BIOMETRICS 35, 25–39
March, 1979

Hazard Rate Models with Covariates

ROSS L. PRENTICE

Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, U.S.A.
and
Department of Biostatistics, University of Washington, Seattle 98195, U.S.A.

J. D. KALBFLEISCH

Department of Statistics, University of Waterloo, Ontario, Canada

Summary

Many problems, particularly in medical research, concern the relationship between certain covariates and the time to occurrence of an event. The hazard or failure rate function provides a conceptually simple representation of time to occurrence data that readily adapts to include such generalizations as competing risks and covariates that vary with time. Two partially parametric models for the hazard function are considered. These are the proportional hazards model of Cox (1972) and the class of log-linear or accelerated failure time models. A synthesis of the literature on estimation from these models under prospective sampling indicates that, although important advances have occurred during the past decade, further effort is warranted on such topics as distribution theory, tests of fit, robustness, and the full utilization of a methodology that permits non-standard features. It is further argued that a good deal of fruitful research could be done on applying the same models under a variety of other sampling schemes. A discussion of estimation from case-control studies illustrates this point.

1. Introduction

The recent statistical literature as well as recent meetings and symposia attest to a considerable current interest in specialized statistical methods for the analysis of time to occurrence or 'failure' time data. Much of this interest appears to have been stimulated by problems arising in medical research though rather similar problems arise, for example, in industrial life-testing and demography.

Frequently the primary objective of a failure time study concerns the association between certain covariates $\mathbf{z} = (z_1, \ldots, z_p)$ and the time $T \geq 0$ to the occurrence of a certain event. For example, a clinical study may be designed to compare several treatment programs in respect to the time, T, to recurrence of a disease. The regression vector $\mathbf{z}$ would include indicator components for treatment as well as other prognostic factors.

The distribution of failure time, T, can be represented in the usual manner in terms of density or distribution functions as well as in more specialized ways such as the hazard function. If the underlying failure process is continuous the instantaneous failure rate or hazard function at time t among individuals with covariate $\mathbf{z}$ is defined as

$$\lambda(t; \mathbf{z}) = \lim_{\Delta t \to 0} P[t \leq T \leq t + \Delta t \mid T \geq t, \mathbf{z}]/\Delta t.$$

Key Words: Failure time data; Hazard function; Regression, Proportional hazards model; Case-control studies; Competing risks; Time-dependent covariates; Relative risk.

The hazard function gives the 'risk' of failure at any time t, given that the individual has not failed prior to t. Since the notion of failure rate is basic and conceptually simple, $\lambda(t;\mathbf{z})$ provides a convenient starting point for modelling the association between $\mathbf{z}$ and t.

One such model, introduced by Cox (1972) provides the basis for most of the discussion of this paper. This model presumes that covariates affect the hazard function in a multiplicative manner according to

$$\lambda(t; \mathbf{z}) = \lambda_0(t)e^{\mathbf{z}\beta}, \tag{1}$$

where λ_0 is an unspecified function of time and β is a column p-vector of parameters. The factor $e^{\mathbf{z}\beta}$ describes the risk of failure for an individual with regression variable $\mathbf{z}$ relative to that at a standard value $\mathbf{z} = \mathbf{0}$ (at all time points). This is a formalization of the relative risk concept basic to epidemiological research. Since the ratio of hazard functions corresponding to any two $\mathbf{z}$-values is a constant, (1) is often referred to as the proportional hazards model.

A second partially parametric model, more akin to ordinary linear regression, specifies the covariates to act multiplicatively on failure time itself (or linearly on log-failure time) rather than multiplicatively on the hazard function: Suppose that $t' = te^{\mathbf{z}\beta}$ has a fixed hazard function $\lambda_0(t')$. This gives a hazard function for failure time t that can be written

$$\lambda(t; \mathbf{z}) = \lambda_0(te^{\mathbf{z}\beta})\, e^{\mathbf{z}\beta}. \tag{2}$$

Since covariates alter by a scale factor the rate at which an individual traverses the time axis, (2) may be referred to as the accelerated failure time model. Estimation techniques based on (1) and (2) are reviewed and compared in Section 2.

In order to yield a unified approach to data analysis, it is important that a failure time model adapt readily to a variety of data types more complex than that considered above. One such generalization, the full value of which is perhaps yet to be realized, allows the covariate $\mathbf{z}$ to vary over time (Section 3). For example, an industrial safety study may be designed to examine the association between $\mathbf{z}(u)$, the accumulated radiation exposure of a uranium miner at age u, and the age, T, at incidence of respiratory cancer or other diseases. This example points to a further generalization; namely, the presence of multiple failure types or competing risks. With multiple failure types the questions of interest typically involve both the effect of covariates on specific causes of failure as well as the relationship between the failure types (Section 4). As a final generalization, failure time itself may be multivariate as occurs, for example, in the analysis of times to attack in epileptic patients or times to infection in cancer patients receiving high doses of chemotherapy (Section 5).

The discussions of Sections 2 through 5 deal with a prospective follow-up type of sampling in which individuals are identified, usually at some specified origin of time measurement, and followed forward to observe their respective failure times. It is also important that a statistical model be adaptable to other sampling procedures. In Section 6 the same statistical models are considered in relation to case-control or retrospective sampling. In this type of study, subjects are selected on the basis of their failure times and types after which one 'looks back' to ascertain the corresponding covariate values or covariate functions.

In each section the literature is briefly summarized and suggestions for further work are given. The main message that emerges is twofold: First, the proportional hazards model provides a unified approach to inference from a variety of types of failure time data, though further theoretical and applied work is indicated in a number of directions. Second, fruitful research can be anticipated on the adaptation of failure time models to case-control and other sampling schemes. Many of the topics mentioned above are discussed in detail in the book manuscript, Kalbfleisch and Prentice (1979).

2. *Prospective Sampling with Fixed Covariates*

2.1. *The Likelihood Function*

Consider a prospective study in which n individuals with covariate values $\mathbf{z}_1, \ldots, \mathbf{z}_n$ are selected and followed forward to observe their failure times $T_1, \ldots, T_n$. The need for specialized statistical methods arises in this context because of the usual presence of censoring in the recording of failure times. Right censored data typically occur because some study subjects are still without failure or have been 'lost to follow-up' at the time of data analysis. Let $t_1, \ldots, t_n$ denote the observed failure times that may or may not be right censored and let $\delta_1, \ldots, \delta_n$ represent the corresponding censoring indicators defined by $\delta_i = 0$ if t_i is censored $(T_i > t_i)$ and $= 1$ if t_i is uncensored $(T_i = t_i)$. The usual assumption concerning censoring is that, conditionally on $\mathbf{z}$, the censoring and failure mechanisms are independent in the sense that, at any point in time, individuals are not selectively censored because of a relatively poor or relatively good prognosis. Such censoring schemes are termed independent.

If the censoring scheme is independent, it is intuitively clear that a censored value t_i contributes only the information that T_i exceeds t_i. Such a censored failure time thereby contributes to the likelihood a factor $F(t_i; \mathbf{z}_i) = P\{T_i > t_i; \mathbf{z}_i\}$. The function $F(t_i; \mathbf{z}_i)$ is known as the survivor function. The likelihood on the data $(t_i, \delta_i, \mathbf{z}_i)$, $i = 1, \ldots, n$ is thus proportional to

$$L = \prod_{i=1}^{n} [f(t_i; \mathbf{z}_i)^{\delta_i} F(t_i; \mathbf{z}_i)^{1-\delta_i}] \tag{3}$$

where $f(t; \mathbf{z}) = -dF(t; \mathbf{z})/dt$ is the density function of T. The above "derivation" of (3) is heuristic and a more careful derivation for general independent censoring schemes would involve building the likelihood up conditionally at successive time points.

In order to fit such models as (1) and (2) it is instructive to rewrite the likelihood in terms of the hazard function. From the above definition $\lambda(t; \mathbf{z}) = f(t; \mathbf{z})/F(t; \mathbf{z})$. Solving for $F(t; \mathbf{z})$ gives

$$F(t; \mathbf{z}) = \exp\left[-\int_0^t \lambda(u; \mathbf{z})\, du\right]$$

so that (3) becomes

$$L = \prod_{i}^{n} \left\{\lambda(t_i; \mathbf{z}_i)^{\delta_i} \exp\left[-\int_0^{t_i} \lambda(u; \mathbf{z}_i)\, du\right]\right\}. \tag{4}$$

2.2. *Parametric Methods*

If $\lambda(t; \mathbf{z})$ is specified up to a finite number of parameters $\boldsymbol{\theta}$, ordinary parametric likelihood methods can be applied to (4). It should be noted, however, that the presence of censoring, except in a few extremely special cases, precludes the possibility of tractable exact distribution theory. Standard asymptotic likelihood methods would, for example, ascribe an asymptotic normal distribution to the maximum likelihood (ml) estimator $\hat{\boldsymbol{\theta}}$ (the solution to $\partial \log L/\partial\boldsymbol{\theta} = 0$), with mean the 'true' value of $\boldsymbol{\theta}$ and variance estimator $I^{-1}(\hat{\boldsymbol{\theta}})$, where $I(\boldsymbol{\theta}) = -\partial^2 \log L/\partial\boldsymbol{\theta}\, \partial\boldsymbol{\theta}$ is the observed Fisher information. A Newton-Raphson iteration is often a convenient procedure for the computation of $\hat{\boldsymbol{\theta}}$ and $I^{-1}(\hat{\boldsymbol{\theta}})$.

One parametric model that has found extensive application is the (two-parameter) Weibull regression model given by (1) with $\lambda_0(t) = \lambda p(\lambda t)^{p-1}$. It is of some interest to note that this model is in the intersection of the partially parametric models (1) and (2) and further it may be shown that any model in this intersection is on the Weibull type. Thus, the Weibull regression model is uniquely characterized as that model in which the regressor variables act multiplicatively both upon the hazard function and upon failure time itself. The asymptotic theory outlined above is readily applied to the Weibull model with $\boldsymbol{\theta} = (\lambda, p, \boldsymbol{\beta})$. In fact, such a model and the exponential special case ($p = 1$) have been frequently used in both the analysis of clinical trial data and industrial life-testing data. Since it is also a special case of (2), the Weibull model can be written in log-linear form as

$$y + \mathbf{z}\boldsymbol{\beta}' = \alpha + \sigma v,$$

where $y = \log t$, $\alpha = -\log \lambda$, $\sigma = p^{-1}$, $\boldsymbol{\beta}' = p^{-1}\boldsymbol{\beta}$ and the 'error' variable v has an extreme value density, $f(v) = \exp(v - e^{v})$.

Other parametric special cases of (2) include the log-normal, the log-logistic and the generalized gamma regression models (e.g., Farewell and Prentice 1977) corresponding to error quantities v that are normal, logistic and the logarithm of gamma variates, respectively.

Even in this fully parametric setting, further work would be desirable on precise conditions on the censoring mechanism, the regressor variables, and the form of $\lambda(t; \mathbf{z})$ that would be sufficient to ensure asymptotic normality. In the context of specific models, further work is merited on such procedures as parameter transformations to improve the adequacy of the asymptotic normal approximations with small or moderate sample sizes.

2.3. Inference on the Proportional Hazards Model

2.3.1. Estimation of the Regression Coefficients

A remarkable feature of the proportional hazards model (1) is that it admits feasible estimation of the relative risk parameter, $\boldsymbol{\beta}$, without restricting the "shape" function λ_0. Cox *(1972) gave a rather informal justification of a likelihood function for $\boldsymbol{\beta}$ which he (Cox 1975)* later considered in more detail under the term partial likelihood. In order to illustrate the argument, we first introduce some notation and consider an alternative derivation of the likelihood function (4).

Suppose, to begin, that there are no ties among the uncensored failure times and let $t_{(1)} < t_{(2)} \cdots < t_{(r)}$ denote the $r = \delta_1 + \cdots + \delta_n$ uncensored failure times in the sample of size n. Let $\mathbf{z}_{(i)}$ denote the covariate vector corresponding to $t_{(i)}$ and let $R(t_{(i)})$ be the set of labels of individuals at risk just prior to $t_{(i)}$. Let A_i represent the event that item (i) fails at time $t_{(i)}$ and let B_i be the event that describes the observed process up to time $t_{(i)}$ including all failure and censoring information as well as the information that a failure occurs at $t_{(i)}$. The likelihood function can then be written

$$L = \prod_{i=1}^{k} P(A_i | B_i) \prod_{i=1}^{k+1} P(B_i | B_{i-1}, A_{i-1})$$

where A_0, B_0 give the history to time 0 and B_{k+1}, the history to time ∞. In view of the definition of the hazard function

$$P(A_i | B_i) = \lambda(t_{(i)}; \mathbf{z}_{(i)}) \Big/ \sum_{l \in R(t_{(i)})} \lambda(t_{(i)}; \mathbf{z}_l).$$

Under a proportional hazards model, $\lambda(t; \mathbf{z}) = \lambda_0(t) \exp(\mathbf{z}\boldsymbol{\beta})$, the first term $\Pi P(A_i | B_i)$ is independent of λ_0 and can be written

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{r} \left\{ e^{\mathbf{z}_{(i)}\boldsymbol{\beta}} \Big/ \sum_{l \in R(t_{(i)})} e^{\mathbf{z}_l \boldsymbol{\beta}} \right\} \tag{5}$$

while the second will generally involve both λ_0 and $\boldsymbol{\beta}$. For such a factorization, Cox (1975) showed that the "partial" likelihood (5) can, under conditions that are evidently quite mild, be manipulated in the usual manner for asymptotic likelihood inference on β.

A number of other approaches give further insight into (5). Kalbfleisch and Prentice (1973) showed that, in the absence of censoring, (5) is precisely the likelihood based on the marginal distribution of the ranks of the failure times. A group invariance argument showed the rank vector to be sufficient for β in the 'absence of knowledge' of λ_0. With censored data, (5) is the likelihood corresponding to the marginal probability of possible underlying rank vectors consistent with the observed data. Breslow (1974) derived (5) as a 'maximized' likelihood for β by restricting the λ_0 function in (4) to be a step function with discontinuities at observed failure times. Breslow's development also required censored failure times to be shifted to the nearest smaller uncensored failure time. Holford (1976) took a similar tack by specifying step function intervals that were independent of the data and then allowing the intervals to become uniformly small.

Several generalizations of (5) have been proposed to accommodate tied failure times. The marginal likelihood approach of Kalbfleisch and Prentice (1973) retains the same relative risk parameter β as the target of estimation. The corresponding likelihood, however, involves unduly complex calculations if the number of ties is at all large. The same comment applies to the tied data partial likelihood of Cox (1972) which was based on an approximating logistic regression model with 'odds ratio' β. With a small fraction of tied failure times, however, both likelihoods will approximate.

$$L(\beta) = \prod_{1}^{r} \left[e^{\mathbf{s}_{(i)}\beta} \Big/ \left(\sum_{l \in R(t_{(i)})} e^{\mathbf{z}_l \beta} \right)^{d_i} \right], \tag{6}$$

where d_i is the number of failures at $t_{(i)}$ and $\mathbf{s}_{(i)}$ is the sum of the covariate values for the d_i individuals. Expression (6) arises directly from the approach of Breslow.

Standard asymptotic likelihood procedures applied to (5) or (6) give rise to straightforward estimation of β and tests of hypotheses about β. At $\beta = \mathbf{0}$, the score statistic from (6) is

$$\mathbf{U} = [\partial \log L(\beta)/\partial \beta]_{\beta=\mathbf{0}} = \sum_{i=1}^{r} \left[\mathbf{s}_{(i)} - d_i n_i^{-1} \sum_{l \in R(t_{(i)})} \mathbf{z}_l \right], \tag{7}$$

where n_i is the number of individuals in $R(t_{(i)})$. This statistic generalizes the Mantel-Haenzel (e.g., Mantel 1966) or log-rank (Peto and Peto 1972) test for the comparison of several survival curves to arbitrary null tests on regression coefficients.

If the number of tied failure times is large, the approximation, (6), becomes poor and it is preferable to use a discrete failure time model. The regression model (1) has been generalized in two ways to include discrete cases. Kalbfleisch and Prentice (1973) showed that grouping the continuous model (1) gave a discrete model with simple properties that has the advantage of retaining the same relative risk parameter β. Prentice and Gloeckler (1978) give asymptotic likelihood methods based on this model. Cox (1972) proposed a different discrete model which allowed generalization of the partial likelihood method. In this model, the conditional probability of failure in an interval was modeled as a binary logistic regression model. Although the regression parameter is different from that in (1), for small grouping intervals the two are very close. Thompson (1977) considers asymptotic inference from this model.

2.3.2. Properties of β Estimators

Let us turn now to properties of estimation based on (5). Though there can be little doubt concerning the applicability of usual asymptotic formulae under mild conditions, such conditions have yet to be precisely written down. An exception is a recent technical report (Tsiatis 1978) that brings weak convergence results to bear on the distribution of the maximum partial likelihood estimator. Even this work includes such assumptions as bounded regression variables which undoubtedly can be relaxed.

Assuming that asymptotic likelihood methods apply, one can inquire as to the efficiency of the ML estimator $\hat{\beta}$ from (5) or the efficiency of the test based on (7). As noted above, Holford (1976) and Breslow (1974) give derivations of (5) as the maximized likelihood of β corresponding to increasingly rich parametric models in (1). These results suggest that no efficiency improvement is likely to be possible as long as the λ_0 function remains unrestricted. Efron (1977), using another parametrization and more formal arguments, has shown this to be the case. Attention then turns to the efficiency of β estimation from (5) relative to that based on specific parametric models.

First, with uncensored data, it can be readily shown that the score statistic for testing $\beta = \mathbf{0}$ is (asymptotically) fully efficient relative to any parametric model in (1) of the form $\lambda_0(t) = \lambda h_0(t)$ with $h_0(t)$ known and λ, a scale parameter to be estimated. The situation changes somewhat as β differs from $\mathbf{0}$. Still with uncensored data and with a single regressor variable, Kalbfleisch (1974) showed that the efficiency of $\hat{\beta}$ from (5) relative to the estimator from a model $\lambda_0(t) = \lambda h_0(t)$ is approximately exp $(-\beta^2\mu_2)$ where μ_2 is the second central moment of the distribution of z-values. While this loss may not be severe in many situations of interest, the dependence of efficiency on the value of the regression parameter is a situation unfamiliar to ordinary linear regression. A reduction in efficiency may occur, even if $\beta = 0$, if the censoring depends on the values of the regression variables. Thus, even the score test for $\beta = 0$ will involve some loss of efficiency if the censoring depends on $\mathbf{z}$.

General approaches to the efficiency of β estimation from (5) are given by Efron (1977) and Oakes (1977). Estimation from (5) is based on the conditional probability that item (i), with covariate $\mathbf{z}_{(i)}$, fails given that one item is known to have failed among the risk set $R(t_{(i)})$, $i = 1, \ldots, k$. The relative efficiency of the partial likelihood estimator depends on the degree of variation in the expected $\mathbf{z}$-values from these conditional distributions across the failure times $t_{(i)}$, $i = 1, \ldots, k$. More precisely the difference between the 'marginal' information matrix for $\hat{\beta}$ from the parametric model with $\lambda_0(t) = \lambda h_0(t)$ and that from (5) is estimated by the covariance function of expected $\mathbf{z}$-values for failures, across the time points.

Similarly the difference in information matrices between that based on a more flexible parametric model, with $\lambda_0(t) = \lambda \exp [\alpha \mathbf{w}(t)]$ where $\mathbf{w}(t) = [w_1(t), \ldots, w_r(t)]$ consists of specified functions of time, and (5) is estimated by the deviation from regression matrix when the expected z-values for individuals that fail at $t_{(1)}, \ldots, t_{(k)}$ are regressed on $\mathbf{w}(t)$. Variations in expected $\mathbf{z}$-values over time can be induced by β-values that differ from zero or by differential censoring rates over $\mathbf{z}$. Such variations introduce correlations, between estimators of β and estimators of the λ_0 functon, that can be exploited by the parametric estimators but not by the likelihood (5). As an extreme case, consider a two sample problem in which one sample is uncensored while the second is totally censored at time t_0. The contribution to the partial likelihood at any failure time after t_0 is independent of β since all items in the risk set have the same covariate value. Parametric analyses, on the other hand, can utilize failure times past t_0 to estimate parameters in the λ_0 function more precisely and hence yield more precise estimates of β.

An attractive feature of inference based on (5) is the robustness implied by the arbitrariness in the λ_0 function. Steve Samuels, in a 1977 unpublished doctoral dissertation at the University of Washington, has examined the robustness of $\hat{\beta}$ from (5) more formally. As expected, since (5) utilizes only the ranks of the failure times, the influence curve (Hampel 1974) for $\hat{\beta}$ is bounded in respect to 'errors' in the failure times or censoring indicators.

Unfortunately, however, the possible influence of errors in **z** can be of any magnitude whatever. Modified methods that guard against an undue influence of a small number of extreme **z**-values may be worthwhile.

Rather little formal consideration of inference based on (6) has taken place. Most of the above remarks are likely to apply to $\hat{\beta}$ from (6), though an asymptotic bias that increases with the severity of the failure time grouping can be anticipated.

2.3.3. Survivor Function Estimation

The survivor function $F(t; \mathbf{z}) = P(T \geq t; \mathbf{z})$ from the proportional hazards model (1) can be written $F(t; \mathbf{z}) = F_0(t)^{[\exp(\mathbf{z}\beta)]}$, where $F_0(t) = \exp\,[-\int_0^t \lambda_0(u)\,du]$. At any specified β a 'nonparametric ML' estimation of F_0 is readily carried out (e.g., Kalbfleisch and Prentice 1973) that gives an estimator for $F(t; \mathbf{z})$ of the form

$$\hat{F}(t; \mathbf{z}, \beta) = \prod_{\{i \mid t_{(i)} < t\}} \hat{\alpha}_i^{[\exp(\mathbf{z}\beta)]}. \tag{8}$$

At $\beta = \mathbf{0}$ this estimator reduces to the product-limit estimator of Kaplan-Meier (1958)

$$\prod_{\{i \mid t_{(i)} < t\}} \left[\frac{n_i - d_i}{n_i}\right] \tag{9}$$

which generalizes the usual empirical distribution function to censored data. If only a single failure occurs at $t_{(i)}$

$$\hat{\alpha}_i = \left[1 - e^{\mathbf{z}_{(i)}\beta} \Big/ \sum_{\ell \in R(t_{(i)})} e^{\mathbf{z}_{(i)}\beta}\right]^{\exp(-\mathbf{z}_{(i)}\beta)}.$$

More generally, however, the $\hat{\alpha}_i$'s will involve iterative calculations. The step function approximation of λ_0 utilized by Breslow leads to an estimator of the form (8) with

$$\hat{\alpha}_i = \exp\left[-d_i \Big/ \sum_{\ell \in R(t_{(i)})} e^{\mathbf{z}_\ell \beta}\right]. \tag{10}$$

Estimators $\hat{\beta}$ from (5) or (6) may be inserted in (8) in order to estimate the survivor function at any specified **z**.

Rather little attention has been directed to the asymptotic distribution of the process $\hat{F}(t; \mathbf{z}, \hat{\beta})$ though Tsiatis (1978) has shown weak convergence to a Gaussian process based on $\hat{\beta}$ from (5) and $\hat{\alpha}_i$ from (10), again under rather severe restrictions such as bounded covariates **z**.

2.4 Inference on the Accelerated Failure Time Model

Consider now the estimation of β in the accelerated failure time model, $\lambda(t; \mathbf{z}) = \lambda_0(t\, e^{\mathbf{z}\beta})\, e^{\mathbf{z}\beta}$. As indicated above this model is equivalent to the class of linear models for $y = \log t$ with error densities in direct correspondence with the shape function λ_0. It follows that problems concerning the estimation of β will be the same as those in ordinary linear regression with the additional complication of censored data. One approach to β estimation would select a rich and plausible class of error densities in the log-linear model and apply parametric methods as outlined in Section 2.2 (e.g., Farewell and Prentice 1977). A related approach to robust estimation from this model would involve the generalization of M (ML-type) estimation to censored data. No such effort appears to have yet taken place.

Several authors have considered the generalization, to censored data, of rank tests for the comparison of several samples (Gehan 1965; Mantel 1966; Breslow 1970; Peto and Peto 1972). Prentice (1978) used the marginal probability of possible rank vectors for $w_i = y_i - \mathbf{z}_i\boldsymbol{\beta}_0$ to derive a censored data rank test for $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ corresponding to any generating density f (or hazard function λ_0). In the manner of Jurečková (1971), an estimator was also defined as the $\boldsymbol{\beta}$ value for which the score statistic (a step function of $\boldsymbol{\beta}$) is as close as possible to zero. Further work is required on both the formal properties and numerical aspects (as in the uncensored case) of such rank regression estimators. At least with uncensored data, however, such a rank regression estimator is fully efficient if the sampling error density agrees with the assumed score generating density. Further, the estimator has good efficiency, relative to fully parametric estimators, for sampling densities that are 'near' the score generating density. Note that since these procedures involve the ranks of log-failure times 'centered' about $\boldsymbol{\beta}$, no loss of efficiency is experienced as $\boldsymbol{\beta}$ moves away from $\mathbf{0}$. The same loss of efficiency as was noted with the partial likelihood estimator can, however, be expected if censoring varies with $\mathbf{z}$.

In correspondence with the methods of Section 2.3.1, one might even hope to obtain practical inference techniques for $\boldsymbol{\beta}$ that are uniformly applicable under any λ_0 function in (2). One possibility in this regard would be to approximate λ_0 by a step function and then to require the step function interval to become uniformly small. Though such a procedure has not been developed it could possibly be useful for linear regression analysis even outside of the failure time setting.

2.5. Generalizations, Tests of Fit and Other Problems

It should be commented that the form of the parametric component, $e^{\mathbf{z}\boldsymbol{\beta}}$, of (1) and (2) can be changed to other parametric forms with no essential complication in the inference techniques discussed above. For example, in carcinogenesis modelling (e.g., Hartley and Sielken 1977) a proportional hazards model (1) is frequently assumed but with a multiplicative factor $(\alpha_1 + \beta_1 z) \ldots (\alpha_k + \beta_k z)$, where z is the dosage of the carcinogen and k is the number of 'stages' in the carcinogenic process.

A second generalization arises since the λ_0 functions in either (1) or (2) may be relaxed to differ among subsets or strata in the population. The likelihood functions for $\boldsymbol{\beta}$ from such stratified models are simply products over strata of the likelihood functions given above. The corresponding log-likelihood derivatives and test statistics are therefore formed by summing the earlier expressions over strata. In fact, censored data rank tests along with a stratification of confounding factors provide a viable means of analysis for many data sets. Peto, Pike, Armitage, Breslow, Cox, Howard, Mantel, McPherson, Peto and Smith (1977) give a non-technical discussion of this approach.

Another rather direct generalization of the methods of Section 2.3 occurs in a cross-sectional follow-up study in which individuals are identified at times that may exceed the zero of the follow-up axis and are followed forward to observe their failure times. The likelihoods (5) and (6) are still applicable with risks sets $R[t_{(i)}]$ consisting only of individuals under follow-up at $t_{(i)}$.

As discussed in the next section, a very important generalization is afforded by the fact that the proportional hazards model (1) adapts readily to the inclusion of time-dependent covariates. This admits formal tests of (1) against specific alternatives and generalizes (1) in a manner that, at least conceptually, spans all possible $\lambda(t; \mathbf{z})$. Some suggestions for graphical checks on (1) have also been given (Kalbfleisch and Prentice 1979). Further work is clearly desirable toward testing the goodness of fit of (1) and (2) and discriminating between the two models.

Additional problems that merit further consideration in relation to (1) and (2) include sequential testing methods, multiple comparisons (see Koziol and Reid 1978) and sample size determinations. Missing data among the regression variables is very common, especially in clinical trial data. In many instances, excluding cases with one or more variables missing can result in exclusion of a large portion of the data. Improved methods of handling this problem would be desirable.

3. *Prospective Sampling with Time Dependent Covariables*

As mentioned above, one generalization of these regression models is to allow the covariate vector to vary with time. Let $\mathbf{z}(u)$ be the covariate vector at time u and $Z(t) = \{\mathbf{z}(u): u \leq t\}$ denote the whole covariate function up to time t. Several types of time dependent covariates can be distinguished:

(*i*) The covariate may be deterministic in character. This occurs when the covariate is under the control of the experimenter. One use of this type of covariate arises in a fixed covariate situation where one may define components of $\mathbf{z}(u)$ to test the fit of the model (1). Thus, one might specify $\mathbf{z}(u) = \{\mathbf{z}, \mathbf{z}\, g(u)\}$ for some known function such as $g(u) = u$ or $g(u) = \log u$ and test that the corresponding regression coefficients are 0. Such covariables also arise in reliability applications where, for example, in testing insulation in electrical cable, the experimenter may increase the voltage in a deterministic way.

(*ii*) The covariate may represent the realization of some "external" process that may affect, but is unaffected by, the failure experience of the population under study. Environmental or occupational factors would often be of this type. This type of covariable is similar to *i*) and analysis would be made by conditioning on the realized value of $z(u)$.

(*iii*) Finally, $\mathbf{z}(u)$ may describe a process that is dependent on, or even attempts to assess, the propensity of the corresponding study subject to fail at time u. Since such covariables may well be influenced by other components of the regression vector, a great deal of care is required in their interpretation. One use of such a covariable is discussed in the next section.

The hazard function definition given in the introduction can be extended to

$$\lambda[t; Z(t)] = \lim_{\Delta t \to 0} P[t \leq T \leq t + \Delta t \mid T \geq t, Z(t)]/\Delta t. \tag{11}$$

This function gives the instantaneous risk of failure at time t given that failure has not occurred prior to t and given the covariate function up to time t. If (11) depends on $Z(t)$ only through the covariate vector $\mathbf{z}(t)$ at time t (this would not appear to be a severe restriction for suitable definition of $\mathbf{z}(u)$), (1) generalizes naturally to

$$\lambda[t, \mathbf{z}(t)] = \lambda_0(t) \exp [\mathbf{z}(t)\beta] \tag{12}$$

with no change in the multiplicative interpretation of covariate effects.

Generalizations of (2) to time-dependent covariates do not seem to have been considered though the Weibull special case can, of course, be generalized as in (12) because of its proportional hazards interpretation. More generally, in order to preserve the accelerated failure time aspect of (2) one might require $t' = t \exp [\mathbf{z}(t)\beta]$ to have a fixed hazard function $\lambda_0(t')$. The resulting hazard function for t' may be too complex to be useful. The special case of a step function $\mathbf{z}(t)$, however, would simply give

$$\lambda[t, \mathbf{z}(t)] = \lambda_0\{t \exp [\mathbf{z}(t)\beta]\} \exp [\mathbf{z}(t)\beta]. \tag{13}$$

If the covariate function is of type *i*) or *ii*), (4) again gives the likelihood based on the distribution of T given Z upon replacing $\lambda(u, \mathbf{z})$ by $\lambda[u, \mathbf{z}(u)]$. As will be discussed in more detail elsewhere, inference can be based on this same likelihood with covariates, of type *iii*), but the likelihood then has only a partial rather than a conditional likelihood interpretation. The inference techniques discussed in Sections 2.3 and 2.4 will apply also to (12) and (13) simply be replacing $\mathbf{z}$ at time u by $\mathbf{z}(u)$. For example, the partial likelihood (5) becomes

$$\prod_{1}^{r} \left\{ \exp\,[\mathbf{z}_i(t_{(i)})\beta] \Big/ \sum_{\ell \epsilon R(t_{(i)})} \exp[\mathbf{z}_\ell(t_{(i)})\beta] \right\}. \tag{14}$$

Conditions for the corresponding asymptotic likelihood theory to apply are even less clear than in the fixed covariate case. The efficiency discussion of Section 2.3.2 would also apply to the maximum partial likelihood estimator based on (12) for the case of an external covariate process $Z(t)$ (see Oakes 1977 and Efron 1977). It is clear that a good deal of valuable additional work could be carried out, for example, on properties and applications of inference based on (14). Crowley and Hu (1977) give an interesting application of (14) to heart transplant survival data. The application of time-dependent covariates to competing risk problems is considered in the next section.

4. *Regression Analysis with Multiple Failure Types*

Suppose in addition to failure time T and covariate $\mathbf{z}$, or covariate function $Z(t)$, that each failure is of one of m distinct types or causes denoted by $J \epsilon \{1, 2, \ldots, m\}$. The hazard function concept (2) is readily generalized to multiple failure types by defining the cause-specific failure rate (Chiang 1968, p. 244; Holt 1978) from cause j at time t to be

$$\lambda_j[t, Z(t)] = \lim_{\Delta t \to 0} P[t \le T \le t + \Delta t, J = j \,|\, T \ge , Z(t)]/\Delta t \tag{15}$$

for all (t, j).

A proportional hazards model would specify

$$\lambda_j[t, Z(t)] = \lambda_{0j}(t) \exp\,[\mathbf{z}(t)\beta_j] \tag{16}$$

for failure type j. Note that both the shape of the underlying cause-specific failure rate and the regression coefficient are allowed to vary with j. A similar generalization of (13) can also be written (at least for constant or step function covariates) that permits accelerating factors $\exp\,[\mathbf{z}(t)\beta_j]$ to differ among failure types. As shown in Prentice, Kalbfleisch, Peterson, Flournoy, Farewell and Breslow (1978), the conditional or partial likelihood function (depending on the nature of Z) for the cause-specific hazard functions factors into a separate component for each j. In fact the factor involving $\lambda_j[t, Z(t)]$ is exactly (4) with $\lambda(t, z)$ replaced by $\lambda_j[t, Z(t)]$ and with $\delta_i = 0$ for all failures of causes other than j. It follows that the inference techniques discussed above may be applied for inference on a particular β_j or λ_{0j} in a proportional hazards or accelerated failure time model simply by regarding failures of causes other than j as censored at their failure times. Note that no assumption has been made concerning the relationship between the failure types. Also note, however, that the cause-specific regression coefficients describe covariate effects on specific causes of failure under current study conditions. There is no implication that similar effects would prevail if the failure process were somehow altered so that some failure types could no longer occur.

The classical approach to studying the interrelation between failure types has been to assume the existence of latent failure times $T_1, \ldots, T_m$ corresponding to each of the m failure types, from which $T = \min\,(T_1, \ldots, T_m)$, $J = (j \,|\, T_j \le T_k$ all $k)$. Several authors (e.g., Cox 1959; Tsiatis 1975) have shown that data of the type (T, J) do not permit one to 'identify' the joint or marginal distributions for $T_1, \ldots, T_m$, nor to test independence of latent failure times, without introducing additional untestable assumptions (such as a parametric model for the joint distributions of $T_1, \ldots, T_m$). Evidently data beyond (T, J) are required to study interrelations among the failure types. One possibility is to utilize the 'pattern' of pathological entities or equipment faults at failure (e.g., Breslow, Day, Tomatis and Turusov 1974; Wong 1977). A rather different approach (Prentice *et al.* 1978) would define risk indicators for some failure types and relate those as time-dependent variables to the cause-specific failure rates for other failure types using the methods of Section 3. For example, in order to test independence of failure and censoring (viewed as a competing risk) one could define a variable, such as performance status in a clinical trial, that attempts to indicate an individual's propensity to fail. A dependence of the censoring rate on such a variable, relative to other study subjects at the same follow-up time, would give evidence for a lack of independence between censoring and failure.

5. *Multivariate Failure Time Data*

One type of multiple failure time data arises when subjects are followed until the occurrences of two or more essentially different events. For example, in a clinical trial in cancer chemotherapy, patients may be followed and times to remission recorded for both primary and secondary disease sites. Such multivariate data is closely related to the competing risks type of data in Section 4 and in fact, the time until the first occurrence might be viewed as a competing risk failure time. In addition to the cause-specific hazard rates (15), conditional hazards must be specified and details are given in Cox (1972). The analysis of the Stanford heart transplant data given by Crowley and Hu (1977) is an interesting example. Here, time to transplant, W, and time to death, T, may be thought of as separate "failure time" variates. Ignoring covariates, one needs to specify $\lambda_1(t)$, $\lambda_2(w)$ as in (15) and $\lambda_1(t \mid w)$ for $t > w$ which gives the conditional hazard of death at time t given $W = w$. No observation is made on W when $W > T$. The problem is simplified and reduced to one of time dependent covariables by making the assumption $\lambda_1(t \mid w) = \lambda_1(t)\, e^{\beta}$, $t > w$, where β now measures the transplant effect. Additional covariates can be added as required. This general approach of using time-dependent covariables to study the regression of one time variable on another is promising. Study of the practical value of a joint analysis of data on several failure variables as compared to separate analyses using time-dependent covariables would be interesting.

A second type of multivariate failure time data arises when individuals experiencing a failure recover or are repaired and then go on to be at risk for a subsequent failure. Thus each individual corresponds to a point process of failure information. Examples of this type would be the times between seizures in epileptic patients, or the times between breakdowns of an automobile. One approach to modelling such data would be to presume that corresponding to each individual is a renewal process with independent interfailure times and specify a model of type (1) or (2) to account for differences between subjects. This approach would usually be suspect since one would expect correlation between failure times for a given subject. Such correlations would arise, for example, if failure weakened the subject and increased the hazard for the next failure or if the model did not include some covariate that was related to failure time even if, given that covariate, the interfailure times were independent. An independent interfailure time model may thereby give rise to analyses that are unduly dominated by a few individuals with frequent failures. An alternative approach would be to introduce suitable time-dependent covariates that would allow for such correlations, though more work is needed in this direction. When only time-dependent covariates are under study, regression coefficients can be estimated from a model of the type (1) or (2) for each failure time on a study subject while permitting each study subject to define a stratum. Such a procedure should very well allow for intra-subject correlations. Whittemore and Keller (1978) use this method to analyze asthma attack data in relation to air quality measurements. Their analysis was based on a discrete logistic model that is closely related to (1).

A similar type of data arises when a subject is observed to pass through several states in the progress of the study. For example, in a clinical trial, leukemia patients may be followed through remission to relapse to death or possibly to repeated remissions. One way to analyze such data would be to specify a stochastic process model for each subject. Recent work in this area has been reported by Lagakos, Sommer and Zelen (1978) who assume that the process is of the Markov renewal type.

6. The Analysis of Case-Control Studies

In some situations, sampling schemes other than prospective are more practical in terms of cost and study duration. These considerations have rendered retrospective or case-control sampling a primary tool for epidemiologic research into factors associated with disease incidence. The case-control approach involves the selection of all, or a random sample of, cases of a certain disease or diseases that occur in a defined population during a certain time period. Retrospectively obtained exposure histories on such cases are then compared to those from a suitable selected disease-free or control sample in order to assess the effects of exposure variables on disease incidence. An important issue in the conduct of case-control studies concerns the quality of retrospectively obtained exposure data.

As in previous sections the population under study may be characterized by cause-specific hazard functions $\lambda_j[t, Z(t)]$. For example $j \epsilon (1, \ldots, m)$ may denote for an individual the first of m study diseases to occur, t may denote the corresponding age at disease incidence, while $Z(t)$ describes exposure and other variables up to age T. The prospective study permits direct estimation of $\lambda_j[t, Z(t)]$ based on samples from the probability distributions $P[t, j \,|\, Z(t)]$ or $P[t, j, Z(t)]$, where 'P' is used generically to denote probability of probability density functions. The case-control study, on the other hand, involves sampling from $P[Z(t) \,|\, (t, j)]$ or, for controls, from $P[Z(t) \,|\, T > t]$.

Case-control data do not permit estimation of the whole of the hazard function $\lambda_j[t, z(t)]$, $j = 1, \ldots, m$. The ratio of hazard functions corresponding to two covariate (exposure) functions can however be estimated: Let $Z_0(t)$ represent some standard covariate function. Simple calculations give

$$P[Z(t) \,|\, (t, j)]\; P(t, j) = \lambda_j[t, Z(t)]\; P[Z(t) \,|\, T > t]\; P(T > t)$$

so that

$$\frac{\lambda_j[t, Z(t)]}{\lambda_j[t, Z_0(t)]} = \frac{P[Z(t) \,|\, (t, j)] / P[Z(t) \,|\, T > t]}{P[Z_0(t) \,|\, (t, j)] / P[Z_0(t) \,|\, T > t]} \,. \tag{17}$$

The quantities on the right side of (17) can all be estimated in a case-control study so the left side can also be estimated. The left side is the relative risk for disease j at time t for an individual with 'exposure' function $Z(t)$ relative to an individual with standard exposure $Z_0(t)$. Under the proportional hazards model (16) the left side of (17) reduces to $\exp[\mathbf{z}(t) - \mathbf{z}_0(t)]\beta_j$.

Much of the literature on estimation from case-control studies has utilized data on whether or not disease occurs in a specified time period rather than the actual times (ages) of incidence. The analogue of (17) then specifies an 'odds-ratio' equality between prospective and retrospective sampling (e.g., Cornfield 1951). Discrete disease incidence models of the type mentioned in Section 2.3.1 may be used to estimate such odds ratios (e.g., Anderson 1972). The remainder of this section discusses the use of times of disease incidence or censoring to estimate relative risks.

Consider a case-control study in which each case as it occurs is matched to one or more controls that are disease-free at the same time (age). Let $Z_1(t)$ be the covariate function for an individual that develops disease j at time t. Let $Z_2(t), \ldots, Z_n(t)$ be a covariate function at time t for the corresponding controls. Given the set $\{Z_l(t),\ l = 1, \ldots, n\}$ of exposure functions, the probability that $Z_1(t)$ corresponds to the case is (Prentice and Breslow 1978)

$$P[Z_1(t)\,|\,(t,j)] \prod_{r=2}^{n} P[Z_r(t)\,|\,T > t] \Big/ \sum_{l=1}^{n} \left\{P[Z_l(t)\,|\,(t,j)] \prod_{r \neq l} P[Z_r(t)\,|\,T > t]\right\}$$
$$= \lambda_j\,[t, Z_1(t)] \Big/ \sum_{l=1}^{n} \lambda_j[t, Z_l(t)]$$

which under (16) reduces to

$$e^{z_1(l)} \Big/ \sum_{l \in R(t)} e^{z_l(t)\beta_j}, \tag{18}$$

where $R(t)$ consists of the case and matched controls at time t.

A conditional likelihood function for $\beta_1, \ldots, \beta_m$ is then given by

$$\prod_{j=1}^{m} \prod_{i=1}^{k_j} \left[e^{z_1(t_{ji})\beta_j} \Big/ \sum_{l \in R(t_{ji})} e^{z_l(t_{ji})\beta_j}\right] \tag{19}$$

where $t_{j1}, t_{j2} \ldots$ represent the incidence times for disease j. Prentice and Breslow give several generalizations of (19). Note that (16) may be relaxed to permit the λ_{0j} functions to differ among matched sets and still yield (19). This means that (19) should provide a thorough accommodation of 'cohort effects' if controls are selected concurrently. Note also the formal similarity between (19) and (5). In (19), however, $R(t)$ refers only to the set of individuals matched at time t while in (5) $R(t)$ consists of all sample subjects known to be at risk at or beyond time t.

Additional methodological work that would adapt (16) to case-control studies would appear to be quite worthwhile. Such work could, for example, avoid the matched aspect of the design mentioned above and consider simple random sampling from the case and control populations at each t. Gains in efficiency and a fuller utilization of the whole covariate function $Z(t)$, rather than a summary vector $\mathbf{z}(t)$, may be anticipated.

Some other sampling schemes that have arisen in our consulting activities include (*i*) a prospective random sample along with additional cases from a larger sample, (*ii*) a prospective sample for times to infection among 'long term survivors' along with data on infection cases among the short term survivors and (*iii*) cross-sectional prevalence sampling. It would be of interest to adapt the hazard function models discussed above to these and other sampling situations.

Acknowledgments

This work was supported by Grants CA-15704, CA-21193 and GM-24472 and Contract N01-ES-8-2125 awarded by the National Institute of Health, DHEW, and by a grant from the National Research Council of Canada.

References

Anderson, J. A. (1972). Separate sample logistic discrimination. *Biometrika 59,* 19–35.

Breslow, N. E. (1970). A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship. *Biometrika 57,* 579–594.

Breslow, N. E. (1974). Covariance analysis of censored survival data. *Biometrics 30,* 89–99.

Breslow, N. E., Day, N. E., Tomatis, L. and Turusov, V. S. (1974). Associations between tumor types in a large-scale carcinogenesis study of CF-1 mice. *Journal of the National Cancer Institute 52,* 233–239.

Chiang, C. L. (1968). *Introduction to Stochastic Processes in Biostatistics.* Wiley, New York.

Cornfield, J. (1951). A method of estimating comparative rates from clinical data. Application to cancer of the lung, breast and cervix. *Journal of the National Cancer Institute 11,* 1269–1275.

Cox, D. R. (1959). The analysis of exponentially distributed lifetimes with two types of failure. *Journal of the Royal Statistical Society, Series B 21,* 411–421.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B 34,* 187–220.

Cox, D. R. (1975). Partial likelihood. *Biometrika 62,* 269–276.

Crowley, J. and Hu, M. (1977). The covariance analysis of heart transplant data. *Journal of the American Statistical Association 72,* 27–36.

Efron, B. (1977). Efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association 72,* 557–565.

Farewell, V. T. and Prentice, R. L. (1977). A study of distributional shape in life testing. *Technometrics 18,* 69–76.

Gehan, E. (1965). A generalized Wilcoxon test for comparing arbitrarily singly censored samples. *Biometrika 52,* 203–223.

Hartley, H. O. and Sielken, R. L. (1977). Estimation of "safe doses" in carcinogenesis experiments (with discussion). *Biometrics 33,* 1–30.

Holford, T. R. (1976). Life tables with concomitant information. *Biometrics 32,* 587–598.

Holt, J. D. (1978). Competing risk analysis with special reference to matched pair experiments. *Biometrika 65,* 159–165.

Hampel, F. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association 69,* 383–393.

Jurečková (1971). Nonparametric estimate of regression coefficients. *Annals of Mathematical Statistics 42,* 1328–1338.

Kalbfleisch, J. D. (1974). Some efficiency calculations for survival distributions. *Biometrika 61,* 31–38.

Kalbfleisch, J. D. and Prentice, R. L. (1973). Marginal likelihood based on Cox's regression and life model. *Biometrika 60,* 267–278.

Kalbfleisch, J. D. and Prentice, R. L. (1979). *The Statistical Analysis of Failure Time Data.* In press, Wiley, New York.

Kaplan, E. L. and Meier, P. (1958). Non-parametric estimation from incomplete observations. *Journal of the American Statistical Association 53,* 457–481.

Koziol, J. A. and Reid, N. (1977). On multiple comparisons among K samples subject to unequal patterns of censorship. *Communications in Statistics A6 (12),* 1149–1164.

Lagakos, S. W., Sommer, C. J., and Zelen, M. (1978). Semi-Markov models for partially censored data. *Biometrika 65,* 311–318.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports 50,* 163–170.

Oakes, D. (1977). The asymptotic information in censored survival data. *Biometrika 64,* 441–448.

Peto, R., and Peto, J. (1972). Asymptotically efficient rank invariant test procedures (with discussion). *Journal of the Royal Statistical Society, Series A 135,* 185–206.

Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K., Peto, J., and Smith, P. G. (1977). Design and analysis of randomized clinical trials requiring prolonged observation on each patient. Part 2. Analysis and examples. *British Journal of Cancer 35,* 1–39.

Prentice, R. L. (1978). Linear rank test with right censored data. *Biometrika 65,* 167–179.

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

Prentice, R. L. and Gloeckler, L. A. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics 34,* 57–67.

Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. T., and Breslow, N. (1978). The analysis of failure time data in the presence of competing risks. *Biometrics 34,* 541–554.

Thompson, W. A. (1977). On the treatment of grouped observations in life studies. *Biometrics 33,* 463–470.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Science 72,* 20–22.

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 No. 526, Department of Statistics, University of Wisconsin.

Whittemore, A. and Keller, J. B. (1978). Asthma and air pollution: A quantitative theory. To appear in Proceedings of SIMS Conference on Environmental Health, SIAM.

Wong, O. (1977). A competing-risk model based on the life table procedure in epidemiologic studies. *International Journal of Epidemiology 6,* 153–159.