JOURNAL OF INSURANCE MEDICINE

J Insur Med 2000;32:217–225

ORIGINAL ARTICLE

An Analytic Method for Longitudinal Mortality Studies

David Strauss, PhD; Robert Shavelle, PhD; Michael J. DeVivo, DrPH; Steven Day

Our knowledge of mortality risks comes largely from longitudinal (cohort) studies. The most commonly used analytic tool is the Cox proportional hazards model for survival analysis. An alternative approach is a simple cross-sectional analysis of person-years. The key to the method is logistic regression, where the outcome variable is lived/died in the given year and the explanatory variables are age, sex, and other potential risk factors. This approach can be used to model any dichotomous outcome and has several important advantages over the more traditional survival analysis. As an example, we compare the two methods using a large data base of patients with spinal cord injury.

Address: Department of Statistics (Strauss, Shavelle, Day), University of California, Riverside, CA 92521-0138; Spain Rehabilitation Center (DeVivo), University of Alabama, Birmingham, AL 35294.

Correspondent: David Strauss, PhD, FASA.

Key words: Cox regression model, mortality, life expectancy, person-year, survival.

Received: August 12, 2000.

Accepted: September 1, 2000.

A central topic in insurance medicine is the mortality risk associated with serious medical conditions such as brain or spinal cord injuries, cerebral palsy, or cancer. Our knowledge about these risks comes primarily from research studies. The research literature is scattered across the various disciplines of application: rehabilitation medicine, neurology, cancer, and others. In some respects, this is unfortunate because the studies raise similar research questions and share common methodological issues.

In these studies, the predominant study design is longitudinal, subjects being followed until death or the end of the study period. Typical questions are

(1) What patient characteristics most affect mortality risk?
(2) How is mortality affected by time since injury (in the case of spinal cord or brain injuries) or time since diagnosis or stage of disease (in the case of cancer)?
(3) Has there been a secular trend (rates perhaps falling over time in response to improvements in treatment)?

What the medical director requires is information on the mortality risk for any specific combination of the above. For example, one might wish to know the risk under current conditions for a 20-year-old male who is 2 years postinjury and has a Frankel grade A (the most severe) spinal injury to the third cervical vertebra.

Any research study attempting to answer such questions uses (1) a longitudinal data base tracking a large number of patients and (2) an analytic method for separating and identifying the effects of various patient char-

Table 1. Some Major Data Sources for Longitudinal Mortality Studies

Cerebral palsy	
United States	California Developmental Disabilities Database[1–3]
Britain	Merseyside Database[4,5]
Canada	British Columbia data[6]
Vegetative state	
United States	California Developmental Disabilities Database[7,8]
Israel	Sazbon et al[9,10]
Japan	Higashi et al[11–13]
Spinal cord injury	
United States	National Model Systems data base[14–17]
Britain	Frankel et al,[18] Coll et al[19]
Australia	Yeo et al[20]
Denmark	Hartkopp et al[21]
Traumatic brain injury	
United States	Traumatic Coma Data Bank[22]
	California Developmental Disabilities Database[23,24]
Britain	Roberts,[25] Lewin et al[26]

acteristics. The chief sources of data for some major chronic disabilities are indicated in Table 1.[1–26] Among the methods for analyzing multiple risk factors on longitudinal (cohort) data, generally known as survival analysis, by far the most widely used is the Cox proportional hazards model. In this framework one estimates the survival curve, which gives a subject's probability of surviving any specified number of months or years. The curve for a given subject depends on his or her age, sex, severity of condition, and other risk factors. The texts by Collett[27] and Lee[28] give good coverage of the methods, and the review article of Singer and Willett[29] is an excellent introduction.

For the needs of the medical director, however, the Cox survival model has serious drawbacks. Specifically, (a) it does not conveniently handle cases where subjects may enter the study several years after injury date or first diagnosis and (b) it is cumbersome when the subject's characteristics, such as stage of cancer or severity of disability, change over time. The method is particularly awkward at separating the intertwined effects of age, current calendar year, and time since onset or diagnosis. For example, one often wishes to model a secular trend to capture the reduction in mortality risk over time due to improved treatment. This will generally violate the proportional hazards assumption in the Cox model.[27] It is possible to cope with this by introducing time-varying covariates,[27] but the analysis quickly becomes very cumbersome.

The problems arise because the method is cohort-based: the unit of analysis is a single patient followed over time. What is needed is an analysis of risk over the short-term, such as a year. Then the research questions can be asked in a more convenient and direct way: How is the risk of dying in the next 12 months, say, affected by the subject's age, sex, and other characteristics together with the time since injury and current calendar year? In this approach, the unit of analysis is a person-year rather than a person.

The person-year methodology, which is simple and convenient to use, is not new. It has been extensively used in the Framingham study of heart disease.[30] Guilkey and Rindfuss[31] and McLanahan[32] have applied the method in sociology, and it has been used by the present authors in numerous studies of mortality of persons with developmental disabilities.[23,33–36] To our knowledge, however, it has not been widely used in research on cancer, physical injuries, or other medical conditions. Our aim here is to show in a nontechnical fashion how it can be used in practice. We illustrate with analysis of a large data base on spinal cord injury. As we will see, the person-year analysis reveals findings on spinal cord mortality that had not been apparent with the earlier methods.

THE METHOD

Construction of the Person-Year Data Set

The construction is best introduced by an example. Consider a hypothetical study that runs for the 6 calendar years 1988–93. We are

Table 2. Potential Risk Factors Defined at Each Person-Year

AGE	at the midpoint of the person-year
SEX	1 if male, 0 if female
CONDITION	a measure of severity of condition; for illustration, we assume here that there are three levels: mild (level 1), moderate (level 2), and severe (level 3); in practice, there may be several such variables, perhaps corresponding to disabilities in the cognitive, motor, and self-care domains
TSI	time since injury, in years
YEAR	calendar year of the person-year, suitably coded; for illustration, we have coded YEAR as 1 if the year is 1990 or later and 0 if not
DIED	a binary variable set to 1 if the subject died in the year in question and 0 if not

interested in the effect on mortality risk of the variables in Table 2.

Consider Subject A, a male whose injury (or onset of the disease) occurred in late 1985. He entered the study at the beginning (January 1988) at age 50 and died in October 1992. His severity was moderate (level 2) for the first 2 years in the study and became severe thereafter. This subject contributes 5 person-years to the study, as shown in Table 3.

It is important to note that each row in Table 3 contains the information for a single year. The last row, for example, corresponds to the fifth year of the study (1992). In this year, the subject's age was 54, his condition was 3 (severe), he was 6 years postinjury, the year was in the 1990s ($YEAR = 1$), and he died ($DIED = 1$). Because he died in 1992, he does not contribute a sixth person year for 1993.

Note that the ordinal scale, severity of disability, is treated here as an interval scale with values 1, 2, or 3. It would be possible instead to code the scale with 2 dummy variables and then test whether the interval scale assumption should be rejected.

The Analytical Method

In the above example, each subject contributes up to 6 person-years to the data set. From this point, one works with this new person-year data set rather than the original subject-based information. For example, it may be helpful to provide descriptive statistics on the person-years broken down by calendar year, time since injury, age, etc.

To analyze the relationship of the various factors to mortality risk, the key tool is logistic regression.[37] Here the unit of analysis is a single person-year, the dependent variable is *DIED* (0 or 1), and the predictor variables are the covariates age, sex, severity, etc., for that year. Formally, we are fitting the model

$$\text{logit}\{P(\text{died})\} = b_0 + b_1 AGE + b_2 SEX + b_3 CONDITION + b_4 TSI + b_5 YEAR, \quad (1)$$

where logit(p) is defined to be $\log\{p/(1-p)\}$.

There is great flexibility in the choice of covariates to associate with a given year. In addition to the usual risk factors, it is, for example, possible to include a patient's characteristics from previous years. To take the above illustration, one might include not only

Table 3. Construction of person-years for subject A

Subject	Age	Sex	Condition	TSI	Year	Died
A	50	1	2	2	0	0
A	51	1	2	3	0	0
A	52	1	3	4	1	0
A	53	1	3	5	1	0
A	54	1	3	6	1	1

the patient's severity of condition in the current year but also the condition in the previous year and/or the average condition over the several preceding years. One can thus investigate whether the current conditions are enough to determine the current risk. This could be viewed as a test of a Markov condition.[38] We are not aware of any published application of this kind using the Cox model (although it would be possible using time-varying covariates). Such person-year analyses have, however, been used in our own work.[33–34]

Choice of Time-Interval

For convenience, we refer to person-years in this article, but any sufficiently short time interval would give almost identical results. Suppose, for example, that we switch from person-years to person-months. The data set becomes 12 times larger, as each row (a person-year) is replaced by 12 person-months. All the variables in these 12 rows are identical to those of the original person-year except that, if the individual died in the year, then *DIED* = 1 for the last month and 0 for all the earlier months. In addition, the time since injury variable would need to be redefined.

One reason to consider a switch from years to a shorter interval is the requirement that the interval be short enough that p, the chance of dying in the interval, is small. Technically, the condition is that p is small enough that terms involving p^2 may be neglected. In practice, it is usually sufficient if the p's are no more than about 5%.

If the above condition holds, then logit(p) can be approximated by log(p). It can then easily be shown that, in equation (1), the estimates of all the b's will be unchanged by the switch of units except that the intercept b_0 will decrease. A change from person-years to person-months, for example, would change b_0 to $b_0 - \log(12)$.

An advantage of using a shorter interval, such as a month, is that a more refined treatment of time is possible. For example, a subject entering the study in October can contribute 2 person months (November and December) to the data set. Similarly, if a subject's condition changes from level 1 to level 2 in October, the earlier months can be coded as 1 and the later months as 2.

It may seem from the above that one should ideally work with very short intervals, such as person-days. In theory this is so, but there is a practical problem: the number of person-days in a study may run into the millions, and for some purposes, this will make the computer runs unreasonably slow.

Are the Person-Years Independent?

An objection to the method that is commonly raised at first encounter is that, for example, it treats 10 person-years from one subject in the same way as if 10 different subjects each contributed 1 year. It appears that we are assuming the 10 person-years to be independent even though they may all be contributed by 1 subject. In fact, however, the person-years are not assumed to be independent. To illustrate, consider a subject who contributes 2 person-years and dies during the second year. The contribution to the likelihood function[39] can be factored into the product of (1) the probability that the subject survives the first year and (2) the probability that he dies during the second year given that he is alive at the beginning of it.

Technically, the latter is a conditional probability,[39(p61)] and the factorization is a standard procedure in probability theory that requires no independence assumption.[39] The argument is completely general. It follows that the entire likelihood function for the data is a product over all the person years of the probabilities of survival or death. Since all inferential procedures (estimation, hypothesis testing) are based on the likelihood function, this means that, when using logistic regression, we may treat the person-years as if they were independent. As noted by Singer and Willett,[29] all the standard inferential and diagnostic procedures for logistic regression may be applied.

This point is not always appreciated. Hosmer and Lemeshow, for example, incorrectly state in their well-known text[37] that

> It is not clear how or if it is possible at all to use the diagnostics computed by logistic regression software to assess the fit of the model. . . . [E]ach subject may contribute more than one line of data to the analysis and the software assumes each line corresponds to an independent subject. Summary statistics such as the deviance, as computed in logistic regression software, may not be meaningful.

Relationship of the Method to Traditional Survival Analysis

The logistic regression analysis of person-years is actually a form of discrete survival analysis. More specifically, in the limit as the time interval (years, months, etc.) becomes small, the method becomes equivalent to the Cox model with time-varying covariates. We summarize the relevant mathematical facts that justify this statement.

(1) Consider a Cox model for survival analysis and assume that the underlying baseline hazard rate[27(p206)] is piecewise constant (ie, we can partition the study period into intervals within which the hazard rate is constant). This assumption can always be satisfied if one works with a sufficiently small time interval. Thus, if a year is too long for the hazard rate to be assumed constant, one can work with months or days. Suppose, for simplicity, that the hazard rate is constant throughout each year.

(2) Given this, it can be shown[40] that the Cox model is exactly equivalent to a Poisson regression model.[41] In this discrete analysis, there is a binary lived/died variable and a covariate vector x for each person-year. All person-years with the same covariates x are grouped, and the group total number of deaths $y(x)$ is assumed to follow a Poisson distribution with mean $\exp\{\beta' x\}$. Here β is a vector of regression parameters to be estimated.

(3) Provided that the probability of dying in any given year is small (which will be true if the time interval is sufficiently small), this Poisson regression model can be conveniently fitted and analyzed using logistic regression.[41,42]

EXAMPLE: SPINAL CORD INJURY

We apply the person-year logistic analysis to data from the National Model Systems Database[15] on persons with spinal cord injury (SCI). Additional records on persons who were treated at model systems but were not eligible for the national database were also included. This data has been extensively analyzed previously by DeVivo and associates using the Cox model.[16,17,43,44]

The data has been described elsewhere.[14] Briefly, persons with a traumatic SCI from 1973 to 1998 and seen within 1 year of injury at either a model SCI care system or a Shriner's Hospital SCI unit were included in the augmented database. We considered only persons over age 10 who were not ventilator dependent (Table 4).

This data was used to construct a data set of 224,594 person-years from 18,872 subjects, of whom 3114 died during the study period. This corresponds to a crude mortality rate of 14 per 1000 person-years. We used standard methods for logistic regression modeling.[37]

Regarding model selection methods, we concur with Hasmer and Lemeshow[37] that it is better to take account of biological plausibility and knowledge of the subject matter than to rely entirely on automatic methods such as stepwise logistic regression. When analyzing large person-year data sets, we have found likelihood-based model comparisons to be especially helpful. Specifically, the deviance statistic[45]—minus twice the log-likelihood ratio—is perhaps the most useful way of comparing two models when one is nested in the other, and the Akaike information statistic[27] is helpful when the models are not nested.

Table 5 is our preferred model. Row 1 shows that males have 1.30 times the odds of dying as females, other things being equal. As would be expected, high quads—persons with a C1–C4 injury—are at much higher risk than others. For example, C1–C4's with com-

Table 4. Variable Names and Definitions

Sex	Male compared with female
White	Caucasian versus other
Violence	Violent etiology (cause of injury) versus other (accidental) causes
Age	Chronological age of the person during the person-year
Calendar year	Calendar year corresponding to the person-year
Neurological level	Highest spinal vertebra with normal sensory and motor function. This could be any of the 8 cervical vertebrae (referred to as C1–C8), 12 thoracic, 5 lumbar or 5 sacral
Frankel grade	Severity of the injury, with categories A (complete injury: no sensory or motor function below the level of injury), B (incomplete injury: sensory but not motor), C (incomplete: some motor function preserved), and D (incomplete: a majority of motor function preserved)
Time since injury	Integer number of years since date of injury for the current person-year

plete injuries (Frankel grade A) have 3.30 times the odds of dying compared with those in the reference group (others) and 1.66 (= 3.30/1.99) times the odds of those with a C5–C8A injury.

The results from Table 5 are broadly similar to those reported earlier by DeVivo et al.[17] In particular, mortality is higher during both the first (odds ratio [OR] = 1.00) and the second (0R = 0.66) years postinjury than it is subsequently (OR = 0.47), other factors being equal. The focus on person-years, however, enabled us to examine the interaction of secular trend and time since injury more closely. To this end, we refitted the model allowing a separate dummy for each combination of time period (calendar year interval) and time since injury. The reference group was the 1973–79 period for the first year after injury. The odds ratios are shown in Table 6.

Table 6 shows a steady decline in mortality during the first postinjury year over the 1973–97 study period. This no doubt reflects the improvement in care and treatment. Mortality during the second and subsequent years has also decreased, but the decline is less marked and there is even a hint of an upturn during the second postinjury year in the most recent period.

Using the Cox model, DeVivo et al[17] reported a stronger upturn in mortality rates during this most recent period. It is possible that this is an artifact of that statistical method since persons injured more recently tend to be censored after a much shorter interval than those injured many years ago, and early deaths tend to have a greater impact on Cox model results than later deaths. However, the Cox model analysis used a slightly more recent version of the database that included 1998 and part of 1999, and this might also explain part of the difference.

DISCUSSION

Survival analysis, usually considered a longitudinal analysis of cohort data, can be carried out by logistic regression—a simple cross-sectional method. The method is easy to apply using standard computer packages such as SAS[46] that come with a variety of options and diagnostics. As noted in the previous section, tests based on the likelihood function can be particularly helpful. Further, a graphical method is available if one wishes to compare risk-adjusted mortality rates for two or more groups, such as a treatment and a control group.[47]

The method can be used with endpoints other than death. For example, one could model the likelihood that a cancer metastasizes during a given year as a function of treatment, age, duration of time in the present stage of cancer, and other factors. For a patient with chronic disabilities, the chances of improvement or decline in functioning over the year (or other time period) can be similarly modeled.

The focus on the person-year as unit of analysis, rather than an individual's entire survival history, makes it easy to disentangle the effect of factors such as current age, time

Table 5. Logistic regression model

Term	Odds Ratio	95% confidence interval
Gender		
Male	1.30	(1.18, 1.44)
Female	1.00	
Race		
Caucasian	0.96	(0.88, 1.04)
Other	1.00	
Cause of injury		
Violent	1.26	(1.12, 1.42)
Accidental	1.00	
Age group		
10–20	1.00	
20–30	1.78	(1.38, 2.29)
30–40	3.30	(2.57, 4.23)
40–50	5.84	(4.54, 7.50)
50–60	10.62	(8.25, 13.67)
60–70	20.35	(15.80, 26.20)
70+	56.55	(43.88, 72.87)
Calendar year		
1973–79	1.00	
1980–84	0.74	(0.65, 0.85)
1985–89	0.65	(0.57, 0.74)
1990+	0.57	(0.51, 0.65)
Neurological level and frankel grade of injury		
C1–C4 A	3.30	(2.80, 3.89)
C5–C8 A	1.99	(1.72, 2.31)
T1–S5 A	1.08	(0.92, 1.25)
C1–C4 BC	1.92	(1.58, 2.34)
C5–C8 BC	1.37	(1.16, 1.61)
T1–S5 BC	1.16	(0.97, 1.39)
C1–C4 D	0.88	(0.71, 1.07)
C5–C8 D	0.73	(0.61, 0.87)
T1–S5 D	0.68	(0.56, 0.82)
All others	1.00	
Time since injury		
<1 year		
1 year	1.00	
2 years or	0.66	(0.57, 0.76)
more	0.47	(0.43, 0.51)

since injury, and calendar year. Further, there is no difficulty in handling persons who enter the study some years after injury. These tasks are much more difficult in the framework of the Cox regression analysis.

Table 6. Odds ratios for combinations of calendar year and time since injury

Calendar year	Years since injury 1	2	3+
1973–79	1.00*	0.60	0.42
1980–84	0.73	0.49	0.32
1985–89	0.72	0.35	0.28
1990–93	0.52	0.36	0.26
1994–97	0.41	0.42	0.26

* The reference group.

Researchers rarely raise questions when they are unfamiliar with the tools to answer them. As a result, they may miss the opportunity to extract useful information from their studies. For this reason, we encourage the use of the person-year approach in mortality studies. As an example, Strauss et al[8] found a secular trend in survival of children in a vegetative state: survival rates of infants aged under 2 years improved markedly during the 1981–96 study period, whereas little or no improvement was observed for older children. These findings emerged quite naturally from the person-year approach, but researchers equipped only with cohort methods may not have raised the question at all.

This work was supported in part by grant H133N50009–96A from the National Institute on Disability and Rehabilitation Research, Office of Special Education and Rehabilitative Services, US Department of Education, Washington, DC, to the University of Alabama at Birmingham. The authors express their gratitude to Richard D. Rutt for his assistance with data base development and data processing.

REFERENCES

1. Strauss DJ, Shavelle RM, Anderson TW. Life expectancy of children with cerebral palsy. *Pediatr Neurol.* 1998;18:143–149.
2. Strauss DJ, Shavelle RM. Life expectancy of adults with cerebral palsy. *Dev Med Child Neurol.* 1998;40: 369–375.
3. Strauss DJ, Cable W, Shavelle RM. Causes of excess mortality in cerebral palsy. *Dev Med Child Neurol.* 1999;41:580–585.
4. Hutton JL, Cooke T, Pharoah POD. Life expectancy

in children with cerebral palsy. *Br Med J.* 1994;309: 431–435.

5. Maudsley G, Hutton JL, Pharoah POD. Cause of death in cerebral palsy: a descriptive study. *Arch Dis Child.* 1999;81:390–394.
6. Chrichton JU, Mackinnon M, White CP. The life expectancy of persons with cerebral palsy. *Dev Med Child Neurol.* 1995;37:567–576.
7. Ashwal S, Eyman RK, Call TL. Life expectancy of children in a persistent vegetative state. *Pediatr Neurol.* 1994;10:27–33.
8. Strauss DJ, Shavelle RM, Ashwal S. Life expectancy and median survival time in the permanent vegetative state. *Pediatr Neurol.* 1999;21:626–631.
9. Sazbon L, Groswasser Z. Medical complications and mortality of patients in the postcomatose unawareness (PC-U) state. *Acta Neurochir (Wien).* 1991;112:110–112.
10. Sazbon L, Zagrega R, Ronen J, Solzi P, Costeff H. Course and outcome of patients in vegetative state of nontraumatic aetiology. *J Neurol Neurosurg Psychiatry.* 1993;56:407–109.
11. Higashi K, Hatano M, Abiko S, et al. Clinical analysis of patients recovered from persistent vegetative state. *No To Shinkei.* 1978;30:27–35.
12. Higashi K, Sakata Y, Hatano M, et al. Epidemiological studies on patients with a persistent vegetative state. *J Neurol Neurosurg Psychiatry* 1977;40: 876–885.
13. Higashi K, Hatano M, Abiko S, et al. Five year follow-up study of patients with persistent vegetative state. *J Neurol Neurosurg Psychiatry.* 1981;44: 552–554.
14. Richards JS, Go BK, Rutt RD, Lazarus PB. The national spinal cord injury collaborate database. In: Stover SL, DeLisa JA, Whiteneck GG, eds. *Spinal Cord Injury: Clinical Outcomes from the Model Systems.* Gaithersburg, Md: Aspen; 1995;10–20.
15. Stover SL, DeLisa JA, Whiteneck GG, eds. *Spinal Cord Injury: Clinical Outcomes from the Model Systems.* Gaithersburg, Md: Aspen; 1995.
16. DeVivo MJ, Stover SL. Long-term survival and causes of death. In Stover SL, DeLisa JA, Whiteneck GG, eds. *Spinal Cord Injury: Clinical Outcomes from the Model Systems.* Gaithersburg, Md: Aspen; 1995:289–316.
17. DeVivo MJ, Krause JS, Lammertse DP. Recent trends in mortality and causes of death among person with spinal cord injury. *Arch Phys Med Rehabil.* 1999;80:1411–1419.
18. Frankel HL, Coll JR, Charlifue SW, et al. Long-term survival in spinal cord injury: a fifty year investigation. *Spinal Cord.* 1998;36:266–274.
19. Coll JR, Frankel HL, Charlifue SW, Whiteneck GG. Evaluating neurological group homogeneity in assessing the mortality risk for people with spinal cord injuries. *Spinal Cord.* 1998;36:275–179.
20. Yeo JD, Walsh J, Rutkowski S, Soden R, Craven M, Middleton J. Mortality following spinal cord injury. *Spinal Cord.* 1988;36:329–336.
21. Hartkopp A, Bronnum-Hansen H, Seidenschnur A-M, Biering-Sorensen F. Survival and cause of death after traumatic spinal cord injury: a long-term epidemiological survey from Denmark. *Spinal Cord.* 1997;35:76–85 [published errata in *Spinal Cord,* 1997;35:862–864].
22. Foulkes MA, Eisenberg HM, eds. Report on the traumatic coma data bank. *J Neurosurg.* 1991; 74(suppl).
23. Strauss DJ, Shavelle RM, Anderson TW. Long-term survival of children and adolescents after traumatic brain injury. *Arch Phys Med Rehabil.* 1998;79: 1095–1100.
24. Strauss DJ, Shavelle RM, DeVivo MJ. Long-term mortality risk after traumatic brain injury [letter to editor]. *J Insurance Med.* 1999;31:104–105.
25. Roberts AH. *Severe Accidental Head Injury: An Assessment of Long-Term Prognosis.* London: Macmillan; 1979.
26. Lewin W, Marshall TFD, Roberts AH. Long-term outcome after severe head injury. *Br Med J.* 1979; 2(6204):1533–1538.
27. Collett D. *Modelling Survival Data in Medical Research.* London: Chapman and Hall; 1994.
28. Lee ET. *Statistical Methods for Survival Data Analysis.* New York: Wiley; 1992.
29. Singer JD, Willett JB. Modeling the days of our lives: using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psych Bull.* 1991;110:268–190.
30. Cupples LA, D'Agostino RB, Anderson K, Kannel WB. Comparison of baseline and repeated measure covariate techniques in the Framingham Heart Study. *Stat Med.* 1998;7:205–222.
31. Guilkey DK, Rindfuss RR. Logistic regression multivariate life tables. *Sociol Methods Res.* 1987;16:276–300.
32. McLanahan S. Family structure and dependency: early transition to female household headship. *Demography.* 1988;25:1–16.
33. Strauss DJ, Kastner TA. Comparative mortality of people with developmental disability in institutions and in the community. *Am J Ment Retard.* 1996;101:26–40.
34. Strauss DJ, Kastner TA, Shavelle RM. Mortality in persons with developmental disabilities, 1985–94. *Ment Retard.* 1998;30:368–391.
35. Strauss DJ, Shavelle RM, Baumeister AA, Anderson TW. Mortality of persons with developmental disabilities after transfer into community care. *Am J Ment Retard.* 1998;102:569–581.
36. Strauss DJ, Kastner TA, Ashwal S, White JF. Tubefeeding and mortality in children with severe dis-

abilities and mental retardation. *Pediatrics.* 1997;99: 358–362.
37. Hosmer D, Lemeshow S. *Applied Logistic Regression.* New York, NY: Wiley; 1989.
38. Breiman L. *Probability and Stochastic Processes.* 2nd ed. Palo Alto, Ca: The Scientific Press; 1986.
39. Hogg RV, Craig AT. *Introduction to Mathematical Statistics.* 4th ed. New York, NY: Macmillan; 1978.
40. Laird N, Olivier D. Covariance analysis of censored survival data using log-linear analysis techniques. *J Am Stat Assoc.* 1981;76:231–240.
41. Agresti A. *Categorical Data Analysis.* New York, NY: Wiley; 1990.
42. Efron B. Logistic regression, survival analysis, and the Kaplan-Meier curve. *J Am Stat Assoc.* 1988;83: 414–425.
43. DeVivo MJ, Kartus PL, Stover SL, Rutt RD, Fine PR. Seven-year survival following spinal cord injury. *Arch Neurol.* 1987;44:872–875.
44. DeVivo MJ, Stover SL, Black KJ. Prognostic factors for 12-year survival after spinal cord injury. *Arch Phys Med Rehabil.* 1992;73:156–162.
45. McCullagh P, Nelder JA. *Generalized Linear Models.* 2nd ed. London: Chapman and Hall; 1989.
46. SAS Institute Inc. *SAS/STAT User's Guide.* Version 6. 4th ed. Volume 2. Cary, NC: SAS Institute Inc; 1989.
47. Strauss DJ. On Miettinen's multivariate confounder score. *J Clin Epidemiol* 1998;51:233–236.