



From the SelectedWorks of Paula Diehr

December 2005

Accounting for missing data in end-of-life research

Contact
Author

Start Your Own
SelectedWorks

Notify Me
of New Work

Available at: http://works.bepress.com/paula_diehr/45

Accounting for Missing Data in End-of-Life Research

PAULA DIEHR, Ph.D.¹ and LAURA LEE JOHNSON, Ph.D.²

ABSTRACT

End-of-life studies are likely to have missing data because sicker persons are less likely to provide information and because measurements cannot be made after death. Ignoring missing data may result in data that are too favorable, because the sickest persons are effectively dropped from the analysis. In a comparison of two groups, the group with the most deaths and missing data will tend to have the most favorable data, which is not desirable. Results based on only the available data may not be generalizable to the original study population. If most of the missing data are absent because of death, methods that account for the deaths may remove much of the bias. Imputation methods can then be used for the data that are missing for other reasons. An example is presented from a randomized trial involving frail veterans. In that dataset, only two thirds of the subjects had complete data, but 60% of the "missing" data were missing because of death. The available data alone suggested that health improved significantly over time. However, after accounting for the deaths, there was a significant decline in health over time, as had been expected. Imputation of the remaining missing data did not change the results very much. With and without the imputed data, there was never a significant difference between the treatment and control groups, but in two nonrandomized comparisons the method of handling the missing data made a substantive difference. These sensitivity analyses suggest that the main results were not sensitive to the death and missing data, but that some secondary analyses were sensitive to these problems. Similar approaches should be considered in other end-of-life studies.

INTRODUCTION

MISSING DATA REPRESENT a potential problem for all research studies, and especially for those that concentrate on the end of life (EOL). For persons nearest to death, data are likely to be missing, and therefore these individuals are likely to be under-represented in many comparisons. Analysis of only the available data leads to findings that apply only to a subgroup of the original population of interest, and that subgroup cannot be prospectively identified. In a randomized

trial setting, such an analysis no longer has the benefits of randomization and is not strictly "intent to treat." In this paper, we refer generically to a measure of patient "status" that is measured multiple times and is sometimes missing. A common study design in EOL studies is to recruit a cohort of seriously ill persons and to follow them closely until the end of the study or until death. Thus, the data are prospectively collected. (Retrospective data collection, after the person has died, is not discussed here.) Two types of analysis are possible with such data. A "forward" anal-

¹School of Public Health and Community Medicine, University of Washington, Seattle, Washington.

²National Center for Complementary and Alternative Medicine (NCCAM), National Institutes of Health, U.S. Department of Health and Human Services, Bethesda, Maryland.

ysis follows the person from the baseline measurement until the study end or death. A “backward” analysis looks at status at different times relative to the date of death.

COMMON TYPES OF MISSING DATA IN EOL STUDIES

Four scenarios for missing data are particularly important in end-of-life studies. The two most likely to cause bias involve data missing because (1) the person died before the end of the study, and (2) data are missing near death because the participant was too ill to continue in the study. Data that are (3) missing far from death, or (4) missing because the person survived, are important for backward analyses, primarily because they result in loss of power.

Data missing because of death

Once a person dies, that person’s status data are “missing” in that they cannot be collected. In a forward analysis, we might wish to estimate the trajectory of the status score after the baseline measurement, to describe how status declines over time. Ignoring the deaths yields a study group that is too healthy and tends to favor the group with the most deaths.

Data missing near to death

Persons who are very sick are less likely to attend study appointments or respond to questionnaires, which suggests that data near death are more likely to be missing. This is likely to cause bias in both forward and backward analyses, as persons with worse status will be underrepresented in the available data.

Data missing far from death

A backward analysis might compare a person’s status just before death to that person’s status 2 years earlier. A person who dies soon after baseline has no measured data 2 years before death. Such data, missing for administrative reasons, are often assumed to be “missing completely at random,” in which case ignoring the missing data will not cause bias.^{1,2} However, the assumption may not be true. For example, if all persons were the same age at baseline, the missing data would come from persons who died young, meaning

that the unobserved data was for younger but sicker persons.

Data missing because person survived

In the backward analysis, status measurements are classified by the length of time they were made before death. If the person survives, the time from death is unknown for the available data. This type of missing data lowers the power of the backward analyses, as survivors will have to be omitted. It may also yield biased estimates for the population of interest, as those who survive the longest after baseline will be omitted. This situation has not been well studied.

We will consider only the first two types of missing data, as the latter two may not result in bias and affect only the backward analyses.

APPROACHES FOR DATA “MISSING” BECAUSE OF DEATH

Although status data cannot be collected after a person dies, the person’s situation is known with certainty. What is needed is an administrative decision about how to handle deaths. If status had been measured on a utility-based scale, where perfect health is assigned the value 100 and death is assigned 0, the deaths would not be missing, as they would be coded as 0.

When (as is usually the case) the measure of interest is not utility based, one may transform the status variable into a new variable that has a natural value for death. For example, the generic “status” variable could be transformed into a new binary variable with the value 100 if the person had good status and 0 if the person had poor status or was dead. The definition of the new variable is thus whether the person had good status at this time. The mean of the new variable is the percentage of the original cohort who had good status at this time (as opposed to having poor status or being dead). The only requirement is that death can be considered as not having good status.

This binary transformation might discard useful detail in the original variable. A transformation with better face validity is the “probability of being healthy in the future” (PHF) transformation. (“Healthy” can be defined in a variety of ways.) The available data are transformed to a new variable the value of which is the probab-

ity that a person with the observed value will be healthy 1 year in the future. Deaths are set to 0 on the new variable because persons who have died have no probability of being healthy in the future.³ The new transformed variable is never missing once a person dies. This suggests that the data missing for reasons other than death should be imputed in some way, to keep the deaths from having too much influence in the analysis.

Some analyses treat the deaths like other missing data. This involves using one of the standard missing-value techniques (described below) to impute the values after death as if the person had not died.⁴ Another frequently used approach is to set the values after death equal to an arbitrary low value, which may be useful for significance testing but does not yield an interpretable summary measure.⁵ A final approach is to model status and survival jointly, yielding an estimated survival and estimate of status conditional on survival for each person; but as yet no measure that combines survival and status has been proposed.⁶⁻⁸

Methods that summarize the person's entire experience over a time period may be calculated using quality-adjusted life years (QALYs)⁹ or quality-adjusted time without symptoms and toxicity (Q-TWIST)⁹⁻¹¹; but these methods involve utilities and health states that are often hard to define and to reproduce across studies.^{9,12}

APPROACHES FOR MISSING DATA FOR REASONS OTHER THAN DEATH

Some missing data may be prevented by more effective study procedures or the use of proxy respondents. Once data are missing, there is no rule of thumb for whether they cause a problem. It is possible for results to change importantly depending on how the missing data are treated, even when the amount of missing data is small or is the same in the treatment and control groups.¹³ There is also no guaranteed way to improve the situation. Many approaches have been suggested, but all rely on strong assumptions that cannot be tested. One approach is to impute the missing data by substituting another value from that person (e.g., last observation carried forward), a value from a similar person (e.g., the hot deck method), or an estimate from a regression equation. A recent study of 14 simple imputation methods found that most yielded imputed values that were too optimistic and that the variance of

the imputed data tended to be too low.¹⁴ Imputation based on the person's own observed data (measured at another time) had better performance than the other methods examined. Multiple imputation is an approach that creates several different imputed values for each missing observation and combines the results to yield a standard error that accounts for the increased uncertainty due to using imputed data.^{1,2} In this paper we do not discuss the theory behind multiple imputation, which can be shown to be appropriate for certain types of missing data and certain imputation techniques.

A final approach is to deal specifically with the missing data at the time of analysis, modeling the missingness mechanism as part of the analysis. The probability of being missing may depend on information that is not in the dataset, most importantly on the unknown value itself. Two approaches for this nonignorable missingness are pattern mixture models¹⁵⁻¹⁷ and selection models.¹⁷⁻²¹ These methods are all based on strong and untestable assumptions and often require specialized software.¹³

The goals of imputation are both to keep the person in the analysis and to obtain a good estimate of the missing data. In some situations, even a poor estimate may be better than losing a research subject. The best method depends on the context, and an approach that reduces bias for one analysis could increase bias for another analysis. There is no "best" way to impute data. Because no approach for handling missing values can be guaranteed to reduce bias, and some may even make the situation worse,²² the best strategy is a sensitivity analysis, to determine whether the important findings are sensitive to what is done with the missing data. Ideally, researchers would try several missing value approaches known to have different performance, and determine whether the key findings hold under all of these approaches.

Example

To illustrate these points we use data from the Adult Day Health Care Program (ADHC).²³⁻²⁵ In the evaluation of the ADHC, 939 frail veterans were identified who satisfied at least one of the following criteria: currently residing in or admitted to the hospital from a nursing home; used personal assistance or supervision for ambulation, toileting, or dressing; bowel incontinence; or sig-

nificant cognitive impairment. Of these, 397 were randomized to a VA-administered ADHC program (ADHC1), 385 to customary care (controls), and an additional 157 persons enrolled in contract ADHC programs were included as a nonrandomized comparison group (ADHC2). (An additional 57 persons who provided no data are not considered here). Their health was measured at baseline, 6 months, and 12 months. The study was negative, showing higher costs and little benefit for ADHC. An interesting feature of the data, however, was that the frail veterans did not appear to become sicker over time. We now explore these data further, to see how deaths and missing values contributed to this unexpected finding.

Data were available at all three times for 626 subjects, 159 had valid values but then died, 125 had some valid and some missing values, 14 had a valid value and a missing value and then died, and 15 had missing data and then died. Another way to describe the missing data problem is that the 939 persons would have generated $3 \times 939 = 2817$ observations if the dataset were complete. However, 302 observations were missing because of death and 200 were missing for other reasons. Clearly, accounting for the deaths would greatly reduce the problem of missing data.

We will examine the ubiquitous variable "How is your health? Excellent, very good, good, fair, or poor" (EVGGFP). There is no standard way to

code EVGGFP. It is often coded on an ordinal scale, with excellent = 5, and poor = 1. Instead, we coded each level as the probability of being healthy 1 year in the future (PHF), conditional on the observed value of the measure. (In this case, "healthy" is defined as being in excellent, very good, or good health). That is, based on values in the literature for a variety of samples of middle-aged and older adults, we coded excellent as 95 because about 95% of persons in excellent health are healthy 1 year later; very good was coded as 90, good as 80, fair as 30, and poor as 15 because only about 15% of persons in poor health are healthy 1 year later.³ (Age-specific transformations are also available.)²⁶ The mean value of EVGGFP using PHF coding is interpretable because mean PHF is the expected percent who will be healthy 1 year in the future. More important, the PHF coding has a natural value for death, because the probability that a dead person will be healthy 1 year later is zero.

Missing data (not caused by death) were imputed as follows. The observed missing data patterns are shown in Table 1, where V represents a valid value, D represents a value missing because of death, and M represents a missing value. Only 626 persons had complete data (the VVV pattern). Setting the deaths to 0 changed VVD to VV0 and VDD to V00, which then became complete cases. VM0, MV0, and M00 were still not complete be-

TABLE 1. PROFILE OF STATUS OF THREE ANNUAL OBSERVATIONS AFTER ACCOUNTING FOR THE DEAD AND MISSING OBSERVATIONS*

<i>No. of subjects</i>	<i>Initial profiles*</i>	<i>Set dead to zero</i>	<i>+Impute missing data where there are two known values</i>	<i>+Impute remaining missing data</i>
626	VVV	VVV	VVV	VVV
60	VVD	VV0	VV0	VV0
99	VDD	V00	V00	V00
10	VMV	VMV	VIV	VIV
11	VMD	VM0	VIO	VIO
55	VVM	VVM	VVI	VVI
14	MVV	MVV	IVV	IVV
3	MVD	MV0	IV0	IV0
15	MDD	M00	I00	I00
44	VMM	VMM	VMM	VII
2	MVM	MVM	MVM	IVI
No. of complete cases	626	785	893	939

*Profile shows status of health measurement at baseline, 6 months, and 12 months, after accounting for the dead and missing data as denoted in the column heading.

V, valid value obtained; D, dead (no value because person was dead); M, missing (no valid value obtained, even though person was alive); I, imputed missing value (see text).

0, dead set to 0, the appropriate value using PHF coding (see text).

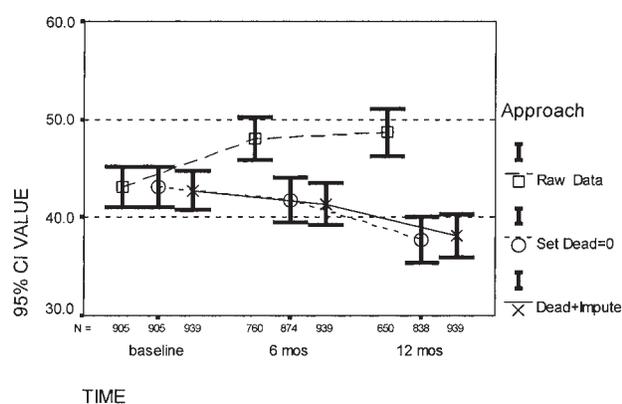


FIG. 1. Mean health over time accounting for death and missing.

cause there were missing values as well as death. In the third column we imputed a value if the other two values were known, using simple methods found to work reasonably well in this type of situation.¹⁴ The missing middle value in VIV and VI0 were imputed as the average of the first and last measures. The missing 12-month values in VVI were set equal to the 6-month values (last observation carried forward) and the baseline values for IVV and IVD were also set equal to the 6-month values (next observation carried back). For the I00 pattern we imputed the missing initial value as "poor health," as the person died shortly after baseline. In the fourth column, we imputed data for the two remaining patterns by setting the missing data equal to the single known observation. Other approaches could, of course, have been used. For example, if we believed that people usually had missing data because they were sick, the imputed value could have been set to be less healthy than the last observation carried forward.

Figure 1 shows the mean values of EVGGFP over time using the PHF coding. Three different ways of handling the deaths and missing are shown. The topmost line (square markers) connects the error bars for the means of the raw (available) data. Surprisingly, there was a significant improvement in health, even though these persons were selected to be frail and were expected to decline in health. The numbers of observations at each time are shown at the bottom of the chart (905, 760, and 650 for the raw data). The second line (round markers) shows the means if the deaths are set to 0. Health now declines significantly over time instead of becoming significantly better, and the sample sizes are larger. Thus, accounting for the deaths using the PHF method made the trend over time for these frail veterans more reasonable. The remaining line (X markers) shows the results of also imputing the missing values not caused by death. This imputation made little difference, but the sample size was increased to 939 at each time point. If we treated the deaths as ordinary missing data and simply imputed them, the result was similar to the top line in Figure 1; that is, the bias was not removed by imputing the deaths rather than accounting for them using the PHF method. If we imputed values only if the other two were known, the line was similar to the two lower lines in Figure 1.

The available data were biased, in the sense that their trends over time were incorrect. It is possible, however, for the bias to be the same in the treatment and control groups, and that it would cancel out when the groups were compared. Table 2 shows the results of *t* tests on the change from 0 to 12 months, with positive values favoring the first-mentioned group. For example,

TABLE 2. ANALYSIS RESULTS (DIFFERENCE IN CHANGE SCORES)

Data	ADHC1 vs. control			ADHC1 vs. ADHC2			Control vs. ADHC2		
	Mean ^a	s.e. ^b	P ^c	Mean ^a	s.e. ^b	P ^c	Mean ^a	s.e. ^b	P ^c
Raw	1.71	3.08	0.58	9.21	3.77	0.02	7.50	3.88	0.06
Death = 0	0.80	2.99	0.79	7.17	3.74	0.06	6.37	3.85	0.10
Impute if 2	2.57	2.86	0.37	6.16	3.54	0.08	3.60	3.65	0.33
Impute if 1	2.24	2.67	0.40	5.83	3.34	0.08	3.59	3.42	0.30
Multiple imputation	1.17	2.87	0.68	5.63	3.77	0.14	4.46	3.99	0.27

^aDifference between the mean change score in the first-mentioned group and the second-mentioned group. Positive values favor the first-mentioned group.

^bStandard error of the difference.

^c*p*-value (two-tailed).

ADHC, Adult Day Health Care program.

comparing ADHC1 to control, the ADHC1 group improved 1.7 more points than the controls, on average, in the raw data. This difference dropped to 0.8 points when deaths were accounted for, then rose to 2.6 and 2.2 when the two levels of imputation were performed. None of these comparisons approached statistical significance. The nonrandomized comparison of ADHC1 to ADHC2, or of control to ADHC2, yielded a different pattern, with statistically significant differences in the raw data, which became less significant or nonsignificant after accounting for deaths and missing data. In the primary comparison of ADHC1 to controls, accounting for death and missing data never yielded a significant difference between the ADHC1 group and the controls. That is, the main findings were not sensitive to the method of accounting for death and missing data. The other two comparisons, however, did yield different results, and so the results must all be reported.

The approaches used thus far were all single imputations. Multiple imputation is not a specific type of imputation, but rather a way to account for the extra variability caused by using imputed data instead of real data. This is done by imputing three (or more) values for each missing observation, resulting in three different complete datasets. The substantive analysis is performed separately on each dataset, and the results combined according to published formulas.^{1,2} In our simple example, the dependent variable was the change from 0 to 12 months, some values of which were calculated from imputed data. Under multiple imputation, for each missing value, the three new multiply imputed values were calculated as the singly imputed value (described above) plus a random number chosen from a normal distribution with mean 0 and standard deviation 35. We chose 35 because it was the approximate standard deviation of the difference between the observed change from 0 to 12 months and the change if we had instead imputed some of the values, averaged across the different missing data patterns. (For example, for the VVM pattern, the 0–6 month difference was used to estimate the 0–12 month difference). Different methods for adding variation, based on a regression model, are more commonly used.

We then compared the amount of change in the two groups of interest using a *t* test on the three different datasets. The three datasets yielded three estimated differences between ADHC1 and

controls: 0.77, 1.65, and 1.10 points. The mean and sample variance of those three numbers are 1.17 and 0.19, respectively. The (squared) standard errors of the three differences were 7.95, 8.03, and 7.93, for an average of 7.97. The multiply imputed treatment effect is the average difference, or 1.17. Letting *m* be the number of imputations (*m* = 3 in this case), the variance of that estimate was $7.97 + ((m+1)/m)*0.19 = 8.23$, and the standard error was its square root, or 2.87. The *t* statistic is $1.17/2.87 = .407$, obviously not statistically significant. (The approximate degrees of freedom for this *t* test is calculated as $df = (m - 1)*[1 + \frac{m*7.97}{(m + 1)*0.19}]^2 = 2007$. As this is much larger than the value at which the *t* distribution closely approximates the normal distribution (about 30 df), degrees of freedom can be ignored.) The multiple imputation results are in the final row of Table 2. The the standard errors are larger in row 4 than in row 3, because of accounting for the imputation variability. Results became less significant with these larger standard errors.

Although not an EOL study, a recent article provides another example where the deaths did matter.²⁷ The authors classified 1915 person as consuming no, moderate, or high amounts of alcohol in 1974. In 2000 the SF-36 was administered to 1216 survivors; 200 more did not respond, and 499 had died. The SF-36 physical component scale (PCS), using the available data, was not significantly different among the three groups. The PCS was then recoded, using the PHF method, to account for the deaths.³ Here, “healthy” was defined as being in the top 75% of the published reference distribution for the PCS. The high-alcohol group had a significantly lower mean value (a lower probability of being physically healthy 1 year later), because of a higher death rate. The only valid interpretation of the original PCS results is retrospective (the health of persons in 2000 does not predict their drinking practices in 1974), whereas PHF calculated from the PCS can be interpreted prospectively as the percentage who were healthy in 2000 predicted from their drinking in 1974. Other missing data were not imputed.

Palmer examined the effect of imputation of missing values in a study of palliative care.²⁸ At 8 days there was one death and nine other missing values. The missing data were imputed in several ways, with values after death imputed as though the person was still alive. The results dif-

ferred in small but substantive ways depending on the method of imputation.

DISCUSSION

End-of-life studies are more likely than other studies to have missing data, as data are likely to be missing near death and because of death. Both the available data and the data from the complete cases are likely to be favorably biased, and the study group with the most deaths and missing observations will thus tend to have the most favorably biased data.

It is important to account for the deaths to address this bias, and accounting for them may make the data substantially more complete (only 65% were complete in the raw ADHC data, but 84% were complete after accounting for death). The PHF method is a simple way to account for the deaths, and yields a measure that is interpretable as the probability of being healthy 1 year later, or equivalently as the expected percentage who will be healthy 1 year later. "Healthy" can be defined in a variety of ways, such as being in excellent/very good/ or good health, being in the top 75% of the reference distribution, or having no ADL difficulties. The PHF method requires only that death be part of the construct that the variable of interest represents. This would probably be the case for variables representing physical health or abilities, perhaps for mental health, and perhaps not for variables measuring satisfaction or spirituality.

Transformation values have been published for EVGGFP (as used here) and for all of the subscales and summary scores of the SF-36,³ as well as for the Sickness Impact Profile and the Quality of Well-Being scale.²⁹ Transformation values have recently been developed for ADLs, IADLs, MMSE, CESD, bed days, timed walk, hospitalizations, number of blocks walked, body mass index, and systolic blood pressure.²⁶ (The authors questioned whether death was really part of the construct represented by the last three variables.) It is not difficult to develop transformations for additional variables if a large general dataset is available. Similar measures have been developed for scales measuring the status of persons after hip fractures.³⁰ The PHF method performs well, although the simpler method of dichotomizing the data and assigning the deaths to "not healthy" often has very similar results.³¹

It may also be important to impute the data missing for other reasons, to avoid overemphasizing the deaths.³² Having a complete dataset permits generalization of the results to the original population and permits an intent-to-treat analysis. We used simple imputation methods, which make it easy to see what assumptions are being made. It is also common to use computationally intensive methods to impute the missing data, which may yield better imputed values under certain assumptions, although there is no way to know whether the assumptions hold. More use of natural experiments to determine the features of the various missing value methods in end-of-life studies would be valuable. Statisticians often recommend imputing the values in several ways, to see if changing the method changes the substantive results. If the results do not change, they are probably not sensitive to missing data, and this sensitivity analysis can be reported. If the results do change, then this must be reported as well, as a study limitation.

It is common to impute missing values but less common to account specifically for the deaths in the manner shown here. In the example, half of the cases with incomplete data became complete once death was accounted for. We recommend that deaths be accounted for first and that the remaining missing values then be imputed. If the results are insensitive to these changes, then perhaps inference on only the raw data may be reported, with footnotes about the other exercises performed. Research studies with subjects near the end of life need to account explicitly for death and missing data, to ensure that the results are appropriate and generalizable to the population of interest.

REFERENCES

1. Rubin DB: *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley, 1987.
2. Raghunathan TE: What do we do with missing data? Some options for analysis of incomplete data. *Ann Rev Public Health* 2004;25:90-117.
3. Diehr P, Patrick DL, Spertus J, Kiefe CI, McDonnell M, Fihn SD: Transforming self-rated health and the SF-36 scales to include death and improve interpretability. *Med Care* 2001;39:670-680.
4. Revicki DA, Gold K, Buckman D, Chan K, Kallich JD, Woolley JM: Imputing physical health status scores missing owing to mortality: Results of a simulation comparing multiple techniques. *Med Care* 2001;39:61-71.

5. Hwang JS, Wang JD: Integrating health profile with survival for quality of life assessment. *Qual Life Res* 2004;13:1–10.
6. Kurland BF, Heagerty PJ: Marginalized transition models for longitudinal binary data with ignorable and non-ignorable drop-out. *Stat Med* 2004;23:2673–2695.
7. Pauler DK, McCoy S, Moinpour C: Pattern mixture models for longitudinal quality of life studies in advanced stage disease. *Stat Med* 2003;22:795–809.
8. Ribaldo HJ, Thompson SG, Allen-Mersh TG: A joint analysis of quality-of-life and survival using a random effect selection model. *Stat Med* 2000;19:3237–3250.
9. Fayers PM, Machin D: *Quality-of-Life*. New York: Wiley, 2000.
10. Parmar MKB, Machin D: *Survival Analysis: A Practical Approach*. New York: Wiley, 1995.
11. Schwartz CE, Cole BF, Vickey B, Gelber RD: The Q-TWiST approach for assessing health-related quality-of-life in epilepsy. *Qual Life Res* 1995;4:135–141.
12. Staquet M, Hayes RD, Fayers PM (eds): *Quality-of-Life Assessment in Clinical Trials*. Oxford: Oxford University Press, 1998
13. Fairclough DL: *Design and Analysis of Quality of Life Studies in Clinical Trials*. New York: Chapman and Hall, 2002.
14. Engels JM, Diehr P: Imputation of missing longitudinal data: A comparison of methods. *J Clin Epidemiol* 2003;56:968–976.
15. Little RJA: Pattern-mixture models for multivariate incomplete data. *J Am Stat Assoc* 1993;88:125–134.
16. Little RJA: A class of pattern-mixture models for normal incomplete data. *Biometrika* 1994;81:471–83.
17. Little RJA: Modeling the drop-out mechanism in repeated-measures studies. *J Am Stat Assoc* 1995;90:1112–1121.
18. DeGruttola V, Tu XM: Modeling progression of CD-4 lymphocyte count and its relationship to survival time. *Biometrics* 1994;50:1003–1014.
19. Diggle PJ, Kenward MG: Informative dropout in longitudinal data analysis. *Appl Stat* 1994;43:49–93.
20. Schluchter MD: Methods for the analysis of informatively censored longitudinal data. *Stat Med* 1992;11:1861–1870.
21. Wu MC, Bailey KR: Estimation and comparison of changes in the presence of informative right censoring: conditional linear model. *Biometrics* 1989;45:939–955.
22. Abraham L: The application of two methods proposed for informative drop-out in longitudinal data analysis. MS thesis, University of Washington, 2000.
23. Hedrick S, Johnson R, Inui T, Diehr P: Factors associated with participation in a randomized trial of adult day health care. *Gerontologist* 1991;31:607–610.
24. Hedrick SC, Rothman ML, Chapko M, Ehreth J, Diehr P, Inui T, Connis RT, Grover PL, Kelly JR: Summary and discussion of methods and results of the adult day health care evaluation study. *Med Care* 1993;31:SS94–S103.
25. Rothman ML, Diehr P, Hedrick SC, Erdly W, Nickinovich DG: Effects of contract adult day health care on health outcomes and satisfaction with care. *Med Care* 1993;31:S75–S83.
26. Diehr P, Johnson LL, Patrick DL, Psaty B: Incorporating death into health-related variables in longitudinal studies: *J Clin Epidemiol* 2005;58:1115–1124.
27. Strandberg AY, Strandberg TE, Salomaa VV, Pitkala K, Miettinen T: Alcohol consumption, 29-y total mortality, and quality of life in men in old age. *Am J Clin Nutr* 2004;80:1366–1371.
28. Palmer JL: Analysis of missing data in palliative care studies. *J Pain Symptom Manage* 2004;28:612–618.
29. Diehr P, Patrick D, Hedrick S, Rothman M, Grembowski D, Raghunathan TE, Beresford S: Including deaths when measuring health status over time. *Med Care* 1995;33:AS164–AS172.
30. Penrod JD, Boockvar KS, Litke A, et al: Physical therapy and mobility 2 and 6 months after hip fracture. *J Am Geriatr Soc* 2004;52:1114–1120.
31. Diehr P, Patrick DL, Burke G, Williamson J: Survival versus years of healthy life: which is more powerful as a study outcome? *Control Clin Trials* 1999;20:267–279.
32. Diehr P, Patrick DL, McDonell MB, Fihn SD: Accounting for deaths in longitudinal studies using the SF-36. *Med Care* 2003;41:1065–1073.

Address reprint requests to:

Paula Diehr, Ph.D.

School of Public Health and Community Medicine

University of Washington

Box 357232

Seattle, WA 98195-7232

E-mail: pdiehr@u.washington.edu