American Journal of Epidemiology

Vol. 145, No. 1
Printed in U.S.A.

Time-to-Event Analysis of Longitudinal Follow-up of a Survey: Choice of the Time-scale

Edward L. Korn,[1] Barry I. Graubard,[2] and Douglas Midthune[3]

Following individuals sampled in a large-scale health survey for the development of diseases and/or death offers the opportunity to assess the prognostic significance of various risk factors. The proportional hazards regression model, which allows for the control of covariates, is frequently used for the analysis of such data. The authors discuss the appropriate time-scale for such regression models, and they recommend that age rather than time since the baseline survey (time-on-study) be used. Additionally, with age as the time-scale, control for calendar-period and/or birth cohort effects can be achieved by stratifying the model on birth cohort. Because, as discussed by the authors, many published analyses have used regression models with time-on-study as the time-scale, it is important to assess the magnitude of the error incurred from this type of incorrect modeling. The authors provide simple conditions for when incorrect use of time-on-study as the time-scale will nevertheless yield approximately unbiased proportional hazards regression coefficients. Examples are given using data from the first National Health and Nutrition Examination Survey (NHANES I) Epidemiologic Followup Study. Additional issues concerning the analysis of longitudinal follow-up of survey data are briefly discussed. *Am J Epidemiol* 1997;145:72–80

cohort studies; epidemiologic methods; proportional hazards models; statistics; survey methods; survival analysis

Large-scale health surveys offer the ability to minimize selection bias in the estimation of associations between variables in a sampled target population. When interest focuses on the association of a risk factor and the development of a disease, however, the inference from a single cross-sectional survey is somewhat limited. One can examine the association between a risk factor and the *prevalence* of a disease, but this may not be a good estimate of the desired association because individuals with the disease may change their behavior, e.g., people with lung cancer who quit smoking. To avoid this bias, one can ask sampled individuals to describe their earlier (pre-disease) risk factors, but their answers may be subject to considerable recall bias. In addition, if the disease is potentially fatal, then patients who die from the disease before the survey will not be sampled, potentially leading to more bias in the estimation of disease/risk factor associations.

With longitudinal follow-up that records the development of various diseases, many of the aforementioned biases are minimized. Additionally, one can study the association of risk factors and mortality. As an example, consider the first National Health and Nutrition Examination Survey (NHANES I) Epidemiologic Followup Study (NHEFS). NHANES I was a multistage, national probability sample of the US civilian noninstitutionalized population (1, 2). The NHEFS is an ongoing series of follow-up surveys of the individuals sampled in NHANES I who were aged 25–74 years at the baseline examination (3). Table 1 displays the types of risk factors and outcomes that were analyzed for papers published in 1993 using NHEFS; the list of publications was abstracted from a computer file provided by the National Center for Health Statistics ("The National Health and Nutrition Examination Surveys, A Selective Bibliography, 1980–93," June 1994).

The purpose of this paper is to discuss appropriate methods of analysis of time-to-event data such as those acquired in NHEFS, and, in particular, the choice of the time-scale in a proportional hazards regression. Statistical issues in the analysis of NHEFS

Received for publication January 30, 1996, and accepted for publication July 25, 1996.

Abbreviations: NHANES I, first National Health and Nutrition Examination Survey; NHEFS, NHANES I Epidemiologic Followup Study.

[1] Biometric Research Branch, National Cancer Institute, Bethesda, MD.

[2] Biometry Branch, National Cancer Institute, Bethesda, MD.

[3] Information Management Services, Silver Spring, MD.

Reprint requests to Dr. Edward L. Korn, Biometric Research Branch EPN-739, National Cancer Institute, Bethesda, MD 20892.

TABLE 1. Examples of time-to-event analyses published in 1993 using the first National Health and Nutrition Examination Survey (NHANES I) Epidemiologic Followup Study

Outcome and primary risk factor(s)	Reference no(s).
Mortality	
Pulmonary impairment	4*
Weight or weight loss	5*, 6*
Dietary diversity	7*
Vitamin and mineral supplements	8*
Health insurance	9*
Coronary heart disease	
Alcohol consumption	10*, 11†
No. of pregnancies	12*
Weight or weight loss	13*
Mortality and coronary heart disease	
White blood cell count	14*
Stroke	
Hormone use	15*
Hip fracture	
Dietary calcium	16*
Multiple myeloma	
Antigenic conditions	17‡

* Analysis used a proportional hazards regression model with time-on-study as the time-scale and baseline age as a covariate.
† Analysis used a parametric regression model with age as the time-scale.
‡ Analysis used a proportional hazards regression model with time-on-study as the time-scale with the cohort matched on age.

data have recently been addressed in a National Center for Health Statistics publication by Ingram and Makuc (18), who in part compared different regression models and recommended that the proportional hazards regression model be used. In the regression models they presented, the time-scale chosen was follow-up time (time-on-study), and baseline age was included as a covariate. With two exceptions (11, 17), all of the publications presented in table 1 have used this recommended model. We describe this model in detail in the next section and argue that a more appropriate proportional hazards regression model would use age as the time-scale, with possible stratification on birth cohort. Simple conditions are given for when using an incorrectly specified model with time-on-study as the time-scale will nevertheless yield approximately unbiased estimates of the regression coefficient for a risk factor. In the third section of this paper, we present some examples using data from the NHEFS with the analyses done with age and time-on-study time-scales to show the differing results. We end with a brief discussion of some related issues in analyzing follow-up of a survey, including the incorporation of the sample design into the analysis, the problem of pre-event conditions that may affect the risk factors, and the use of longitudinally collected risk factor information during the follow-up period.

PROPORTIONAL HAZARDS REGRESSION MODELS

There are three general types of models used for analyzing time-to-event data: parametric, nonparametric, and semiparametric. For a parametric model, the distribution of the times to events given a set of risk factors and covariates is completely specified except for a (finite) set of unknown parameters, which are estimated from the data. For a nonparametric analysis, these distributions are estimated directly from the data with essentially no model assumptions, e.g., by using Kaplan-Meier plots for each combination of risk factor and covariate values. For a semiparametric analysis, these distributions are modeled as a function of an unspecified baseline distribution and a set of unknown parameters. One such semiparametric model, which is the focus of this paper, is the proportional hazards regression model of Cox (19). To define this model, let Y be the time to the event. The hazard function of Y (force of mortality), defined by

$$\lambda_Y(y) = \lim_{\Delta \to 0} \frac{1}{\Delta} P(Y \epsilon [y, y+\Delta) \mid Y \geq y),$$

is the instantaneous rate of an event occurring at time y given that it has not occurred before time y. Let $z = (z_1, z_2, \ldots, z_K)$ be a vector of baseline risk factors and/or covariates, $\beta = (\beta_1, \beta_2, \ldots, \beta_K)$ be a vector of unknown regression coefficients, $\beta' z = \beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_K z_K$, and assume that the individuals under study can be categorized into a stratum j, $j = 1, \ldots, J$. One possible proportional hazards regression model for Y is given by

$$\lambda_Y(y \mid j, z) = \lambda_{0jY}(y) \exp\{\beta' z\}, \qquad (1)$$

where $\lambda_{0jY}(.)$ are unspecified baseline hazard functions that allow the hazard to be different depending on stratum membership.

With data from longitudinal follow-up of a survey, there are different ways one could apply a proportional hazards model. For example, consider an individual aged 60 years in 1980 who has been followed since being sampled in the baseline survey in 1970. His hazard could depend on his age (60 years), his time-on-study (10 years), which is equivalent to the calendar period (1980), his birth cohort (1920), as well as the risk factors and other covariates. Because a person's age plus his birth cohort equals the calendar period, there are well-known identifiability problems associated with untangling the effect of age, cohort, and calendar period on the hazard function (20).

Fortunately, for research questions directed at the association of risk factors with time-to-events, the

identifiability problems associated with age-cohort-period models can be avoided. Consider the model

$$\lambda_A(a|b_0,z) = \lambda_0(a,b_0)\exp\{\beta' z\}, \quad (2)$$

where $A = a$ is the age of the individual during the follow-up period, and b_0 is the birth cohort of the individual. The baseline hazard in model 2 is specified in terms of age and cohort, but a parametrization in terms of age and calendar period would yield equivalent results for the regression coefficients. An analysis using model 2 can proceed by grouping the ages and cohorts into intervals (21). A preferable analysis is of the form of model 1 and groups only on cohort, viz.,

$$\lambda_A(a|b_0 \epsilon B_j,z) = \lambda_{0jA}(a)\exp\{\beta' z\}, \quad (3)$$

where $B_1,B_2,\ldots,B_J$ are birth cohort intervals, e.g., 1906–1910, 1911–1915, etc. With cohort intervals of small width, model 3 controls for period effects as well as age and cohort effects. Model 3 is the model we generally recommend. Two alternative models that will yield similar results to using model 3 are possible. One of these alternative models uses age as the time-scale but stratifies on time-dependent strata consisting of 5-year calendar periods (22), while the other uses time-on-study as the time-scale but stratifies on birth cohort (23); see also Reichman ("Cox Proportional Hazards Survival Analysis with Multiple Time Scales," unpublished Master of Arts thesis, University of Maryland, College Park, Maryland, 1991).

We now present two proportional hazards regression models that do not use stratification. The first model is the commonly used one we mentioned in the introduction. Let T be the time-on-study and a_0 the age of the baseline survey of an individual. Using time-on-study as the time-scale we have

$$\lambda_T(t|a_0,z) = \lambda_{0T}(t)\exp\{\xi a_0 + \gamma' z\}, \quad (4)$$

where $\gamma = (\gamma_1,\gamma_2,\ldots,\gamma_K)$. A direct simplification of model 3 uses age as the time-scale but without stratification:

$$\lambda_A(a|a_0,z) = \lambda_{0A}(a)\exp\{\beta' z\}. \quad (5)$$

This model would be appropriate if there were no concerns about cohort or period effects.

We believe that models 3 or 5 are preferable to model 4 for analyzing outcomes such as those in table 1. This is because for these outcomes we would expect the hazard to change more as a function of age than as a function of time-on-study. For example, we would expect the hazards of persons aged 50 years and 60 years who have both been on study for 10 years to differ more than the hazards of two persons aged 55 years, one of whom has been on study 5 years and the other 15 years. Because the great flexibility of the proportional hazards model is due to not having to specify the form of $\lambda_0(.)$, it is best to use this function to model the variable that is expected to have the largest effect on the hazard (in this case, age); see also reference 22. In other biomedical applications, $\lambda_0(.)$ would be typically modeled as a function of variables other than age. For example, in a randomized clinical trial or in a natural history study of a disease, time since randomization or diagnosis would be used; see Andersen et al. (24) for additional examples. With follow-up of a healthy population, however, we believe age will be the most appropriate time-scale for most outcomes. This assumes that the hazard function given age and other covariates does not change with calendar time. In some situations, this assumption will not be valid because of calendar effects due to advances in medical management. In these cases, it is important to stratify on birth cohort (model 3). In fact, provided that the cohort stratification is coarse enough so that there are sufficient numbers of individuals in each cohort, then practically no precision is lost in unnecessarily stratifying by birth cohort when estimating β. However, fine stratification can lead to bias in the estimators of the variance of the estimated β. For any given analysis, this can be checked using computer simulations, as we will demonstrate empirically in the examples given below. We do not address the possibility of combining multiple time-scales into one time-scale (25, 26), because there would appear to be no advantage of this type of model over model 3.

Suppose that age is the appropriate time-scale and model 5 is consistent with the data. What are the implications of inappropriately using model 4 for analysis? We show in the Appendix that if it happens that the baseline hazard $\lambda_{0A}(a) = c\exp\{\psi a\}$ for some c and ψ, then the γ's estimated using model 4 will be estimating the correct β's from model 5. Even if the baseline hazard function were only approximately exponential, we would expect only a small bias in using model 4. If the baseline hazard function is not anything like an exponential function, there is another condition that will ensure that the estimated γ's will approximately estimate the β's. This condition is that the z and the baseline ages a_0 are statistically independent. The existence of such conditions is good because it suggests that previous analyses performed using model 4 may not be seriously in error. This is because the hazards of many outcomes would be expected to increase rapidly as a function of age, roughly approximating an exponential distribution. In fact, the Gompertz distribution (27), which has been historically used to model mortality, has precisely an exponential hazard function. We give an example in the

next section in which the hazard function is not an exponential function.

A computer program that can be used for analysis using models 3 or 5 which also allows for the possibility of time-dependent covariates is available from the authors. The program has the option of fully incorporating the sampling design (assuming with-replacement sampling at the first stage of sampling). In particular, sample-weighted estimators are calculated, and standard errors are estimated accounting for the sample clustering and stratification in the sample design by using Taylor-series linearization (28). With no time-dependent covariates, the program can also produce estimates of the cumulative hazard for any pattern of covariates, but these estimates may be unreliable because of the small numbers of subjects at risk at the youngest ages.

EXAMPLES

In this section, we present examples to demonstrate the effect of the choice of time-scale on the estimation of a proportional hazards regression coefficient. The examples, which use the 1987 follow-up data (3) on the women in the NHEFS, are meant to be illustrative of points discussed in the last section rather than to be substantive analyses. We consider two outcomes: all-cause mortality and the removal of the ovaries (last ovary, if removed at different times). The outcomes were chosen because the cumulative hazard function (and therefore hazard function) for mortality as a function of age looks exponential (figure 1), whereas for age at ovary removal it does not (figure 2). Because ovary status was not asked at the baseline survey in 1971–1975, we used the method described in the foot-

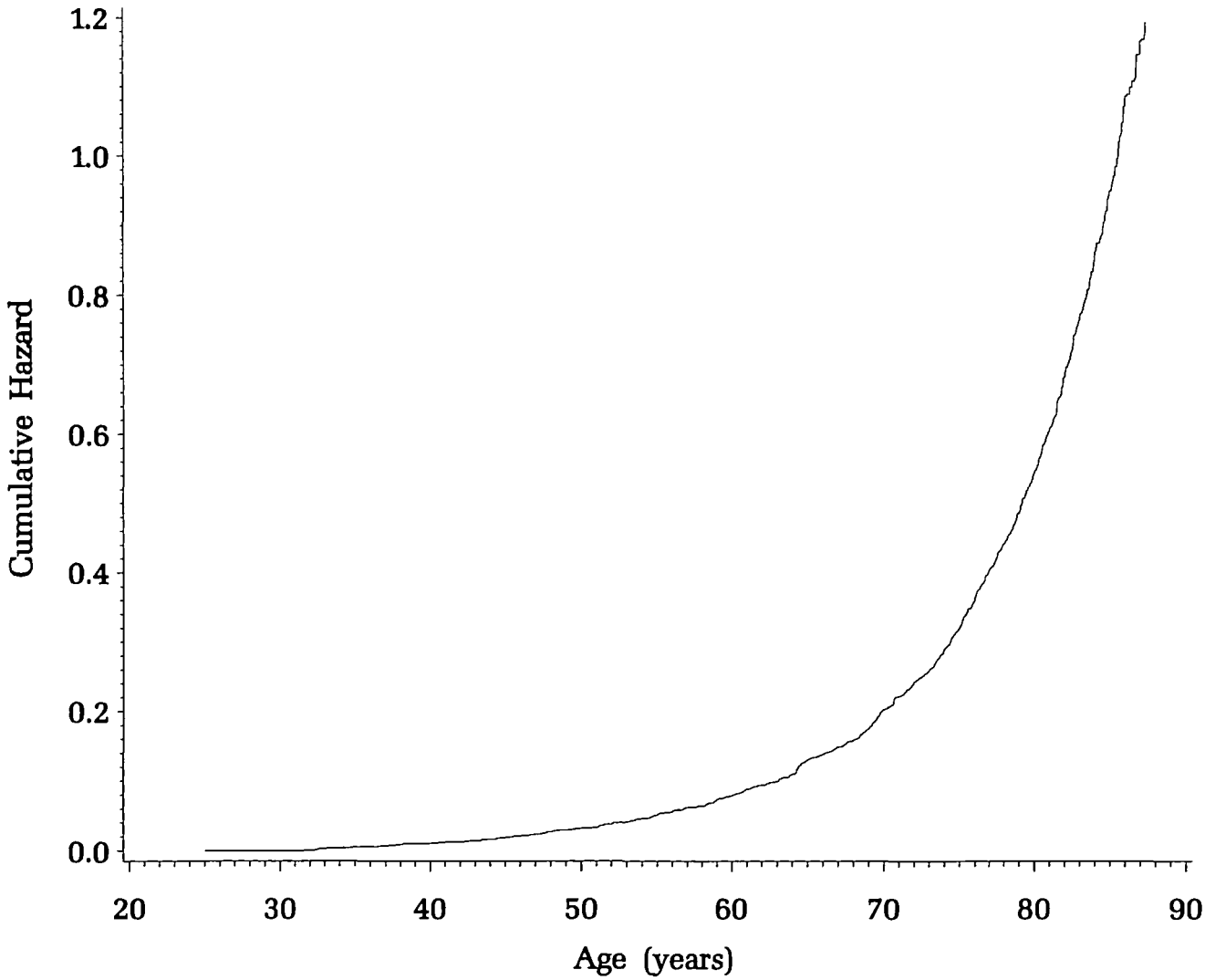

FIGURE 1. Cumulative hazard for mortality as a function of age (age ≥25 years) for women being followed in the NHANES I Epidemiologic Followup Study.

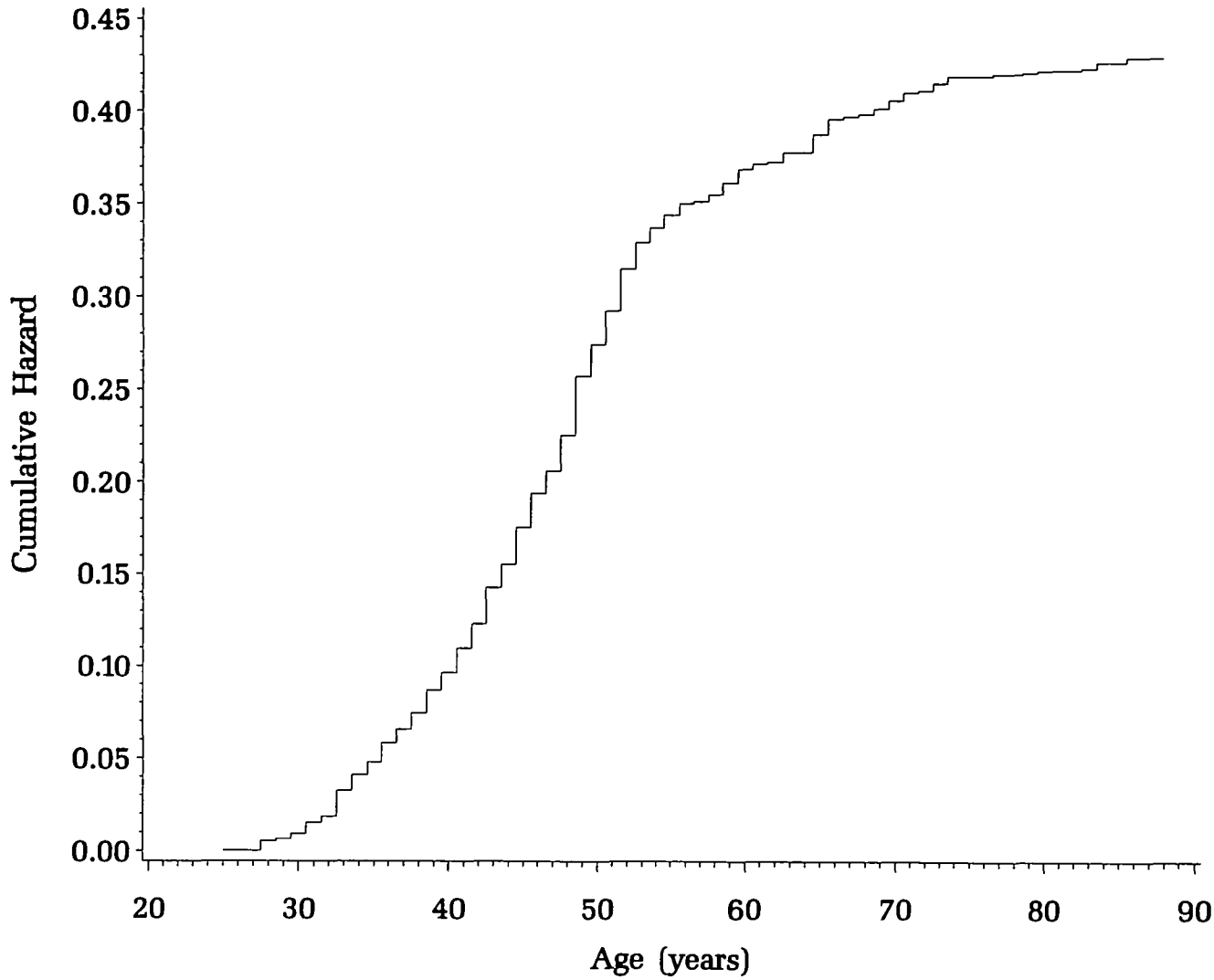

FIGURE 2. Cumulative hazard for ovary removal as a function of age (age ≥25 years) for women being followed in the NHANES I Epidemiologic Followup Study.

note to table 4 to approximate what the data would have looked like if they had been collected at baseline. Additionally, for women who died without ovary removal during the follow-up period, their time-to-ovary removal was considered censored at the time of death; an alternative analysis of cumulative incidence (29) would assign an arbitrarily large value for time-to-ovary removal for these women.

Three baseline binary risk factors are considered: 1) whether the woman lived in an urban or rural area (as derived from the 1960 census), 2) whether the woman was a smoker (current or former) or a nonsmoker, and 3) whether the woman's family income was ≤$4,000 or >$4,000. These risk factors were chosen because they represent a range of disparities on baseline age distributions. For the urban/rural variable, the age distributions are well matched; for the smoker/nonsmoker variable, they are moderately different; and for the family income variable, they are very different; see table 2 for selected percentiles for the subset of women analyzed for ovary removal.

The analyses presented in this section fully utilize the sample design. The sample weights and standard errors were calculated as recommended by Ingram and Makuc (18).

Table 3 presents the proportional hazards regression coefficients for the three risk factors for mortality calculated using three different proportional hazards models. The first (model 3) uses age as the time-scale and stratification on 5-year birth cohorts. The second (model 5) has age as the time-scale but with no stratification. The third (model 4), which we do not rec-

TABLE 2. Baseline age distributions for three risk factors in the NHANES I Epidemiologic Followup Study for a subset of women analyzed for ovary removal (see text)

Risk factor	Age (years) by weighted percentile		
	25th	50th	75th
Urban/rural			
Urban (n = 2,272)	33.6	44.2	55.9
Rural (n = 3,710)	33.7	45.1	57.1
Smoker/nonsmoker			
Smoker (n = 2,495)	32.6	42.2	53.2
Nonsmoker (n = 3,228)	35.5	47.6	60.2
Family income			
≤$4,000 ($n$ = 1,134)	43.6	58.2	67.1
>$4,000 ($n$ = 4,634)	32.7	42.7	53.9

ommend, uses time-on-study as the time-scale with baseline age as a covariate. The first two models yield almost identical results, suggesting that there are no large cohort or period effects interacting with the relative hazard. As expected, because of the exponential-like cumulative hazard (figure 1), the results of the second two models are almost identical. Table 4 presents the results for ovary removal. Again the results for the first two models are almost identical. The comparison of the second two models depends on the risk factor. For the urban/rural variable, the results are the same; and for the smoker/nonsmoker variable, the results are close. For the family income variable, the results are quite different, 0.29 ± 0.19 versus 0.06 ± 0.20. Based on a jackknife, the standard error of the difference of these estimators, 0.23 (= 0.29 − 0.06), is ±0.04, demonstrating that the observed difference is not due to sampling variability. These findings are consistent with the theoretical considerations given in the previous section. Note that in all cases the stratification on 5-year birth cohorts did not increase the standard errors. Based on computer simulations (not shown), these standard errors are estimated accurately.

TABLE 3. Proportional hazards regression coefficients (± standard error) for three risk factors (considered one at a time) for mortality among women in the NHANES I Epidemiologic Followup Study calculated by three methods

Risk factor	Method		
	Age as the time-scale with stratification on birth cohort (5-year intervals)	Age as the time-scale	Time-on-study as the time-scale with baseline age as a covariate
Urban vs. rural (n = 8,183)	0.05 ± 0.08	0.05 ± 0.08	0.05 ± 0.08
Smoker vs. nonsmoker (n = 7,626)	0.40 ± 0.11	0.40 ± 0.11	0.38 ± 0.11
Family income (≤$4,000 vs. >$4,000) (n = 7,878)	0.21 ± 0.08	0.20 ± 0.08	0.23 ± 0.08

TABLE 4. Proportional hazards regression coefficients (± standard error) for three risk factors (considered one at a time) for risk of ovary removal* among women in the NHANES I Epidemiologic Followup Study calculated by three methods

Risk factor	Method		
	Age as the time-scale with stratification on birth cohort (5-year intervals)	Age as the time-scale	Time-on-study as the time-scale with baseline age as a covariate
Urban vs. rural (n = 5,982)	−0.08 ± 0.11	−0.09 ± 0.11	−0.09 ± 0.11
Smoker vs. nonsmoker (n = 5,723)	0.06 ± 0.09	0.06 ± 0.09	0.09 ± 0.09
Family income (≤$4,000 vs. >$4,000) (n = 5,768)	0.30 ± 0.19	0.29 ± 0.19	0.06 ± 0.20

* Women were asked in the 1986–1987 follow-up about their ovary status, and if removed, their age at the time of removal. If the age at ovary removal for a woman was before her age at the baseline survey, her data were not used in the analysis. For women who were alive at the 1982–1984 follow-up, but unable to be interviewed in 1986–1987 (e.g., because they had died), a proxy response was used for time of ovary removal. Data from women who died before the 1982–1984 follow-up were not used in the analysis because no proxy responses were available.

Based on a priori biologic considerations, we believe that age is the most appropriate time-scale for these analyses. However, there is also some evidence provided by the data on this point. We performed an analysis of ovary removal with calendar time as the time-scale with the following independent variables: family income (≤$4,000 vs. >$4,000), baseline age, and the square of baseline age. The coefficient for family income was estimated to be 0.31 ± 0.20, very close to the value obtained when using age as the time-scale (last row of table 4). This suggests that the simpler model with calendar-time as the time-scale and age at baseline as a covariate mismodels the hazard function.

DISCUSSION

Besides the choice of time-scale, there are other issues that need to be addressed when analyzing longitudinal follow-up of a survey. One is how to incorporate aspects of the sampling design into the analysis. Our general recommendations are given elsewhere (30); specific recommendations for NHEFS are given by Ingram and Makuc (18). Standard errors estimated

incorporating the clustering of the sample design may be quite variable when there are few primary sampling units (PSU's) in the design. However, since there are sufficient PSU's in the NHEFS design to estimate reliably the standard errors (75 non-certainty PSU's sampled from 25 strata, and 10 certainty PSU's treated for variance estimation as 30 pseudo-PSU's sampled from 10 pseudo-strata), we recommend using variance estimation that accounts for the sample clustering and stratification for this survey. The use of the sample weights in a weighted analysis yields approximately unbiased estimates of population quantities (which an unweighted analysis may not), but at the cost of increased variability of the estimators. Whether or not to incorporate the sample weights and clustering into the analysis needs to be considered survey by survey, and perhaps even analysis by analysis (18). Somewhat surprisingly, all but one (9) of the proportional hazards regression analyses in table 1 used no aspect of the sampling design of the NHEFS; two articles (10, 14) used the sample design in subsidiary logistic regression analyses. For comparison purposes, we repeated the key analyses of the last section (the last row of table 4) ignoring the sampling design. The regression coefficients for family income for the hazard of ovary removal are estimated to be -0.18 ± 0.13 with age as the time-scale, and -0.37 ± 0.14 with time-on-study as the time-scale and baseline age as a covariate.

A second issue is how to handle pre-event conditions that may affect the risk factor. For example, suppose the outcome event is death from cancer and the risk factor is smoking. It is not hard to imagine that an individual with lung cancer at the baseline survey may have reduced her smoking prior to the survey, leading to an obvious bias. One approach is to eliminate from the analysis individuals with preexisting conditions that might affect the risk factor. This approach is itself not without bias, because the individuals who are potentially the most susceptible to the harm of the risk factor are being removed from the analysis. Another approach is to stratify by the condition in the analysis. For example, in the analysis of health insurance and mortality (9), the presence or absence of morbidity at the baseline survey was included as a covariate. This approach is also potentially biased since one is stratifying on a variable that is on the causal pathway between the risk factor and the disease (31).

Sometimes there may be concerns that there are pre-event conditions that may affect the risk factor but be preclinical at the baseline survey. One approach is to eliminate follow-up data within a specified time interval from the baseline survey from the analysis. For a proportional hazards regression model with time-on-study as the time-scale, this is equivalent to eliminating from the analysis individuals who have the event within the specified time interval. An example is given by the study of weight loss and mortality (6), where additional analyses were done in which deaths that occurred in the first 5 or 8 years of follow-up were excluded. With age as the time-scale, eliminating initial events is not the same as eliminating initial follow-up, with the latter being the correct approach.

Rather than eliminating the data from the early follow-up period, one can attempt to model the effect of pre-event conditions on the risk factor. For example, if one believes that a preclinical pre-event condition could affect the risk factor for up to 2 years, then one might consider the following model for a binary risk factor R:

$$\lambda_A(a|R,a_0) = \lambda_{0A}(a)\exp\{\beta R - \zeta R(2 - a + a_0)_+/2\}, \quad (6)$$

where $(x)_+$ equals x if $x > 0$ or 0 otherwise. In model 6, the log-relative hazard due to R is modeled to be reduced by ζ at the beginning of the follow-up period, $\zeta/2$ at one year of follow-up, and zero at ≥ 2 years of follow-up. If the estimated ζ is small, then this suggests that there is no evidence that a pre-event condition is affecting the risk factor.

The last issue we address is how to utilize longitudinal risk factor information when it is available. For example, in the NHEFS, weight and blood pressure were obtained at the 1982–1984 follow-up as well as the baseline survey. We first note that with a single measurement available (typically from the baseline survey), an implicit assumption in using the regression models previously described is that the single observed risk factor is useful in characterizing the lifetime risk of that individual. Because an individual's risk factor status may change with age, this assumption is open to question. In this regards, model 3 offers an advantage over model 5 in that individuals being directly compared for their risk factor status are being compared for their status obtained at the same general age, e.g., family income at ages 30–35 years.

When multiple measurements of a risk factor are obtained, there are additional analysis possibilities. If one is assessing the utility of a biologic marker of a disease for screening purposes, then carrying forward the last known value of the marker as a time-dependent covariate in the regression models would be appropriate. For the more typical uses of longitudinal follow-up, one should consider the hypothetical connections between the risk factor and the disease. For example, if there is thought to be a long latency period between the action of the risk factor and

the incidence of the disease, then one might utilize only the baseline risk information. Additionally, if one thought pre-event conditions might be affecting the risk factor as described above, then utilizing longitudinally collected risk information right up to the time of the event might not be recommended. A final consideration is that many risk factors have short-term temporal variation due to measurement error (e.g., a laboratory value), biologic variability (e.g., blood pressure), or their inherent nature (e.g., 24-hour dietary recall). A one-shot measurement at the baseline survey is thus subject to "error," which can attenuate associations between the risk factor and disease. Averaging risk factor values collected at different times will lessen this problem. At a minimum, one can use the multiple measurements to assess the extent of the attenuation.

In summary, the analysis of data collected from the longitudinal follow-up of a survey is not straightforward. The purpose of this brief discussion has been to raise some of the issues; we realize that we have not provided answers to all the questions raised. However these issues are settled, the use of a model that is consistent with the data is important. We recommend modeling the events using a proportional hazards model with age as the time-scale. This procedure is more meaningful and less restrictive than using time-on-study as the time-scale.

ACKNOWLEDGMENTS

The authors thank Drs. M. Gail, J. Lubin, E. Slud, and two anonymous referees for their helpful comments.

REFERENCES

1. Miller HW. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971–73. Hyattsville, MD: National Center for Health Statistics, 1973. (Vital and Health Statistics, Series 1: Programs and collection procedures, No. 10a) (DHEW publication no. (PHS) 79–1310).
2. Engel A, Murphy RS, Maurer K, et al. Plan and operation of the NHANES I Augmentation Survey of adults 25–74 years, United States, 1974–75. Hyattsville, MD: National Center for Health Statistics, 1978. (Vital and Health Statistics, Series 1: Programs and collection procedures, No. 14) (DHEW publication no. (PHS) 78–1314).
3. Cox CS, Rothwell ST, Madans JH, et al. Plan and operation of the NHANES I Epidemiologic Followup Study, 1987. Hyattsville, MD: National Center for Health Statistics, 1992. (Vital and Health Statistics, Series 1: Programs and collection procedures, No. 27) (DHHS publication no. (PHS) 92–1303)
4. Bang KM, Gergen PJ, Kramer R, et al. The effect of pulmonary impairment on all-cause mortality in a national cohort. Chest 1993;103:536–40.
5. Rumpel C, Harris TB, Madans J. Modification of the relationship between the Quetelet index and mortality by weight-loss history among older women. Ann Epidemiol 1993;3:343–50.
6. Pamuk ER, Williamson DF, Serdula MK, et al. Weight loss and subsequent death in a cohort of US adults. Ann Intern Med 1993;119:744–8.
7. Kant AK, Schatzkin A, Harris TB, et al. Dietary diversity and subsequent mortality in the first National Health and Nutrition Examination Survey Epidemiologic Follow-up Study. Am J Clin Nutr 1993;57:434–40.
8. Kim I, Williamson DF, Byers T, et al. Vitamin and mineral supplement use and mortality in a US cohort. Am J Public Health 1993;83:546–50.
9. Franks P, Clancy CM, Gold MR. Health insurance and mortality. Evidence from a national cohort. JAMA 1993;270: 737–41.
10. Garg R, Wagener DK, Madans JH. Alcohol consumption and risk of ischemic heart disease in women. Arch Intern Med 1993;153:1211–16.
11. Coate D. Moderate drinking and coronary heart disease mortality: evidence from NHANES I and the NHANES I Follow-up. Am J Public Health 1993;83:888–90.
12. Ness RB, Harris T, Cobb J, et al. Number of pregnancies and the subsequent risk of cardiovascular disease. N Engl J Med 1993;328:1528–33.
13. Harris TB, Ballard-Barbasch R, Madans J, et al. Overweight, weight loss, and risk of coronary heart disease in older women: the NHANES I Epidemiologic Follow-up Study. Am J Epidemiol 1993;137:1318–27.
14. Gillum RF, Ingram DD, Makuc DM. White blood cell count, coronary heart disease, and death: the NHANES I Epidemiologic Follow-up Study. Am Heart J 1993;125:855–63.
15. Finucane FF, Madans JH, Bush TL, et al. Decreased risk of stroke among postmenopausal hormone users. Results from a national cohort. Arch Intern Med 1993;153:73–9.
16. Looker AC, Harris TB, Madans JH, et al. Dietary calcium and hip fracture risk: the NHANES I Epidemiologic Follow-up Study. Osteoporosis Int 1993;3:177–84.
17. Bourguet CC, Logue EE. Antigenic stimulation and multiple myeloma. A prospective study. Cancer 1993;72:2148–54.
18. Ingram DD, Makuc DM. Statistical issues in analyzing the NHANES I Epidemiologic Followup Study. Hyattsville, MD: National Center for Health Statistics, 1994. (Vital and Health Statistics, Series 2: Data evaluation and methods research, No. 121) (DHHS publication no. (PHS) 94–1395).
19. Cox DR. Regression models and life tables (with discussion). J R Stat Soc [B] 1972;34:187–220.
20. Mason WM, Fienberg SE, eds. Cohort analysis in social research, beyond the identification problem. New York: Springer-Verlag, 1985.
21. Breslow NE, Day NE. Statistical methods in cancer research. Vol II. The design and analysis of cohort studies. Lyon: International Agency for Research on Cancer, 1987, chap 4.
22. Breslow NE, Lubin JH, Marek P, et al. Multiplicative models and cohort analysis. J Am Stat Assoc 1983;78:1–12.
23. Cnaan A, Ryan L. Survival analysis in natural history studies of disease. Stat Med 1989;8:1255–68.
24. Andersen PK, Borgan O, Gill RD, et al. Statistical models based on counting processes. New York: Springer-Verlag, 1993:675–82.
25. Farewell VT, Cox DR. A note on multiple time scales in life testing. Appl Stat 1979;28:73–5.
26. Oakes D. Multiple time scales in survival analysis. Lifetime Data Analysis 1995;1:7–18.
27. Read CB. Gompertz distribution. In: Kotz S, Johnson NL, eds. Encyclopedia of statistical sciences. Vol 3. New York: John Wiley, 1983:446.
28. Binder DA. Fitting Cox's proportional hazards models from survey data. Biometrika 1992;79:139–47.
29. Korn EL, Dorey FJ. Applications of crude incidence curves. Stat Med 1992;11:813–29.
30. Korn EL, Graubard BI. Analysis of large health surveys: accounting for the sampling design. J R Stat Soc [A] 1995; 158:263–95.

31. Breslow NE, Day NE. Statistical methods in cancer research. Vol 1. The analysis of case-control data. Lyon: International Agency for Research on Cancer, 1980:104–5.

32. Gail MH, Wieand S, Piantadosi S. Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. Biometrika 1984;71:431–44.

APPENDIX

In this appendix, we consider two conditions that will ensure that the γ's estimated using model 4 will approximately estimate the β's from model 5 when model 5 is, in fact, correct. The first condition is that $\lambda_{0A}(a) = c \exp\{\psi a\}$ for some c and ψ. Using the definition of a hazard function we have

$$\lambda_T(t|a_0,z) = \lim_{\Delta\to\infty}\frac{1}{\Delta}P(T\epsilon[t,t+\Delta)\,|\,T\geq t,a_0,z) = \lim_{\Delta\to\infty}\frac{1}{\Delta}P(A-a_0\epsilon[t,t+\Delta)\,|\,A-a_0\geq t,a_0,z)$$

$$= \lim_{\Delta\to 0}\frac{1}{\Delta}P(A\epsilon[a_0+t,a_0+t+\Delta)\,|\,A\geq a_0+t,a_0,z) = \lambda_A(a_0+t|a_0,z)$$

$$= \lambda_{0A}(a_0+t)\exp\{\beta' z\} \tag{7}$$

$$= c\exp\{\psi t\}\exp\{\psi a_0+\beta' z\}, \tag{8}$$

where we have used the special exponential form of $\lambda_{0A}(a)$ to derive expression 8 from expression 7. Model 8 is consistent with model 4 with $\lambda_{0T}(t) = c\exp\{\psi t\}$. Therefore, the estimates of the γ's using model 4 will correctly estimate the β's. Notice that this would not be true if the baseline age a_0 were not included as a covariate in model 4.

If $\lambda_{0A}(a)$ is not an exponential function, the second condition we consider is that the baseline ages a_0 are statistically independent of z. For example, if z is a binary risk factor, then this condition states that the distribution of baseline ages for those with and without the risk factor should be the same. To show heuristically why this condition works, rewrite model 4 as

$$\lambda_T(t|a_0,z) = [\lambda_{0T}(t)\exp\{\xi a_0\}]\exp\{\gamma' z\}. \tag{9}$$

By fitting model 9 to data consistent with model 5 (and therefore expression 7), one approximates $\lambda_{0A}(a_0+t)$ by $[\lambda_{0T}(t)\exp\{\xi a_0\}]$. If a_0 and z are independent, then any misfit of this approximation would not be expected to affect much how $\exp\{\gamma' z\}$ is approximating $\exp\{\beta' z\}$. To make this argument precise, and to show just how much bias there will be in estimating β with estimates of γ (it is not unbiased), requires asymptotic theory beyond the scope of this paper; see also Gail et al. (32). We examined this issue empirically in the examples given in the main body of the paper.