose that the endpoint in question is death rather than some other event such as relapse. Let T denote the (random) survival time for an individual sampled from the population under study. The initial goal is to estimate the statistical distribution of T, which may be specified alternatively in terms of the survival function

$$S(t) = P\ (T>t),$$

the hazard function (also known as the force of mortality)

$$\lambda(t) = \lim_{\Delta t \to 0} P[T<t+\Delta t \mid T \geqslant t],$$

or the cumulative hazard function

$$\Lambda(t) = \int_0^t \lambda(u)du = -\log S(t)$$

The hazard function $\lambda(t)$ represents the (instantaneous) mortality rate at t among persons still alive at that time, $\Lambda(t)$ its integral and S(t) the proportion of the population surviving until time t.

If complete data are available for all individuals in a study, an obvious estimate of S(t) is the sample proportion actually observed to survive until that time. However if study subjects are not all kept under observation until death, and some are withdrawn from observation before time t, this proportion will underestimate S(t). An early resolution of the difficulty was provided by application of fundamental actuarial concepts, as embodied in the life table. Typically, the time period is partitioned into a number of intervals, say $[0,t_1)$, $[t_1,t_2),\ldots,$ $[t_{j-1},t_j)$. Let n_j denote the number still under observation at the start of the j^{th} interval, of which d_j die, w_j are "withdrawn alive" and $n_{j+1} = n_j - d_j - w_j$ survive to the start of the next interval. Then the probability of survival to t_j may be estimated by

$$\hat{S}(t_j) = (1-\hat{q}_1)(1-\hat{q}_2)\ldots(1-\hat{q}_j)$$

where $\hat{q}_j = d_j(n_j - \frac{1}{2}w_j)^{-1}$ represents the conditional probability of death during the j^{th} interval. Thus individuals who are withdrawn early on in the study contribute to the estimation of survival only over the intervals during which they are at risk of death. Table 1 illustrates this calculation, which is by now well known to clinicians, epidemiologists and other medical scientists (Cutler and Ederer, 1958).

The term $\frac{1}{2}w_j$ in the denominator of $\hat{q}_j$ is intended to adjust for the fact that individuals withdrawn alive during the interval effectively reduce the number at risk. This is a rough approximation which works best if the intervals are short. A more accurate estimate is obtained in the limit as the intervals become infinitely large in number and small in length. If $t_1 < t_2 < \ldots\ldots t_K \ldots\ldots$ now denote the distinct, ordered times of death, and if d_K deaths occur at t_K among the n_K individuals still at risk just prior to that time, the "product limit" (PL) estimate may be written

$$\hat{S}(t) = \prod_{t_K < t} \left(1 - \frac{d_K}{n_K}\right).$$

TABLE I

ACTUARIAL (LIFE TABLE) ESTIMATE OF THE SURVIVAL DISTRIBUTION

Interval j $[t_{j-1}, t_j)$	Number alive at start n_j	Deaths d_j	Losses w_j	Prob of death $\hat{q}_j = \frac{d_j}{n_j - w_j/2}$	Prob of survival $1 - \hat{q}_j$	Cumulative prob survival $\prod(1-\hat{q}_j)$
0 - 11 mos	15	5	0	0.3333	0.6666	0.6666
12 - 23 mos	10	2	3	0.2353	0.7647	0.5098
24 - 35 mos	5	1	1	0.2222	0.7777	0.3965
36+ mos.	3	-	-	-	-	-

Kaplan and Meier (1958), who introduced the estimate to the statistical community, noted that it had been proposed by Böhmer (1912) at the 7th International Congress of Actuaries. When there is no censorship, it reduces to the simple proportion mentioned earlier, known as the empirical distribution function. The corresponding estimate of the cumulative hazard (Altschuler, 1970) is

$$\hat{\Lambda}(t) = \sum_{t_k \le t} \frac{d_k}{n_k} .$$

In order to make statistical inferences about the true survival function, it is necessary to know something about the distribution of $\hat{S}$. According to the work of Efron (1967), Breslow and Crowley (1974) and Meier (1975), the function $W(t) = \hat{S}(t) - S(t)$, considered as a stochastic process in t, is approximately Gaussian with mean 0 and a covariance kernel which may be estimated for $t \le s$ by

$$\hat{Cov}[W(t), W(s)] = \hat{S}(t)\hat{S}(s) \sum_{t_k \le t} \frac{d_k}{n_k(n_k - d_k)}$$

When t = s this yields the variance function for $\hat{S}(t)$ known as Greenwood's (1926) formula.

This result leads to the construction of simultaneous confidence bands for S, and to a Kolmogorov-Smirnov type of test for its equality with some particular S_0 (Hall and Wellner, 1980). More precisely, defining

$$\hat{C}(t) = n_0 \sum_{t_k \le t} \frac{d_k}{n_k(n_k - d_k)}$$

where n_0 is the total sample size (number alive and under observation at $t_0 = 0$), it may be shown that

$$P\left[\sqrt{n_0}\,|\hat{S}(t) - S(t)| \le \lambda \hat{S}(t)\{1 + \hat{C}(t)\}/\hat{C}(t) \text{ for } 0 \le t \le T\right]$$

is approximately equal to

$$P\left[|B(u)| \le \lambda \text{ for } 0 \le u \le \hat{C}(T)\{1 + \hat{C}(T)\}^{-1}\right],$$

where B(u) is the stochastic process known as the Brownian bridge. These latter probabilities may be calculated using results of Anderson (1960).

Medical research typically involves the comparison of one group of subjects with another: the survival of patients treated with Drug A is compared to survival with Drug B, or the cancer incidence of underground miners is compared with that of surface miners. If infinite resources were available, one could approach the comparison by making separate estimates of the survival function for each subgroup of the population under study. However this is rarely feasible with the small samples available to the researcher. He may wish to investigate the effect of several different factors on survival yet can collect information on only

a few hundred deaths. Clearly what is needed in this situation is a mathematical model which allows him to express the effects of group membership or other individual characteristics on the survival distribution in terms of a relatively small number of parameters.

Data typically available for each subject in a longitudinal study consist of a survival or follow-up time t, an indicator δ which denotes whether the subject actually died ($\delta=1$) or rather contributed a censored observation ($\delta=0$), and a number of treatment and covariables z_1, $z_2,\ldots z_p$ whose effect on outcome is to be determined. The most widely used and successful statistical models express the effect of the covariables on outcome by means of a linear function

$$\beta z = \beta_1 z_1 + \ldots + \beta_p z_p,$$

where the β parameters, or unknown regression coefficients, are estimated from the data. A wide range of relationships may be expressed in this form by including among the z's both transformations and cross-products involving some basic set of variables. Thus in clinical trials the available information is summarized into a single "prognostic score" βz for each patient which is used to estimate his survival duration.

Since the survival times are all positive, it is inconvenient and somewhat unreasonable to suppose that βz acts additively on them. However the idea that $\exp(\beta z)$ acts multiplicatively on t leads to the accelerated aging model

$$\log t = \alpha + \beta z + \sigma w$$

where w denotes a random error variable (mean 0, variance 1) and σ and α are unknown parameters of scale and location, respectively. Various parametric specifications for the error distribution lead to such familiar survival distributions as the log-normal and Weibull.

Recent emphasis has been on partially non-parametric formulations of the model in which the error distribution is left unspecified. This leads to consideration of the distribution of the rank vector, appropriately generalized to censored data, of the residuals $\log t - \beta z$ (Kalbfleisch and Prentice, 1980; § 6.4). Unfortunately, estimation of the β regression coefficients using this approach is rather cumbersome, and it has not seen much application to date.

Cox (1972) proposed an alternative regression model in which $\exp(\beta z)$ acts multiplicatively on the hazard function rather than on the survival time. The hazard for an individual with covariates z is written

$$\lambda(t;z) = \exp(\beta z)\lambda_0(t) ,$$

where λ_0 is the hazard for an individual with a standard (z=0) set of covariates. A consequence is that the hazard functions for two individuals with different covariate vectors are proportional, meaning that their ratio is constant in time. The relation among survival functions under this proportional hazards (PH) model is

$$S(t;z) = S_0(t)^{\exp(\beta z)}$$

One main advantage of the PH model is that the β coefficients have simple interpretations as log relative risks (rate ratios) which are familiar to epidemiologists and other medical scientists. A second advantage is that the likelihood function used to make statistical inferences about β takes the rather simple form

$$L(\beta) = \prod_{t_k} \frac{\exp(\beta z_k)}{\sum_{j \in R(t_k)} \exp(\beta z_j)} ,$$

where z_k denotes the covariate vector for the individual who dies at t_k, while $R(t_k)$ denotes the "risk set" of subjects who are still alive and under observation just prior to t_k. This may be derived alternately as a marginal likelihood based on the generalized rank vector of the censored survival times (Kalbfleisch and Prentice, 1973) or, using a chain of conditional probabilities reminiscent of actuarial techniques, as a partial likelihood (Cox, 1975). From this latter viewpoint, the contribution to the likelihood at $t = t_k$ is simply the conditional probability, given that some one of the individuals in $R(t_k)$ dies at that time, that it is the particular one whose death was actually observed.

Maximum likelihood estimates, tests of significance, and confidence intervals for individual β coefficients are all obtained by applying standard large sample likelihood methods to $L(\beta)$. A non-parametric estimate of the underlying survival distribution in the PH model is

$$\hat{S}_0(t) = \prod_{t_k \le t} 1 - \frac{d_k}{\sum_{j \in R(t_j)} \exp(\hat{\beta} z_j)}$$

(Breslow, 1975), which reduces to the PL estimate when $\hat{\beta} = 0$. Tsiatis (1978) has shown that $\hat{S}_0$ enjoys similar weak convergence properties to those noted earlier.

For purposes of illustration we fitted the PH regression model to data on 268 leukemia children enrolled in a clinical trial to investigate modifications of the standard maintenance chemotherapy in use at the time (Miller et al., 1974). The endpoint in question was the duration of remission, i.e. time until relapse, from completion of an initial induction course of therapy. At the time the data were analyzed, 181 children had relapsed while 87 remained in remission and thus had censored observations.

The regression variables of interest included: z_1 the initial (diagnostic) white blood count, expressed in log units; z_2 the child's age at diagnosis; z_3 age squared; and z_4 a dichotomous treatment variable which indicated whether ($z_4 = 1$) or not ($z_4 = 0$) the drug actinomycin-D had been added to the standard regimen. Table 2 shows the results of fitting the model in a hierarchical fashion, wherein each of the four variables are added to the regression equation in succession. The three test statistics computed at each fit evaluate the statistical significance of the contribution of the new variable, after accounting for the effects of the previous ones. These are the likelihood ratio statistic, the score test (based on the first derivative of the log likelihood), and the squared ratio of the estimated β divided by its standard error, all of which are known to be approximately equal in large samples under the null hypothesis (Rao, 1966, §6e).

The tests indeed yield rather similar numerical results with these data. The initial white count is the most dramatic determinant of the time of subsequent relapse. Age has a lesser but nevertheless important quadratic effect, such that the very youngest and oldest children are at highest risk, while those in the mid age range (2-10 years) have a somewhat better prognosis. After accounting for the effects of these baseline factors, the treatment variable is not statistically significant, although there is a trend towards a lower relapse rate for those receiving the new agent. Quantitatively, the regression coefficients indicate that there is an approximate doubling ($\exp(.721) \approx 2.06$) in the relapse rate for each 10 fold increase in the initial white count. The overall effect of treatment is to reduce the relapse rate by a factor of $\exp(-0.220) = 0.80$.

As a means of evaluating the goodness-of-fit of the model, the 268 patients were divided into 4 groups according to the value of their βz prognostic scores and PL estimates of the remission duration curve were computed separately for each one. In addition, fitted curves from the model were calculated for fixed βz values corresponding to the boundaries of the 4 groups. The results shown in Figure 1 illustrate some rather serious problems with the fit. The observed curves are more widely separated during the early time period than are the fitted curves, whereas the reverse is true in the later period. Such behavior indicates that the model is underestimating the effects of the covariates on the hazard function for small t, and overestimating them for larger t. In other words, the multiplicative effects depend on time.

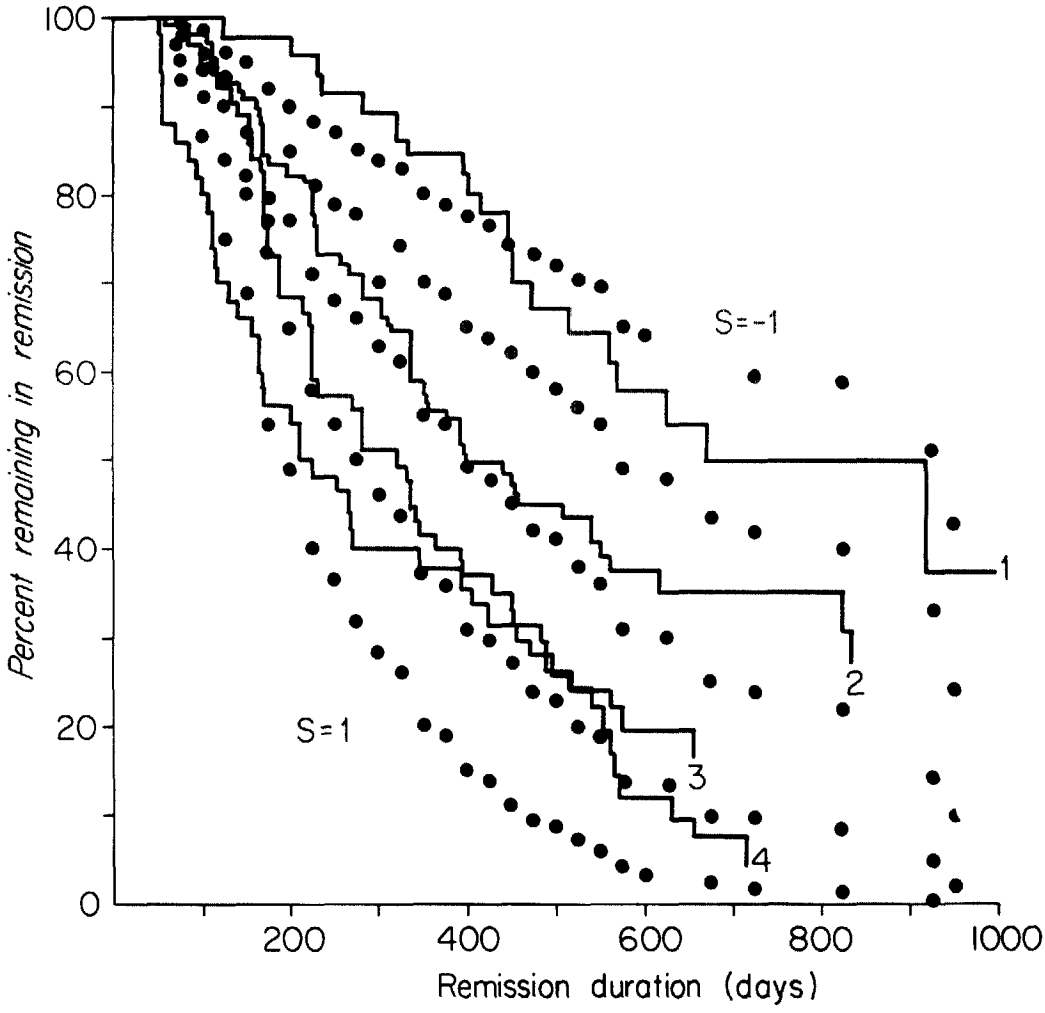
As an informal means of confirming this interpretation of the lack of fit, we re-analyzed the data by dividing the time period into two intervals with a cut-point at 270 days and making separate estimates of regression coefficients for each one. Operationally this is accomplished by factoring the partial likelihood into two pieces, one for relapses which occur prior to 270 days and another for those which occur later. The results in Table 3 show that the coefficients for the second period are indeed smaller in absolute value than those for the first. In

TABLE 2

FITTING OF PH MODEL TO DATA
ON 268 CHILDREN WITH ACUTE LEUKAEMIA
(ENTIRE OBSERVATION PERIOD)

STEP	VARIABLE ADDED	TESTS OF SIGNIFICANCE (χ^2_1)			REGRESSION COEFFICIENTS (Standardized Coefficients in Parentheses)			
		LR (N-P)	SCORE (Rao)	$\hat{\beta}^2/\widehat{Var}(\hat{\beta})$ (Wald)	LOG(WBC)	AGE/10	$AGE^2/100$	RX (0/1)
1	LOG(WBC)	38.61	41.56	40.55	0.783 (6.37)			
2	AGE	0.58	0.57	0.57	0.785 (6.41)	0.166 (0.75)		
3	AGE^2	6.62	6.97	6.87	0.737 (5.88)	-1.859 (-2.29)	1.454 (2.62)	
4	RX	2.12	2.12	2.11	0.721 (5.71)	-1.889 (-2.32)	1.483 (2.65)	-0.220 (-1.45)

FIGURE 1. Remission duration curves for four groups of leukemia children according to their prognostic score, and predicted curves on the basis of the PH model. (Reprinted, with permission, from Environmental Health Perspectives 32:181-189, 1979.)



Prognostic score: $S = 0.72 \log(WBC) - 0.19\,age + 0.015\,age^2 - 0.22\,AMD$

● Predicted remission duration curves from model for $S = -1, -\frac{1}{2}, 0, \frac{1}{2}$, and 1

– PL estimates of remission duration for four groups of patients defined by:

1	$S < -\frac{1}{2}$	(N = 47)
2	$-\frac{1}{2} \leq S < 0$	(N = 108)
3	$0 \leq S < \frac{1}{2}$	(N = 63)
4	$\frac{1}{2} \leq S$	(N = 50)

TABLE 3

COMPARISON OF REGRESSION COEFFICIENTS
FROM EARLY AND LATE OBSERVATION PERIODS
ON 268 CHILDREN WITH ACUTE LEUKAEMIA

PERIOD	LOG LIKELIHOOD	REGRESSION COEFFICIENTS ± S.E.			
		LOG(WBC)	AGE/10	$AGE^2/100$	RX (0/1)
EARLY	-480.729	0.866±0.172	-2.353±1.076	1.863±0.746	-0.468±0.210
LATE	-385.257	0.513±0.190	-1.206±1.294	0.786±0.907	0.037±0.229
ENTIRE	-876.863	0.721±0.126	-1.889±0.815	1.483±0.560	-0.220±0.151

Test for early vs. late: $\chi^2_4 = 2\{876.863-480.729-385.257\} = 21.75$ ($p<0.0002$)

particular, there appears to be a statistically significant effect of treatment on the early relapses such that patients receiving the new agent relapse at only exp(-0.468) = 63% of the rate of those who did not, whereas the treatment effect on the relapse rate during the later period is virtually nil.

While this sort of informal reasoning suffices to demonstrate the lack of fit of the PH model in this instance, it is an inadequate basis for extending the model.

One would not expect an abrupt change in the effects of covariates at 270 days, but rather a smooth transition over time. Such changes may be incorporated through the introduction into the model of time dependent covariates, some of which may represent interactions of baseline factors with time itself. For example, in the present situation, we define

$$z_5(t) = z_1 \times t$$
$$z_6(t) = z_4 \times t$$

to represent the attenuation of the multiplicative effects of white blood count and treatment on $\lambda_0(t)$. The model itself becomes

$$\lambda(t; z(t)) = \exp\{\beta z(t)\}\lambda_0(t),$$

from which it is clear that the hazard ratios for two different individuals need no longer be proportional. Formal tests of significance of the PH model are obtained by examining the coefficients of the time dependent covariates for evidence of departure from zero. The partial likelihood function used for inferences about the coefficients is extended to

$$L(\beta) = \prod_K \frac{\exp\{\beta z_K(t_K)\}}{\sum_{j \in R(t_K)} \exp\{\beta z_j(t_K)\}},$$

wherein the contribution to the likelihood at the time $t = t_K$ of the K^{th} death is based on the evaluation of the covariates of all individuals in the risk set at that time (Kalbfleisch and Prentice, 1980).

While time dependent covariates have been used here primarily to test the goodness-of-fit of the PH model, in other contexts, their employment contributes to the resolution of some major methodological problems. In occupational health studies, for example, they may be used to represent cumulative exposures which each worker receives during the follow-up period, or the duration of time since cessation of exposure, either one of which may influence the subsequent risk of disease development. Similarly, in clinical trials, they may be used to represent the effects of prognostic variables whose values are changing during the post treatment period. While particular care must be exercised in interpretation of the coefficients of such covariates, their use nevertheless significantly extends the range of possible analyses.

The above examples illustrate that the PH model is a flexible tool for the multivariate regression analysis of censored survival data, and it is not hard to see why is has been adopted wholeheartedly by medical statisticians. New areas of application are being discovered all the time. In conclusion, we mention briefly its application to case control studies.

When conducting etiologic studies of rare chronic diseases such as cancer, it is not possible logistically to keep under constant surveillance the large samples which would be required for diagnosis of a reasonable number of cases. An alternative is to collect the cases as they appear in hospitals and outpatient clinics, and compare their exposure histories with those of a control sample drawn from the population at risk. With matched studies, one or more disease-free controls of the same age and sex are sampled each time a case is diagnosed. An extension of the PH model for this situation leads to the same partial likelihood for the relative risk parameters B as given above, except that the risk sets for each case consist only of the case and its matched controls rather than the entire disease-free population (Prentice and Breslow, 1978). This approach provides a method for the multivariate analysis of match case-control studies which is closely related to many procedures already in use by epidemiologists (Breslow et al, 1978; Breslow and Day, 1980).

REFERENCES

Altschuler, B. (1970) Theory for the measurement of competing risks in animal experiments. Math. Biosc. 6, 1-11

Anderson, T.W. (1960) A modification of the sequential probability ratio test to reduce the sample size. Ann. Math. Statist. 31:165-197.

Böhmer, P.E. (1912) Theorie de unabhängigen Wahrscheinlichkeiten. Rapports, Memoires et Proces-verbaux de Septieme Congres International d'Actuaires,

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. Ann. Statist. 2:437-453.

Breslow, N. (1975) Analysis of survival data under the proportional hazards model. Int. Statist. Rev. 43:43-57.

Breslow, N. E. and Day, N.E. (1980) Statistical Methods in Cancer Research I: The Analysis of Case-Control Studies. Lyon: International Agency for Research on Cancer (WHO).

Breslow, N.E., Day, N.E., Halvorsen, K.T., Prentice, R. L., and Sabai, C. (1978) Estimation of multiple relative risk functions in matched case-control studies. Am. J. Epidemiol. 108:299-307.

Cox, D.R. (1972) Regression models and life tables (with discussion). J. R. Statist Soc. B 34:187-202.

Cox, D.R. (1975) Partial likelihood. Biometrika 62:269-276.

Cutler, S. and Ederer, F. (1958) Maximum utilization of the life table in analyzing survival. J. Chronic Disease 8:699-712.

Efron, B. (1967) The two sample problem with censored data. Proc Fifth Berkeley Symp. Math. Statist. Prob. 4:831-853.

Greenwood, M. (1926) The natural duration of cancer. Reports on Public Health and Medical Subjects 33 London: H. M. Stationery Office, pp. 1-26.

Hall, W.J. and Wellner, J.A. (1980) Confidence bands for a survival curve from censored data. Biometrika 67:133-143.

Kalbfleisch, J.D. and Prentice, R.L. (1973) Marginal Likelihoods based on Cox's regression and life model. Biometrika 60: 267-79.

Kalbfleisch, J.D. and Prentice, R. L. (1980) The Statistical Analysis of Failure Time Data. New York: Wiley.

Kaplan, E. L. and Meier, P. (1958) Non-parametric estimation from incomplete observation. J. Am. Statist. Assoc. 53:457-481.

Meier, P. (1975) Estimation of a distribution function from incomplete observations. In. Gani J. Ed., Perspectives in Probability and Statistics, London: Academic Press. pp. 67-87.

Miller, D., Sonley, M., Karon, M., Breslow, N. and Hammond, D. (1974) Additive therapy in the maintenance of remission in acute lymphoblastic leukemia of childhood: The effect of the initial leukocyte count. Cancer 34:508-517.

Prentice, R. L. and Breslow, N.E. (1978) Retrospective studies and failure time models. Biometrika 65:153-158.

Rao, C. R. (1966) Linear Statistical Inference and Its Applications. New York: Wiley.

Tsiatis, A. (1978) A large sample study of the estimate for the integrated hazard function in Cox's regression model for survival data. Technical Report 562. Department of Statistics, University of Wisconsin, Madison.

DR. KOCH: The second presentation will be given by Professor Regina Elandt-Johnson and Professor Norman Johnson who are both with the University of North Carolina. It is entitled "Statistical Aspects of Mortality Analysis: Actuarial and Biometric Points of View," and it will be concerned with trying to tie together common subject matter in biometric research and actuarial science.

Professor Elandt-Johnson had her graduate training at the University of Poznan. She has a strong interest in probability models in biology and survivorship. She is the author of a textbook entitled Probability Models and Statistical Methods in Genetics. She has also recently co-authored a new book with Norman Johnson entitled Survival Model and Data Analysis.

Professor Norman Johnson is currently the Alumni Distinguished Professor of Statistics at the University of North Carolina where he has been a professor since 1962. His graduate training was at the University of London where he also served on the faculty for a number of years. He is the author of many papers and books and recently he has taken on a new enterprise - the preparation of an Encyclopedia of Statistical Sciences. He also has interests as an actuary and is a Fellow of the Institute of Actuaries and a member of ASTIN in addition to being a Fellow of the Royal Statistical Society and of the American Statistical Association.

DR. ELANDT-JOHNSON AND DR. JOHNSON: As Seal (1977) has pointed out, actuaries had been constructing and studying life tables - in particular, multiple decrement life tables - for some 100 years before the recent (25 years or so) growth of interest in this topic among statisticians. This growth was fostered, in large part, by the very substantial increase in the number of statisticians employed in what may be broadly called "health industries", who were required to analyze the results of controlled clinical trials, and data from longitudinal epidemiological

studies. Using mathematical ideas of probability theory, statisticians have developed their own concepts and techniques for analysis of survival data, sometimes paying very little or no attention to the techniques already existing in actuarial science. There is consequently a considerable difference in notation and concepts.

In recent years, relations between the two approaches have become, in general, more clearly understood; our discussion is intended to be a modest contribution to the continuation of this process.

We use the term "biometric" rather than "probabilistic" or "statistical" to describe the second point of view in our title in recognition of the fact that a large part of recent developments have arisen from work in biological (including medical) applications. The basic probabilistic ideas, however, are derived from reliability theory whose main field of applications is quality control in industry.

A basic tool of the actuary is the life table; the corresponding statistical concept is the survival distribution function (SDF). Denoting age (or survival time) by x, the relations between the two are summarized as follows:

LIFE TABLE	SURVIVAL ANALYSIS
Expected number surviving at age x	
ℓ_x	---
Expected proportion surviving at age x (survival function)	
ℓ_x/ℓ_0	$\Pr\{X > x\} = S_X(x)$
Conditional probability of death between age x and x + 1 given alive at age x	
$q_x = 1 - \frac{\ell_{x+1}}{\ell_x}$	$q_x = \Pr\{x < X \leq x+1 \mid X > x\} = 1 - \frac{S_X(x+1)}{S_X(x)}$
Force of mortality (hazard rate)	
$\mu_x = -\frac{1}{\ell_x} \cdot \frac{d\ell_x}{dx}$	$\mu(x) = -\frac{1}{S_X(x)} \cdot \frac{dS_X(x)}{dx}$

It is easy to see that

$$S_X(x) = (1 - q_0)(1 - q_1) \ldots (1 - q_{x-1}), \text{ and} \qquad (1)$$

$$\ell_x = \ell_0 \cdot S_X(x) \qquad (2)$$

The further life table functions (d_x, l_x, $\mathring{e}_x$, etc.) can be derived using probability theory.

While actuaries certainly used the concept of probability (though not always spelled out), it seems that they primarily thought of the life table as representing the progress of a cohort of l_0 newborn individuals through life, rather than representing the distribution of lifetimes.

Actuaries have from time to time been concerned about formulating mathematical "laws of mortality"; the corresponding statistical problem is that of finding an appropriate parametric distribution function. Statisticians tend to give greater importance to this aspect than actuaries, whose life tables were often constructed by estimating values of q_x, and then combining them by means of formula (1), without ensuring a parametric lifetime distribution. There are of course, not unusual exceptions, from both sides. Gompertz or Makeham-Gompertz laws were used in construction of life tables - the British H^M table is an early example - while the Kaplan-Meier (1958) formula, which results from a "distribution-free" approach, has been in common use among statisticians for over two decades.

There are at least two topics related to life tables in which differences between the two approaches can clearly be seen: (i) multiple decrement life tables and competing risk theory, and (ii) select life tables and use of concomitant variables in survival models.

Suppose that causes of death are classified into k disjoint classes, C_1, C_2,..., C_k; we will, for convenience, refer to "cause C_1", "cause C_2", etc. From mortality records (death certificates) we can identify the cause of death and time (age) at death.

A good deal of recent theoretical development in the theory of competing risks has been based on a concept of potential times "due to die", $X_1,\ldots,X_k$, one for each cause. In fact, this concept is adapted from reliability theory, where it is customary to speak about failures of different components, each component having its own time "due to fail". It is, then, convenient to introduce their joint survival function

$$Pr\{X_1 > x_1, X_2 > x_2, \ldots, X_k > x_k\} = S_{1\ldots k}(x_1, x_2, \ldots, x_k). \tag{3}$$

Under the assumption (A1) that death is ascribed to just one cause, the observed time at death is, in fact, X=min $(X_1,\ldots,X_k)$, and the overall survival function is

$$S_X(x) = S_{1\ldots k}(x, x, \ldots, x). \tag{4}$$

We now have to distinguish forces of mortality, "survival functions", etc., for cause C_α in the presence and in the absence of all other causes. We add the prefix "a" to the corresponding function when we refer to presence, and no prefix for absence of all other causes. For

example, the force of mortality for cause C_α acting in the presence of all other causes is denoted by $a\mu_\alpha(x)$; and when C_α is acting alone it is denoted by $\mu_\alpha(x)$. (In British notation, $a\mu_\alpha(x)$ is denoted by $(a\mu)^\alpha_x$, and $\mu_\alpha(x)$ by μ^α_x while the corresponding U.S. notation is $\mu^{(\alpha)}_x$ and $\mu'^{(\alpha)}_x$, respectively. Similarly, $a\ell_{\alpha x}$, $\ell_{\alpha x}$ coincide with British $(a\ell)^\alpha_x$, ℓ^α_x and U.S. $\ell^{(\alpha)}_x$, $\ell'^{(\alpha)}_x$, respectively.) Note that

$$a\mu_\alpha(x) = -\frac{1}{S_{1,\ldots,K}(x,\ldots,x)} \cdot \frac{dS_{1,\ldots,K}(x_1,\ldots,x_K)}{dx_\alpha}\Bigg|_{\{x_r = x \text{ for } r=1,\ldots,K\}} \tag{5}$$

and in view of assumption A1,

$$\mu(x) = a\mu_1(x) + a\mu_2(x) + \cdots + a\mu_K(x). \tag{6}$$

It is easy to see (c.f., also Gail (1975)) that

$$S_X(x) = S_{1,\ldots,K}(x,x,\ldots,x) = G_1(x) \cdot G_2(x) \cdots G_K(x), \tag{7}$$

where

$$G_\alpha(x) = \exp\left[-\int_0^x a\mu_\alpha(t)dt\right], \quad \alpha = 1, 2, \ldots, K. \tag{8}$$

The latter can be interpreted as a "survival function" for cause C_α alone if the force of mortality in the absence of all other causes was equal to the force of mortality in their presence, that is, (assumption A2),

$$\mu_\alpha(x) = a\mu_\alpha(x), \quad \alpha = 1, 2, \ldots, K. \tag{9}$$

When using the concept of joint SDF, (9) is equivalent to the assumption that $X_1, X_2, \ldots, X_K$ are mutually independent. Unfortunately, this independence cannot be tested. If this assumption, or some similar assumption, is not justified (e.g., on biological grounds), we face the problem of nonidentifiability: the joint survival function cannot be identified from mortality data alone. There exists an infinity of possible joint distributions of $X_1, \ldots, X_K$, which produce the same observable survival distribution $S_X(x)$ and identified cause of death. [e.g., Cox (1962), Tsiatis (1975)].

The difficulty in regard to both the nonidentifiability of $S_{1\ldots k}(x_1, \ldots, x_K)$ and the actuarial assumption $\mu_\alpha(x) = a\mu_\alpha(x)$ can, to some extent, be resolved.

Let

$$aP_\alpha(x) = Pr\{X > x, X = X_\alpha\} = \int_x^\infty a\mu_\alpha(t) S_X(t) dt \tag{10}$$

denote the probability function that the person aged x will eventually die from cause C_α in the presence of all other causes. Then

$$a\mu_\alpha(x) = -\frac{1}{S_X(x)} \cdot \frac{d(aP_\alpha(x))}{dx} \quad . \tag{11}$$

It is easy to see that

$$aP_1(x) + aP_2(x) + \ldots + aP_k(x) = S_X(x) \quad . \tag{12}$$

The definition of the $aP_\alpha(x)$'s (and then of $S_X(x)$ and $G_\alpha(x)$'s does not require introduction of "times due to die" and their joint SDF, so that assumption A2 no longer implies independence.

The approach just discussed is, in fact, an actuarial approach to construction of multiple decrement life tables. The $S_X(x)$ corresponds to al_x/al_0, the $aP_\alpha(x)$ corresponds to $al_{\alpha x}/al_0$, while $G_\alpha(x)$ corresponds to $l_{\alpha x}/l_0$. Formal mathematical bases for deriving $S_X(x)$, using the probability functions $aP_\alpha(x)$'s can be found in Cox (1962) and Berman (1963). The meaning of $G_\alpha(x)$ (and $l_{\alpha x}$) remains, however, obscure.

Why must it be interpreted as the SDF for cause C_α acting alone, when other causes have been eliminated? Chronic diseases, which are now major causes of death, are mutually associated in one way or another; only greater knowledge of the history of diseases could help (though still speculatively) in regard to the effect of "elimination" of a cause. In the present stage of knowledge, it would be helpful to regard $G_\alpha(x)$ as a (hypothetical) "waiting time" distribution for cause C_α with the force of mortality $a\mu_\alpha(x)$. [Elandt-Johnson & Johnson (1980), Chapters 9-12]. No strong physical interpretation has to be associated with $G_\alpha(x)$ (or $l_{\alpha x}$), but it can be useful as a tool in estimation problems, (using mortality data from cause C_α only, while deaths from other causes are treated as withdrawals). In our opinion, the tables on effect of cause "elimination" in the National Life Tables by Causes of Death are misleading and should be omitted, or interpreted in another way.

The results discussed in this section are summarized as follows:

LIFE TABLES BY CAUSES OF DEATH	COMPETING RISK ANALYSIS
Expected number surviving at age x	
al_x	—
Expected proportion surviving at age x	
al_x/al_0	$Pr\{X > x\} = S_X(x)$
Expected proportion of individuals who die after age x from cause C_α in the presence of all other causes	
$al_{\alpha x}/al_0$	$Pr\{X > x \text{ and } X = X_\alpha\} = aP_\alpha(x)$
Force of mortality for cause C_α in the presence of all other causes	
$a\mu_{\alpha x} = -\frac{1}{al_x} \cdot \frac{d(al_{\alpha x})}{dx}$	$a\mu_\alpha(x) = -\frac{1}{S_X(x)} \cdot \frac{d(aP_\alpha(x))}{dx}$

"Waiting time" distribution for cause C_α alone

$$l_{\alpha x}/l_{\alpha 0} \qquad G_\alpha(x) = exp\left[-\int_0^x a\mu_\alpha(t)dt\right]$$

In most controlled clinical trials or longitudinal studies, measurements of many variables (blood pressure, cigarette smoking, etc.) are routinely obtained. Statisticians have attempted to make use of this information by incorporating concomitant variables in the parametric models they construct. This is usually done by representing the hazard rate (force of mortality) in a parametric (or semi-parametric, that is, without specifying the hazard rate completely) form. For example, Cox (1972) suggested a model of hazard rate at time t in the form

$$\mu(t;z) = \mu(t)\exp\left(\sum_{u=1}^{5}\beta_u z_u\right), \tag{13}$$

where $z = (z_1, z_2, \ldots, z_5)$ are concomitant variables, and $\mu(t)$ is an arbitrary underlying hazard. Several other models for $\mu(t;z)$ can be found in the literature (see Elandt-Johnson & Johnson (1980), Chapter 13).

It is perhaps, fair to say that, so far, most of this modeling has been done on an arbitrary basis rather than by application of any specific scientific principles or theory.

Models incorporating effects of concomitant variables can be regarded as rather elaborate forms of select and/or sectional life tables. If the only variables are "age at entry" and "time since entry," then there is, of course, an exact parallel with select tables, though the selection is often negative.

Models with concomitant variables can be estimated from a relatively small volume of observations, but construction of select life tables requires a large body of data - effectively to allow for stratification. Therefore, the results of survival analysis with concomitant information are rarely presented in the form of select life tables. There is, however, an interesting exception in a report by Tallis et al. (1973) on breast cancer among women in Melbourne, Australia. (Here "entry" is registration as having breast cancer.) Not only are select life tables presented, but there are separate tables on each of 12 combinations: (4 stages of disease) x (3 types of tumor malignancy).

Comparison of mortality at various times after operation with that among general population at the same attained age (aggregate mortality), though not in the form of select life tables, is sometimes presented e.g., Cutler and Axtell (1963) .

Adequate presentation of the information continued in fitted models which incorporate several concomitant variables in life table form would often call for so many select life tables as to constitute a small library. A great advantage of mathematical formulation of hazard rate is the condensed presentation and by using computer programs, the flexibility to

incorporate many covariables. A disadvantage is that the assumptions, on which models are based, are often not well justified and consequently, the SDF's corresponding to certain combinations of concomitant variables may be quite unrealistic.

A few related comments should be made at this point. In the estimation of central death rate, m_x, (and so in estimation of q_x) the concept of "person-time units exposed to risk" plays an essential role. This concept appears to be a casualty of the statistical approach. Statisticians tend to prefer to approximate the survival distribution (and sometimes also the distribution of "time due for withdrawal"), in a given age group, by uniform or exponential distribution. Using further assumptions - for example, that, on the average, the withdrawal time is in the middle of the age group interval - the statisticians construct a likelihood function and come up with a so-called "maximum likelihood estimator" of q_x. This approach is called, by some statisticians, "scientific" while the actuarial approach, based on the concept of person-years, is named "intuitive".

First, we notice that the actuarial estimator of q_x,

$$q_x \doteq \frac{m_x}{1+\frac{1}{2}m_x}, \tag{14}$$

is, indeed, a "maximum likelihood" estimator, when the exponential approximation to the survival distribution is used, though it was originally derived without reference to this fact. More details on comparisons of various estimators of q_x can be found in Elandt-Johnson and Johnson (1980), Chapter 6.

Second, the concept of "person-years of exposed to risk" beyond its value as a technical tool in estimating death rates, provides an easily understood index to the amount of data available in various subpopulations (e.g., age groups, sections, etc.) of the experience.

We also note the unfortunate fact that in the analysis of clinical trials data, "life table" is used to mean what, for many years, has been termed "experience". What has been traditionally called a "life table," seems to be called in analysis of survival data from clinical trials, a "theoretical life table."

Even more confusion is caused by using the terms "rate" and "probability" interchangeably. Clearly, q_x is a probability, but in actuarial terminology it is termed "mortality rate" while the central rate, m_x, is termed a "death rate". It is important to recognize the difference between rate and proportion to avoid confusion Elandt-Johnson, (1975) .

REFERENCES

Berman, S.M. (1963). Note on extreme values, competing risk and semi-Markov process. Ann. Math. Statist. 34, 1104-1106.

Cox, D.R. (1959). The analysis of exponentially distributed lifetimes with two types of failure. J. Roy. Statist. Soc. Ser. B, 21, 411-421.

Cox, D.R. (1962). Renewal Theory. Methuen, London. (Chapter 10).

Cox, D.R. (1972). Regression models and life tables. J. Roy. Statist. Soc. Ser. B. 33, 187-202.

Cutler, S.J. and Axtell, L.M. (1963). Partitioning of a patient population with respect to different mortality risks. J. Amer. Statist. Assoc. 58, 701-712.

Elandt-Johnson, R.C. (1975). Definition of rates: Some remarks on their use and misuse. Amer. J. Epid. 102, 267-271.

Elandt-Johnson, R.C. and Johnson, N.L. (1980). Survival Models and Data Analysis. Wiley, New York.

Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53, 457-481.

Seal, H.L. (1977). Studies in history of probability and statistics. XXXV. Multiple decrements and competing risks. Biometrika 64 420-439.

Tallis, G.M., Sarfaty, G., and Leppard, P. (1973). The use of probability model for the construction of age specific life tables for women with breast cancer. Endocrine Research Unit, Cancer Institute, Melbourne, Australia.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. Proc. Nat. Acad. Sci. U.S.A 72, 20-22.

DR. KOCH: I am sure that everyone enjoyed the dialogue and the information on the different perspectives that are in existence with respect to biomedical research and actuarial science as well as the concern over language. The problem of specialized language not only exists in terms of biometricians and actuaries but statisticians and epidemiologists often tend to use somewhat different language for the same thing. It would seem that each person who has some kind of interest in the use of probabilistic or statistical methods in a particular study has a tendency to develop his or her own language and set of methods almost independent of what exists in other areas. We are trying to bridge the gap so that this tends not to happen at least to the extent that it has in the past.

The third presentation will be given by Professor Ralph Buncher who is the Director of the Division of Biostatistics and Epidemiology at the University of Cincinnati. His presentation, entitled "Is the Healthy Worker Effect Due to Select Mortality?," is concerned with the particular type of phenomena seen in epidemiological research, an area which has an actuarial flavor as well as a biostatistical flavor. Dr. Buncher received his graduate training at Harvard University. Since that time he has served at Merrell-National Laboratories as well as the University of Cincinnati as I mentioned previously. He has many professional activities including efforts in the area of occupational health as well as biopharmaceutical research. In the latter context he is co-editor of a soon to be published research monograph entitled, Statistics in the Pharmaceutical Industry. He is also the current chairman of the American Statistical Association Biopharmaceutical Section.

DR. C. RALPH BUNCHER: This paper is being presented because I believe this topic can best be resolved by both actuaries and those who work in biostatistics and epidemiology joining forces and pooling their knowledge. The topic of interest is the so called Healthy Worker Effect - said to be a variety of select mortality. My ultimate goal is to prove that in some circumstances working is healthy. In trying to arrive at that conclusion, one must first resolve a number of issues of bias and analysis which complicate the understanding of the problem. An abbreviated survey of the literature in this field has demonstrated to me that there are many factors that must be considered. We shall discuss these issues in this paper.

In this discussion I shall first describe the studies of occupational mortality as they are usually done, then I shall define the Healthy Worker Effect and give the standard epidemiological explanations, and finally discuss some of the possible alternative explanations and confounding factors.

Let us discuss one variety of occupational mortality study. This type of study goes under many different names such as a historical prospective study or a cohort mortality study. A group of workers who become the cohort to be followed are defined based on a common work environment. Usually this definition involves specific work exposures or potential exposures, certain calendar years of interest, a minimum time on the job, and occasionally, demographic limitations such as males only. The health factors of interest are usually death from chronic disease, especially cancer. Thus those individuals newly hired, or who have only been exposed for a few years in periods of time too short for a chronic disease to appear, are of much less interest than those who have been on the job for many years and for whom many years have passed since their first exposure to that work environment.

Thus, a typical study might have the following characteristics. All employees who worked for the ABC Chemical Company during the interval from 1 January 1948 through 31 December 1974 are eligible for the study cohort. Each person must have worked at least one year in the

manufacturing area during that period of time. Thus, persons such as guards, truck drivers, secretaries in the office, and gardeners would be excluded. White collar employees and other management personnel would be included provided they either worked in the factory along side of the hourly employees and thus were also exposed to the chemical or if they had previously been hourly employees working in the factory prior to their promotion to a management position. The minimum exposure of one year allows sufficient time on the job so that one could anticipate a biologic effect being likely and at the same time eliminates from consideration those employees who worked only for short periods at this plant. Such short term employees are more likely to have worked for only short periods of time at several other locations making exposure status very difficult to measure and making follow-up more difficult.

When one is measuring mortality, one notes that the persons in the cohort are only at risk of dying after they have worked at least one year during the time period under consideration; otherwise, these individuals would not be in the study had they died prior to completing one year on the job. In terms of exposure, one measures all of the time the employee spent on the job whether during the calendar years in the cohort definition or during those years prior to the ones under study since prior years also involve exposure to the chemical. These prior years are involved with calculations of the risk and estimating minimum number of years on the job or after initial exposure before one sees the effect, if any.

We are here discussing only studies that yield death rates for each cause of death. The denominator consists of person-years at risk of dying. Some studies use the deaths as the base (denominator) and look at the proportion of deaths due to each cause; these proportionate mortality studies are not under discussion.

Each worker in the cohort is then followed to some fixed date to ascertain whether the individual is still alive or has died. Thus, all persons in the study might be followed through 31 December 1978 and their vital status as of that day is determined. Usually, most of the workers are still on the job or are retirees whose whereabouts are known to the company because of benefit plans such as an annuity contract. Those individuals who have left employment must be found through various follow-up mechanisms to ascertain their status. This can be done through contacts in the workplace, telephone directories, drivers license records, city directories, and by ascertaining status from the U.S. Social Security Administration. If none of these methods work which is usually true for some 10% or less of the cohort, then special searches must be made. Given that many of these studies involve several decades of old work records, a 90% follow-up rate is usually considered adequate although 95% and above is considered good. The crucial problem is the question of whether those who have not been followed have a vital status and distribution of causes of death similar to the rest of the cohort.

One then creates a life table analysis or more commonly a comparison of the observed death rates with those to be expected in some standard population group. The usual standard is the vital statistics population of the United States which provides adequate numbers and is easily available in publications of the federal government. It is also possible to use the vital statistics of the state in which the plant is located as a comparison group but this introduces smaller numbers and raises the issue of whether the death actually occurred in that state or whether the work force was sufficiently mobile that persons, especially after retirement, moved to other states and died there thus calling into question the standard of using the vital statistics from the state in which the plant is located. Obviously, it would be very useful to have a set of vital statistics for employed persons or employed persons plus retirees as a standard rather than include all of the vital statistics population. Unfortunately, this concept is both difficult on a theoretical basis (try to define exactly who should be in such a comparison group and get others to agree) and on a practical basis is not generally available to occupational epidemiologists in any form at this time.

When total mortality of the work cohort is compared to total mortality of the vital statistics population, the usual finding is that the number of observed deaths is less than the number to be expected based on the vital statistics population, not infrequently being 70 to 80 percent of the expected. The expected numbers are derived using individuals of the same sex and race in five year age groups and for either specific calendar years or perhaps five year calendar intervals. Deaths (the numerators) come from vital statistics while the population base (the denominators) comes from census data. Variations on this theme are possible. The point to be made here is that these other demographic variables are controlled for and hence are not responsible for differences between the observed and the expected number of deaths.

Since the observed number of deaths in total is typically less than the expected number of deaths in these occupational epidemiology studies, this finding has been called the Healthy Worker Effect. The title implies that those who worked in the factory were selected individuals from the vital statistics population such that only healthier individuals were in the work force.

Two explanations are usually given as the explanation of the Healthy Worker Effect. The first is that the medical preemployment examinations select out the unhealthy and the second is that there is self selection such that the unhealthy do not apply for employment. Each of these explanations contains obvious elements of truth whether they provide a complete explanation for the effect or not. Physical examinations can select out those with a bad heart, malfunctioning lungs, kidney or liver disease, and other signs of impending mortality. On the other hand, the ability of even modern medicine to predict mortality several years hence applies only to a small fragment of the population. Thus this type of

selection could only explain a small portion of the commonly observed 20-30% reduction in mortality found and called the Healthy Worker Effect. Moreover, the nature of this medical selection is such that the Healthy Worker Effect should only appear for about a decade after initial employment. Finally, many preemployment **physicals** have been sufficiently perfunctory that only the most gross abnormalities were likely to have been found and become cause for rejection. We note that current employment law, which rejects criteria predictive of mortality unless those criteria also indicate that a person cannot perform a job, may in future studies produce changes in this portion of the effect.

A second portion of the effect is the bias produced by self selection on the part of the individuals who might apply for the jobs. Thus workers with cystic fibrosis and other lung diseases are unlikely to apply for positions in dusty environments. Physically frail persons may not apply for heavy physical labor. Persons who react to various chemicals or have allergies or other departures from the usual may choose to avoid working in a factory type environment looking instead for a job in an air conditioned office building.

Is there an analog to the Healthy Worker Effect in other studies? One answer is to recall the well known life insurance "select period" found as a routine part of the preparation of mortality tables for persons with life insurance. This select period is explained because an applicant for life insurance must meet certain underwriting standards. Thus those insured on a standard basis have started in a pool of applicants from which were eliminated those who are insured on other than a standard basis and those who are considered uninsurable. Actuaries then find that the rate of mortality is lower for a group of lives just insured on a standard basis at age 30 than another group of lives age 30 similarly insured a year ago. In turn each of these has lower mortality than another group age 30 who were insured 2 years ago at age 28, and so forth.

Jordan (1967) tells us that a select period may last as long as ten years, and sometimes even up to 15 years. I have never seen a reference that would indicate that a select period would last longer than 20 years even given a special life insurance physical examination which most persons would consider far more rigorous than a preemployment physical examination.

The life insurance information seems to tell us that no matter how rigorously we select a group of individuals who represent one-third to two-thirds of the total group, after two decades the survivors of the select group will be the same as the survivors of the unselected group. This observation should not come as a great surprise to anyone since it is merely the factual representation of our inability to predict medical events far into the future and our observation that entropy increases so that the survivors of a highly selected group are eventually the same as those not selected. Thus we are left with the conclusion that selection is a self-limiting process.

Clearly then, if we are to interpret the Healthy Worker Effect as a selection-at-time-of-employment effect, our other experience suggests that this selectivity will disappear after a couple of decades. Any investigator who wishes to point out that a Healthy Worker Effect has been found in an occupational epidemiology study should show us that effect separately based on the first decade after selection and the second decade after selection. If there is a component of better health after the first two decades of work, then an alternative explanation must be sought.

Is it possible that work is good for you especially in the sense of a lowered mortality rate caused by working? Obviously there are many steps to be taken before a causal relationship could be concluded. In this section, let us consider the likely paths by which work might reduce mortality. Let us explore the ways in which work might serve as an "etiologic agent" for reducing mortality.

First, many occupations involve a constant low level of exercise. Few occupations still involve a great deal of physical exercise in the North American world of the 1980's. Many occupations require walking perhaps as much as several miles in a day and climbing stairs or their equivalent. The literature on exercise and heart disease suggests that constant exercise tends to reduce the cardiovascular death rate. There are a number of problems with this literature, and although the conclusion is generally accepted, there are epidemiologists who either do not believe in the causal relationship or believe that the risk reduction is minimal. Nevertheless, the report by Brand et al. (1979) on a cohort of longshoremen is relatively convincing.

External causes of death, especially accidents, provide another area to be explored. A great deal of effort is expended by employers and by governmental agencies such as the Occupational Safety and Health Administration on the education of workers in accident prevention. Workers are reminded of safe habits on the job which raises the question of whether that training has carryover to the nonjob hours such that those same workers are involved in fewer accidents in the home, in their automobiles, and in pursuit of their hobbies. Although most work groups have lower rates of death due to external causes, an interesting finding in some studies is that high income workers in their 20's may have more automobile accidents since these persons are better able to purchase the high performance automobiles which are statistically more likely to be involved in automobile accidents. The other external causes of death, suicide and homicide, are also frequently found to be less common in occupational cohorts than in vital statistics populations. An interesting possibility of association between a long term job and a reduction in both suicide and homicide is worthy of additional study.

A third possibility to consider involves conformity to the norm. In spite of occasional exceptions, one would have to believe that the norm of the workplace tends to be more healthy than the individual choices which might be exercised. For example, excessive drinking, excessive

quarreling, and other personality problems are activities outside the norm of a working place. The norm would dictate taking care of oneself such as taking antihypertensive medication. In these matters it is possible that the social group of the workplace helps the individual towards healthy behavior which lowers mortality.

Another important characteristic of the workplace is that it provides and requires a scheduled life for the individual. Some of the findings in various mortality studies, e.g., those who eat breakfast have lower death rates, can be interpreted to mean that those people who live a scheduled life tend to have better mortality than those who do not adhere to such biologic rhythms. The anecdotal accounts of workers who have retired and rapidly go downhill to death could be explained by too rapid changes in these biologic rhythms. An individual without a workplace imposed life schedule is faced with creating such a schedule.

Obviously many steps must be taken before these potential healthy effects of the workplace can be claimed. One of the requirements for proving causation involves demonstrating the time basis of a causal relationship so that we don't confuse it with an association. For example, one must be careful to differentiate between work causing better health and the inverse relationship that poor health causes lack of work.

These problems can be grouped into three categories that need verification. These three categories could be called selectivity, balanced risks, and calculation errors and biases. We shall consider each in turn.

A number of problems concerning selectivity must be resolved. For example, short term workers, those excluded from the usual cohort mortality study, may be different from those included in at least two different ways. First, it may be that high turnover of employees may be associated with a high death rate. Alcoholics, heavier drug users, those with strong personality conflicts, etc. may have both a high job turnover and a higher death rate. Alternatively it may be that those who are especially sensitive to the workplace environment are the ones who leave. Those who have an itch, a rash, an uncomfortable feeling, difficulty breathing, watery eyes, and so forth may leave the workplace early and be excluded from studies. The concept is of a person who is exposed only briefly but because of special sensitivity may have been exposed to a large enough personal dose of the chemical under study to show biologic effects. Unfortunately, those effects would be almost impossible to associate with the short duration cause because of the myriad of potentially conflicting causes for that individual. In these instances, the observed effect of the workplace would be a lower death rate because those who would raise that death rate have been excluded from the epidemiologic analysis, or even worse included in the comparison group.

Another alternative is that the high and low risks of death are balanced if we were to look at a census of mortality studies and the numbers of workers involved. It would be excellent to have a sort of input/output table showing numbers of workers in different occupations and the mortality rates of those occupations. This combined information in a form comparable to econometric models could tell us whether those special work groups such as asbestos workers who are at an increased death rate balance those who show the Healthy Worker Effect. Most epidemiologists would point out that the large quantity of workers whose workplace is an air conditioned office (including actuaries and biostatisticians) have been little studied because of the assumption that their mortality is at least as good as those occupational groups who have been found to exhibit the Healthy Worker Effect. The more general question is to find where is the excess of mortality that balances out the Healthy Worker Effect reported in various occupational groups.

Actuaries have done a lot of important work along these lines and published some of those results in the actuarial literature. I believe that some cooperative sessions involving both the actuarial community and the epidemiologic/biostatistical community could produce a great deal of fruitful information on this subject based on both published and unpublished sources. Risk levels by occupation are better documented than numbers of persons exposed to each risk level.

At least five different possible calculation errors and biases should also be considered in these studies. First, we note that in every study there is a small segment of persons who are lost to follow-up. These people tend to be "loners," that is those without very many friends or much family. The assumption must be made that these individuals are similar to those in the rest of the study. If the percentage of lost to follow-up is not very small and these people died from different causes of death and/or at a more rapid rate, then there is an important bias in the study data.

It is possible that some of the analysts have overcounted person years of exposure in the denominator of their death rates. For example, there is an "immortal period" as mentioned previously during which time a worker cannot die and be in the study according to the definition of the cohort. Clearly these person years should not be counted in the death rates since the worker is not at risk of dying during these years of qualifying for the study. While I have great faith in the epidemiologists doing these studies, the complexity of communication amongst those in charge of worker follow-up, those who record and abstract the work history, those doing computer programming, and the epidemiologist suggest to me that this is a possible source of underestimating death rates.

Problems with causes of death certification must also be understood. There are a limited number of special coders (nosologists) who do the work on occupational epidemiology studies and most of these persons are aware of the special interest in cancer in these studies. If the data are to be compared to vital statistics death coding, then it is important

that the special coders do not code differently than those who code vital statistics. This assumption could use much greater verification. It may even be that a few death coders sometimes note the occupation in choosing an underlying cause of death. Physicians filling out the forms might also be subject to this bias. The ultimate difficulty, for the physician filling out the death certificate and the coder in choosing between those who die from cancer compared to those who die with cancer but from another cause, is almost without objective solution. Clearly the items in this paragraph relate more to the distribution of the causes of death rather than the total death rate.

Changes in the death rates also produce some problems since vital statistics generally take several years before they are published. For example, we all know that death rates have fallen rapidly through the decade of the 70s. Within the total, there have been major declines in cardiovascular death rates even as the rates for respiratory cancer are rising rapidly. Thus if the 1975 vital statistics are used to provide expected deaths for the years 1976 through 1978, the actual death rate in those years will be less than the expected.

A final problem for discussion is the standard of comparison. Working groups which have been studied tend to come from industrial states which are more urban and have higher death rates than the rest of the United States. What then is the appropriate standard of comparison? Is it the total United States or the single state of interest? For example, I am now involved in a study in which the work group has a higher lung cancer death rate than the total United States but not higher than the local state rate.

These issues could be resolved by pooling resources of the actuarial community and the occupational epidemiology community. Many of these issues could be resolved almost immediately based on personal experience and reports in the literature; others would require the acquisition of a more extensive data base. If we are able to understand these secondary issues, then we can proceed with trying to understand better the Healthy Worker Effect and whether working causes lower mortality.

REFERENCES

Brand, Richard J., Paffenbarger, Ralph S., Jr., Sholtz, Robert I., and Kampert, James B. (1979). Work activity and fatal heart attack studied by multiple logistic risk analysis. Amer. J. Epid. 110: 52-62.

Jordan, Chester W., Jr. (1975). Society of Actuaries' Textbook on Life Contingencies. (2nd ed) The Society of Actuaries, Chicago.

DR. KOCH: We will now have a formal discussion by Tom Herzog. Tom will dicuss the three papers from the viewpoint of the actuary. He is an actuary as well as a mathematical statistician for FHA and HUD. His present duties include statistical surveys, sample design consultation and analysis using procedures such as log linear models and logistic and robust regression. He had previously worked with the Social Security Administration. Discussion from the floor will follow.

MR. THOMAS N. HERZOG: I would like to thank Bob Johansen and Gary Koch for organizing this session. I would also like to thank the speakers for their stimulating talks. Since the only paper I have had for more than a week was that of the Johnsons, my remarks will be brief and of a general nature.

With the advent of modern computers, the development and application of multivariate statistical procedures has increased rapidly during the last two or three decades. Such techniques as multidimensional contingency table analysis (which Gary Koch has been advocating for some time now) and regression models of the type described by Norman Breslow are quite powerful and useful techniques. Some of these multivariate procedures can be used to construct statistical models that indicate which predictor variables and interaction effects are important and which are not.

Multivariate statistical techniques have great potential for successful application in the actuarial field. One important area of potential application is risk-classification. For example, actuaries might use these techniques to determine what characteristics account for mortality differences among insureds. Do nonsmokers have a greater life expectancy than smokers because they do not smoke or is this difference explained by the observation that nonsmokers tend to be richer than smokers and that the rich tend to live longer? The actuary should have sufficient familiarity with multivariate statistical procedures so that, when the need to use them arises, he can at least request the appropriate outside help if he feels unable to carry out the analysis himself.

While these procedures are quite important, they are not always completely trivial to carry out. One issue which frequently arises here concerns sample size requirements. Are there enough observations to examine all of the predictor variables likely to be important? Another practical issue is how, or perhaps more appropriately, whether to write your own computer program to perform the desired analysis. Some of these procedures, such as the regression procedure discussed here by Norman Breslow, require the use of quite sophisticated numerical analysis techniques in order to (1) ensure that the procedure converges in the minimal amount of time and (2) that the convergence is to the "correct" solution. There are, of course, several widely used statistical packages which may contain the required software. Two such statistical packages are

(1) The Biomedical Computer Programs, Series P (known as BMDP) which is produced and distributed by the UCLA School of Medicine, and

(2) The Statistical Analysis System (known as SAS) which is brought to you by the folks at the SAS Institute, which was formerly associated with North Carolina State University.

Both of these packages contain logistic regression programs which are similar to the type of regression analysis suggested by Dr. Breslow.

Another package which is reputed to be very good is the IMSL package--standing for International Mathematics and Statistics Library.

Let me now move on to some additional specific remarks about the individual papers just presented.

As an addendum to the Johnsons' discussion of competing risks, I would like to suggest an additional reference: A monograph written by Z. W. Birnbaum (1979) entitled On the Mathematics of Competing Risks. This monograph is an introductory report on the mathematics of competing risks which attempts to unify the theory common to several disciplines, including actuarial science and biostatistics. It is available, free of charge, from the publications office of the National Center for Health Statistics.

I will now make some remarks which some of our biometrician friends here may possibly regard as controversial. I would, therefore, be most interested in their reaction to my ensuing comments.

The recent actuarial literature, as well as the biostatistics literature, contains a number of papers on Bayesian approaches to the construction of life tables. In particular, I refer to the papers on Bayesian graduation by Kimeldorf and Jones (1967), and Hickman and Miller (1977 and 1979). In addition, Cornfield and Detre (1977) have written a paper entitled Bayesian Life Table Analysis. There is also a Lindley (1979) article entitled Analysis of Life Tables with Grouping and Withdrawals. This work describes a Bayesian approach to a problem considered earlier by Breslow and Crowley (1974). Some of the advantages of a Bayesian approach to graduation are that:

(1) It forces the analyst to explicitly state his principal assumptions;

(2) It permits the analyst to make use of his prior subjective notions of the salient aspects of the data;

(3) It is a unified and formal approach; and finally,

(4) It provides a posterior distribution of the statistics of interest, thereby allowing the analyst to compute estimated variances of the statistics of interest.

In my own work at HUD, I have used a Bayesian graduation procedure to construct a double-decrement table for single-family FHA 30-year term mortgage insurance contracts.

Moving briefly to Dr. Breslow's presentation, I had a slight concern with his application. If I consider his final model as a multidimensional contingency table, the table would have 7 dimensions since the regression model has 1 dependent and 6 predictor variables. Thus, I would have only 268 observations partitioned into a minimum of 128 cells--a very sparse table indeed. Hopefully, Dr. Breslow and/or Dr. Koch can allay my concerns here.

In conclusion, I urge my fellow actuaries not to end their statistical education with Part V of the Society's examinations, but to follow the lead of Bob Johansen and at least become familiar with some of the multivariate statistical procedures discussed here today.

REFERENCES

Birnbaum, Z. William, On the Mathematics of Competing Risks, Vital and Health Statistics: Series 2, Data Evaluation and Methods Research; No.77, Public Health Service #79-1351, Hyattsville, Maryland, 1979.

Breslow, N. and Crowley, J., A large sample study of the life table and product limit estimates under random censorship, Annals of Statistics, Vol. 2, pages 437-453, 1974.

Cornfield, J. and Detre, K., Bayesian Life Table Analysis, The Journal of the Royal Statistical Society, Series B, Vol. 39, No. 1, pages 86-94, 1977.

Hickman, J. and Miller, R., Notes on Bayesian Graduation, Transactions of the Society of Actuaries, Vol. 29, pages 7-21, 1977.

Hickman, J. and Miller, R., Bayesian Bivariate Graduation and Forecasting, ARCH, 1979.3 Issue, pages 99-136, 1979.

Kilmeldorf, G. and Jones, D., Bayesian Graduation, Transactions of the Society of Actuaries, Vol. 19, pages 66-112, 1967.

Lindley, D. V., Analysis of Life Tables with Grouping and Withdrawals, Biometrics, Vol. 35, pages 605-612, 1979.

MR. ROY GOLDMAN: I would like to ask Dr. Breslow how his model relates to a log linear modeling. Is it the same thing?

DR BRESLOW: Yes. If one develops the model in terms of the conditional probability of survival over subintervals, the model states that the logarithm of the survival rate can be expressed as a linear function of the co-variables.

DR. KOCH: With respect to the related question that Tom Herzog raised about the estimation in the example, one of the strengths of the types of methods that have been used in biomedical research is that if one has a strong enough rationale to write down the likelihood function for the data, that is, if one can presume that the mathematical model that one is using

does adequately represent reality, then it indeed does provide a framework for generating estimates of the parameters that will have statistical validity. This can be done even if one has small samples. In actuarial studies, one often has available large amounts of data and estimates are produced from such data. In biomedical research, there are small amounts of data. The question is - what is the difference between the two frameworks. The difference is simply the mathematical model. In other words, if one wants to produce estimates that only depend on the data, one needs a large amount of data. In that case, one can produce those estimates, and there are neither assumptions nor necessarily any apologies for them.

On the other hand, if it turns out that one has a small amount of data, what one proceeds to do is to use a mathematical model. The mathematical model produces what one might otherwise have obtained with large amounts of data. Now in this particular case, one has to evaluate the assumptions. But if one can do analyses like what Dr. Breslow illustrated, one can often find that the data are compatible with the model to a reasonable degree of accuracy - in which case one then has the power of the results.

DR BRESLOW: I'd like to follow up on that. That's very much the attitude here. It was mentioned that we had 128 descriptions of the population, and certainly with thousands and thousands of observations, one could make a separate estimate of the survival within each one of those 128 cells. What we are trying to do with the multivariate model is to borrow strength from neighboring cells to make that estimate, and the kind of assumption which is involved is, for example, that the effect of white blood count is the same for young children as for older children, or that the effect of treatment is the same whether you have a high white count initially or a low white count. Actually it is of particular interest to look at interactions between treatment and the prognostic variables. Very often one finds that for those patients who are at a very high risk of death, e.g., persons with cancer, treatment simply does not work as a variable. Most of the treatment effects seen are in patients who were reasonably good risks.

MR. EDWARD A. LEW: Time and circumstances do not permit me to do justice to the three presentations on biometric methods which we have been privileged to hear. I regret not having been able to see in advance the manuscripts relating to Dr. Breslow's and Dr. Buncher's presentation. Fortunately I am familiar with one of Dr. Breslow's recent papers, entitled "Statistical Methods for Censored Survival Data", published in 1979 in Environmental Health Perspectives and I judge that Dr. Breslow has referred to that paper for his presentation.

The key concepts stressed by Dr. Breslow are the Kaplan-Meier product limit for estimating survival and Mantel's test for appraising differences between survival curves. The Kaplan-Meier product limit is essentially the same as the actuarial calculation of ${}_tp_x$, when the time intervals are small and the distribution of deaths and withdrawals is reasonably stable. Mantel's test statistic is a more sophisticated version of the chi-square formula and is based on the fact that under the null hypothesis the vector of observed deaths in a comparison of γ survival curves has a γ dimensional hypergeometric distribution. The Mantel test statistic has sometimes been referred to as the proportional hazards (PH) model, because it assumes that the forces of mortality in each of the γ survival curves are in constant ratios throughout the time period under consideration. The chi-square formula gives a conservative estimate for Mantel's test.

The catch in the more sophisticated satistical tests is whether the implicit assumptions made correspond reasonably well to actual circumstances. An example of where actual circumstances may not justify oversimplified assumptions is offered by epidemiological studies where the end point is a diagnosis or death from a particular cause and the situation is treated as presenting a "yes - no" dichotomy. The diagnosis of a particular disease or certification of death involves a spectrum of probabilities as to the accurate diagnosis of a particular disease or accurate designation of the cause of death. It would be more realistic to describe the situation in terms of a multinomial distribution of probabilities that the diagnosis of a particular disease or designation of a particular cause of death is in fact correct.

We are indebted to Drs. Elandt-Johnson and Johnson for an illuminating exposition of the different perspectives on mortality analysis held by actuaries and medical statisticians. These different points of view stem from the different kinds of problems and different forms of data which have confronted actuaries and medical statisticians.

Actuaries have usually been required to derive death rates for practical purposes, mostly to provide conservative estimates of future mortality. They have generally been presented with a relatively large volume of data and have customarily smoothed the crude death rates derived from such data to facilitate actuarial calculations. This procedure has rarely involved studies of the mortality fluctuations in order to learn more about the ranges of values containing the underlying death rates at some predetermined probability level. In most actuarial problems a deterministic approach has been sufficient for the purpose. In situations where differences in death rates appeared to reflect mainly the effects of different types of selection - rather than variations due to random sampling - a stochastic approach to judging the significance of differences in death rates did not seem to be called for.

Medical statisticians, on the other hand, have as a rule been asked to determine survival rates, because this criterion was easily understood by surgeons and physicians as appropriate for judging the efficacy of surgical or medical treatment. They were usually presented with a small body of data where variations due to random sampling could play a major role, so that the significance of the findings had to be appraised in probability terms. R. A. Fischer had pointed out that studies of variations due to random sampling automatically led to the concept of a frequency distribution, and that in such problems it was necessary to investigate both the nature of the distribution in random samples and the statistics designed to test the validity of the specifications of the distribution.

Actuaries did not begin to think of the mathematics of life contingencies in stochastic terms until risk theory had evolved to encompass life insurance problems (e.g. reinsurance). It is noteworthy, nevertheless, that Filip Lundberg's doctoral thesis on risk theory written in 1903 introduced the length of time until the event as the basic random variable, long before R. A. Fischer had laid the foundations of modern statistical methods. Lundberg developed his thesis while employed by a life insurance company, so that he presumably had the length of time elapsed until death in mind as the random variable in life insurance.

While the classical actuarial approach to life contingencies was deterministic and rested on expected values, it was reformulated with the advent of risk theory in terms of the random variable "length of time until death of a life age x." The distribution of this random variable is the life table. If we begin with this random variable and its distribution, we are led to a deeper conception of survival analysis. I would recommend the Johnsons' book on survival analysis, just published by John Wiley, as a very useful elaboration of the methods used by actuaries and medical statisticians in mortality studies.

I find myself in complete agreement with the Johnsons that death rates from the chronic diseases are not independent. Accordingly, I believe the National Life Tables by Cause to be misleading. A number of life insurance mortality investigations lend support to the proposition that most chronic diseases are interrelated.

I would like to emphasize that actuaries have been more cautious than medical statisticians in drawing conclusions about death rates from small samples. They have been more conscious of the various types of biases introduced by observational selections and less prone to judge differences in death rates in terms of test statistics which assume simple urn models. This has been particularly true of medico-actuarial investigations which offer many examples of how random fluctuations in small samples are overshadowed by biases introduced by self selection, selection due to screening and particularly class selection. In the words of E. B. Wilson, the excellence of a person as a statistician may depend largely on his or her ability to recognize the pitfalls due to observational selection.

Dr. Buncher's discussion of the question whether the Healthy Worker Effect is due to select mortality can partially at least be answered by the 1970-74 group life insurance experience for all industries combined. It shows that the death rates among actively employed men and women covered by group life insurance were distinctly lower than the contemporaneous ultimate ordinary insurance death rates in the age range 40 to 70 and slightly lower in the thirties. Accordingly, the indications are that selection for active employment is a significant force, especially as the group life insurance experience for all industries include some occupations with special hazards. We can attribute this selection to the screening of a healthy population for employment and to the survival in active employment of the healthier men and women. Studies by Fox and Collier in the Office of Population Censuses and Surveys in England have also shown that the low mortality in industrial cohorts is primarily due to selection for work and survival on the job.