

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Publications, Department of Psychology

Psychology, Department of

---

2007

## Population Inference with Mortality and Attrition in Longitudinal Studies on Aging: A Two-Stage Multiple Imputation Method

Ofer Harel

*University of Connecticut*, [oharel@stat.uconn.edu](mailto:oharel@stat.uconn.edu)

Scott M. Hofer

*Oregon State University*, [scott.hofer@oregonstate.edu](mailto:scott.hofer@oregonstate.edu)

Lesa Hoffman

*University of Nebraska-Lincoln*, [lhoffman2@unl.edu](mailto:lhoffman2@unl.edu)

Nancy L. Pedersen

*Karolinska Institutet*

Follow this and additional works at: <https://digitalcommons.unl.edu/psychfacpub>



Part of the [Psychiatry and Psychology Commons](#)

---

Harel, Ofer; Hofer, Scott M.; Hoffman, Lesa; and Pedersen, Nancy L., "Population Inference with Mortality and Attrition in Longitudinal Studies on Aging: A Two-Stage Multiple Imputation Method" (2007). *Faculty Publications, Department of Psychology*. 434.

<https://digitalcommons.unl.edu/psychfacpub/434>

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *Experimental Aging Research*, (2007) 33: 187-203. Copyright 2007,  
Taylor & Francis Group, LLC. ISSN: 0361-073X print/1096-4657 online.  
DOI: 10.1080/03610730701239004. Used by permission.

## Population Inference with Mortality and Attrition in Longitudinal Studies on Aging: A Two-Stage Multiple Imputation Method

Ofer Harel

Department of Statistics, University of Connecticut,  
Storrs, Connecticut, USA

Scott M. Hofer

Department of Human Development and Family Sciences,  
Oregon State University, Corvallis, Oregon, USA

Lesa Hoffman

Department of Psychology, University of Nebraska,  
Lincoln, Nebraska, USA

Nancy L. Pedersen

Department of Medical Epidemiology and Biostatistics,  
Karolinska Institutet, Stockholm, Sweden

Received April 28, 2005; accepted April 2, 2006, posted to University of  
Nebraska-Lincoln Digital Commons February 25, 2010.

Data analyzed in this study were from the longitudinal Origins of Variance in the  
Old-Old: Octogenarian Twins study that was supported by a grant from the National  
Institute on Aging (AG08861).

Address correspondence to Scott M. Hofer, Department of Human Development  
and Family Sciences, Oregon State University, Corvallis OR, USA:  
scott.hofer@oregonstate.edu or to Ofer Harel, Department of Statistics,  
University of Connecticut, Storrs, CT, USA: oharel@stat.uconn.edu.

Boo Johansson

Department of Psychology, Göteborg University, Göteborg, Sweden; Department of Biobehavioral Health, Pennsylvania State University, University Park, Pennsylvania, USA

A major challenge for inference regarding aging-related change in longitudinal studies is that of study attrition and population mortality. Inferences in longitudinal studies can account for attrition and mortality-related change as distinct processes, but this is made difficult when follow-up of all individuals (*i.e.*, age at death) is not complete. This is a common problem because most longitudinal studies of aging either have incomplete follow-up or are still collecting data on subsequent outcomes, including time of death. A statistical approach is suggested for including time-to-death as a predictor in models with incomplete follow-up using a two-stage multiple-imputation procedure. An empirical example using data from the OCTO-Twin study is presented that shows the utility of his procedure for making inferences conditional on mortality when mortality data are incomplete.

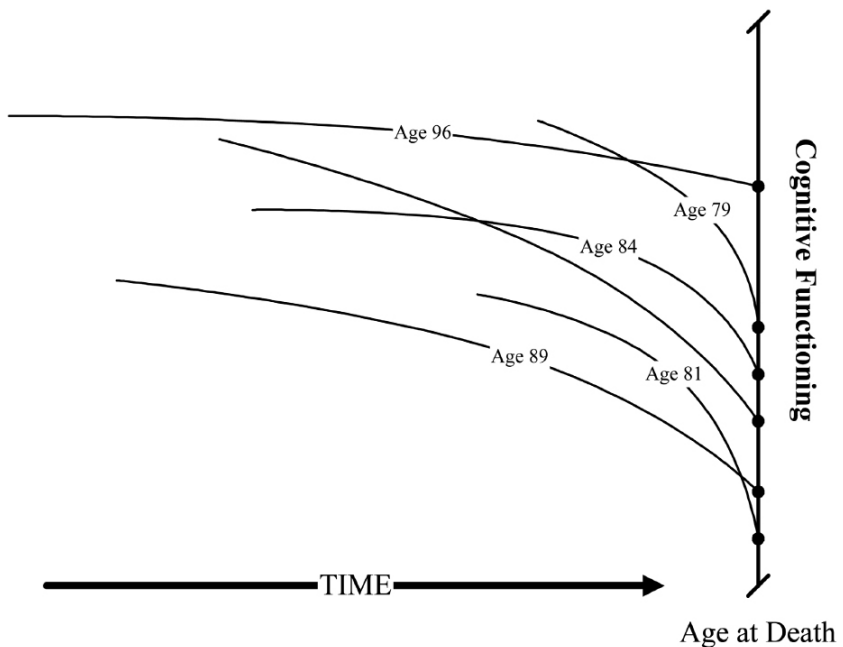
Although there are numerous challenges for the investigation of aging-related changes in older adults, statistical analysis with incomplete data and the conceptualization of population processes related to mortality is one of the most difficult. Selective attrition and mortality selection within longitudinal studies on aging are intrinsically related to many aging-related changes and must be carefully considered in the analysis and interpretation of results (*e.g.*, Baltes, 1968; Hofer & Sliwinski, 2006; Schaie, Labouvie, & Barrett, 1973). A key distinction is made between attrition (*i.e.*, selective dropout) and mortality selection (*i.e.*, selective survival) in that attrition affects characteristics of the particular sample under investigation, whereas mortality selection affects both the definition of the population as well as the sample under study (Baltes, 1968). Including time-to-death as a predictor in models for estimating change in outcomes of interest permits conditional inferences to defined populations based on age and survival (and their interaction) and is easily performed when complete data are available for both chronological age and age of death. In most studies, however, complete data for all individuals are not currently available and may not be available for a substantial period of time. The purpose of the current work is to present a two-stage multiple-imputation approach for treating mortality and attrition as distinct processes leading to incomplete data and which permit the use of time-to-death in the predictive models when follow-up is incomplete.

In major studies of cognitive aging, attrition rates between test occasions may be as high as 50% (Cooney, Schaie, & Willis, 1988). Attrition is not likely to be random in relation to important age-related characteristics. For example, older individuals and individuals with lower levels of education, socioeconomic status, compromised health, and cognitive ability are more likely to dropout from longitudinal studies (Cooney, Schaie, & Willis, 1988; Rabbitt, Watson, Donlan, Bent, & McInnes, 1994). Differences in attrition patterns are generally not found in terms of race, religion, or marital status (Siegler & Botwinick, 1979; Streib, 1966). Patterns of nonresponse are related to change in performance prior to attrition as well as to individual characteristics at the initiation of the study (Cooney *et al.*, 1988; Rabbitt *et al.*, 1994).

Raising the issue of selective attrition in the 1960s, Riegel, Reigel, and Meyer (1967) initiated exploration of the *terminal drop* phenomenon, in which individuals who drop out of a study may exhibit a marked decline in both physiological health and concomitant cognitive performance prior to death. Because these individuals typically go unmeasured, the results will be upwardly biased towards inferences of less decline in the population than is actually true. Findings of decline across a range of cognitive abilities associated with time to death provide evidence for "terminal decline" or "terminal drop" (for reviews see Berg, 1987, 1996; Siegler, 1975; Small & Bäckman, 1999; Bosworth & Siegler, 2002). Of the longitudinal studies that have evaluated this hypothesis, significant findings have typically included spans between measurement occasions of less than 3 years (*e.g.*, Deeg, Hofman, & van Zonneveld, 1990; Johansson & Berg, 1989; Smits, Deeg, Kriegsman, & Schmand, 1999).

The issue of population mortality, however, challenges most standard analyses of change and undermines inference to a single, representative, population of aging individuals. An important question to consider is *should death be considered a missing value?* In some types of longitudinal studies, such as clinical trials of treatments for cancer or acquired immunodeficiency syndrome (AIDS), death itself is the end point of interest, and the analysis focuses on modeling time of survival. In other studies, death is infrequent and unrelated to the outcome being measured; in these cases it might be appropriate to consider the counterfactual (*i.e.*, possible) outcome a participant could have if he or she had not died. Although a rationale and state-of-the-art techniques for addressing incomplete data have been developed (Diggle & Kenward, 1994; Diggle, Liang, & Zeger, 1994; Little & Rubin, 1987; Rubin, 1976, 1987; Schafer, 1997; Schafer & Graham, 2002), the application of these methods to later-life longitudinal samples remains problematic, conceptually and computationally (Hofer & Hoffman, *in press*; Hofer & Sliwinski, 2006).

In gerontological studies, death is a common and natural end point that defines the limits of inference in both individuals and populations and therefore cannot be simply viewed as a source of missing values. Consider a study with four time points in which one of the participants died after the second time point. The data record for this individual is complete; no additional information could be added. Alternatively, consider the perfect scenario in longitudinal studies of older adults in which there is no drop-out other than mortality. One common method to model the data is to use a population-average or marginal curve in order to represent the population trends. However, due to the selection processes occurring over time that change the nature of the population, the population of participants at age 80 is quite different than the population comprised of individuals at age 95. An average curve, obtained from cross-sectional between-person data or aggregate longitudinal data will not represent the behavior of any given individual in this older population. An alternative to the population-average curve is the representation of multiple populations that are conditional on chronological age, time to death, and their interaction.



**Figure 1. Modeling terminal decline.**

Figure 1 shows a hypothetical example of a set of conditional curves for individuals who die at specific ages. Notice that in this hypothetical example, participants were in the study for varying lengths of time based on their age at the initial assessment. To estimate these curves, time to death is introduced as an additional covariate in a linear mixed-effects or semiparametric model. This method will tend to reflect the multiple influences on individual outcomes more accurately, allowing a distinction between the effects of proximity to death and aging in terms of distance from birth (Berg, 1996; Bosworth & Siegler, 2002; see also Sliwinski, Hofer, & Hall, 2003), and permitting modeling of the interaction of age and mortality.

The presence of dropout (*i.e.*, non-mortality-related attrition) in addition to mortality complicates matters because any appropriate missingness model must account for these two separate kinds of missing values. The dropout process right-censors the time of death and masks cognitive functioning between time of dropout and time of death. Currently recommended methods that assume data are *missing at random* (MAR) do not distinguish between missingness due to death and due to dropout. However, a recently developed approach, two-stage multiple imputation (MI), extends single-stage conventional MI by permitting separation of missing values into two types through sequential imputation, and also permits estimation of rates of missing information. Moreover, two-stage MI is useful for applying a framework of assumptions (*i.e.*, models) for two types of missing values (Harel, 2003; Harel, Hofer, & Schafer, 2003; Rubin, 2003; Shen, 2000).

The focus of the present study is on the application of a two-stage multiple-imputation method to obtain inferences in the presence of incomplete data when the primary cause for dropout is mortality-related and when age at death is partially missing. Inferences can then be defined as conditional on the probability of surviving and/or remaining in the study (*e.g.*, DuFouil, Brayne, & Clayton, 2004; Kurland, 2002; Ribaudo, Thompson, & Allen-Mersh, 2000). The rationale for this approach is that it will help to alleviate concerns regarding inference to conditional populations defined by both age and survival by permitting mortality outcomes (age at death) to be modeled when there is incomplete sample follow-up.

## METHOD

### Participants

Participants included 507 community-dwelling elders (178 men, 329 women) drawn from the ongoing longitudinal study, *Origins of Variance*

in the Old-Old (OCTO-Twin Study; McClearn *et al.*, 1997). This is a subsample of the 702 individuals (351 complete same-sex twin pairs; 149 monozygotic [MZ] pairs and 202 same-sex dizygotic [DZ] pairs) who were not diagnosed with dementia over the 6-year study (according to the DSM-III-R criteria; American Psychiatric Association, 1987). Participants were assessed at four different occasions spaced by a 2-year interval. At the baseline occasion of measurement participants ranged in age from 79 to 98 years ( $M = 83.2$ ,  $SD = 2.98$ ) and averaged 7 years of education ( $SD = 2.42$ , range = 0-23). The gender ratio, education, socioeconomic status, marital status, and housing of the OCTO-Twin sample correspond to population statistics for this age segment of the Swedish population (Simmons *et al.*, 1997).

### Measures

The sample was tested with a broad spectrum of biobehavioral measures of health and functional capacity, personality, well-being, and interpersonal functioning. Cognitive functioning was assessed with a widely used Swedish psychometric battery (SRB; Dureman & Sälde, 1959). For purposes of demonstration and analysis in the present study, a test of visuospatial ability was used as the outcome measure (Koh's Block Design I, similar to Block Design subtest in the WAIS battery). In this test (Dureman & Sälde, 1959) the participant is presented with colored blocks and several patterns on cards. The task is to arrange the proper blocks to form the design shown on the card. The maximum score is 42, with higher scores indicating better performance. Scores at all occasions were standardized to a pseudo *t*-score ( $M = 5$ ;  $SD = 1$ ) using the first occasion mean and standard deviation as the standardization base to preserve mean and variance changes over time.

### Statistical Analysis

#### Missing Data

The majority of participant nonresponse over the course of this study was due to mortality. Once mortality is taken into account, less than 10% is due to refusal to participate, mainly related to frailty and compromised health. Based on the work of Rubin (1976; see also Diggle & Kenward, 1994; Diggle, Liang, & Zeger, 1994; Little & Rubin, 1987), participant nonresponse is often categorized as *missing completely at random* (MCAR), *missing at random* (MAR), or *missing not at random* (MNAR). The key distinction is whether the cause of the missingness is related directly to levels of

the outcome variable (MNAR) or whether the missingness is due to other variables that are either irrelevant (MCAR) or measured and included in the statistical model (MAR). Currently used methods for analysis in the presence of incomplete data (*e.g.*, full information maximum likelihood, single-stage multiple imputation) are based on assumptions that responses are *missing at random*, or that the probability of missing information is related to covariates and previously measured outcomes.

In the data set there are 7 variables, 264 observations, and 21 patterns of missing values. Table 1 summarizes the patterns and frequencies of the

**Table 1. Pattern and frequency of missing values by missing value pattern**

Pattern number	Missing value pattern							Number of missing variables	Number of cases
	Predictors			Outcomes					
	Sex	Age at T1	Age at death	BD1	BD2	BD3	BD4		
1	.	.	.	.	.	.	.	0	11
2	.	.	.	.	.	.	<i>m</i>	1	20
3	.	.	<i>m</i>	.	.	.	.	1	85
4	.	.	.	.	<i>m</i> .	.	.	1	2
5	.	.	.	<i>m</i>	.	.	.	1	1
6	.	.	.	.	.	<i>m</i>	<i>m</i>	2	32
7	.	.	<i>m</i>	.	.	.	<i>m</i>	2	7
8	.	.	.	.	<i>m</i>	.	<i>m</i>	2	1
9	.	.	<i>m</i>	.	<i>m</i>	.	.	2	2
10	.	.	.	<i>m</i>	.	.	<i>m</i>	2	1
11	.	.	<i>m</i>	<i>m</i>	.	.	.	2	1
12	.	.	.	<i>m</i>	<i>m</i>	.	.	2	1
13	.	.	<i>m</i>	.	.	<i>m</i>	<i>m</i>	3	5
14	.	.	.	.	<i>m</i>	<i>m</i>	<i>m</i>	3	43
15	.	.	<i>m</i>	.	<i>m</i>	.	<i>m</i>	3	2
16	.	.	.	<i>m</i>	.	<i>m</i>	<i>m</i>	3	5
17	.	.	<i>m</i>	.	<i>m</i>	<i>m</i>	<i>m</i>	4	10
18	.	.	<i>m</i>	<i>m</i>	.	<i>m</i>	<i>m</i>	4	1
19	.	.	.	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	4	29
20	.	.	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	.	4	1
21	.	.	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	5	4
Number missing	0	0	118	44	95	130	160		
Percent missing	0	0	45	17	36	49	61		

*Note.*  $N = 264$ . BD = Block Design. A dot represents an observed value whereas *m* represents missing values.



missing values by variable. Five variables have at least one missing value (a dot represents an observed value while  $m$  represents missing values). There were 253 observations with at least one missing value; 85 participants have age, sex, and all outcomes observed but missing time of death.

### Dependencies Related to Twin Pair

In the longitudinal analyses the twins were treated as unrelated singletons. Therefore, it is important to note that the parameters obtained in the analysis will be unbiased estimates of the population parameters. However, under conditions of positive and moderate to substantial intraclass correlations, standard errors of these parameters will be downwardly biased due to the dependencies associated with analyzing co-twins simultaneously. In these octogenarian twins, the intraclass association at the first occasion is low to moderate (*e.g.*, 0.51 combined sample) and decreases across occasions, and thus would have little to no influence on the significance of the results regarding magnitude of change (see Johansson *et al.*, 2004).

### Imputation Model

Conventional single-stage multiple imputation (MI; Rubin, 1987) replaces the missing values in a data set by  $m > 1$  sets of simulated values. Two-stage MI (Harel, 2003) is an extension in which the missing values are partitioned into two groups and imputed  $N = mn$  times in a nested fashion, and point estimates and standard errors from the  $mn$  complete data analyses are consolidated by simple rules derived by analogy to nested analysis of variance. The choice of  $m$  and  $n$  depends on the inference in mind, as discussed in Rubin (1987) and extended in Harel (2003). A relatively small number of imputations ( $m$  in single-stage MI and  $N = mn$  in two-stage MI) is sufficient when one is interested in the estimates and their standard errors, as was used in the present study.

Let  $Y_{com} = (Y_{obs}, Y_{mis})$  denote the complete data set, which in the present example includes scores from up to four block design measurements every 2 years. Block design is treated as a continuous variable. Predictors of initial status and change in block design included chronological age, sex, and time to death, as calculated as the number of years from the age at the initial measurement occasion to the age at time of death for each individual. Age and time to death were treated as categorical due to the limited number of categories. The missing data are then separated into two types: let the missing time of death be denoted as  $Y_{mis}^A$  and missing cogni-

tive functioning measures be  $Y_{mis}^B$ . using an extended theory of ignorability presented by Harel (2003), one can show that the processes that cause the missing values can be ignored in two-stage MI if (a)  $Y_{mis}$  is MAR, and (b) the process that divides  $Y_{mis}$  into  $Y_{mis}^A$  and  $Y_{mis}^B$  does not depend on any portion of  $Y_{mis}$ . Other conditions under which part or all of the missingness mechanism can be ignored in one or both stages of two-stage MI are described by Harel (2003).

The imputation model used belongs to the family of general location models for multivariate continuous and categorical data (Little & Schluchter, 1985; Olkin & Tate, 1961). The general location model combines a loglinear model for the categorical variables with a multivariate normal linear regression for the continuous variables given the categorical variables. We expect measures of block design to be related not only to a participant's age and sex, but also to time to death. A person may experience a decline in cognitive performance prior to death that is more dramatic than a gradual decline often associated with aging. Therefore, it is crucial that the imputation model preserve interactions between age and time to death for block design.

The loglinear part of the model included all associations among age, sex, and time to death, the linear part of the model allowed the means of the four block design scores to depend on age, sex, and time to death and all interactions among them, with unstructured residual covariances among the block design measures. Under this model, we generated 20 multiple imputations by the Markov chain Monte Carlo procedure described by Schafer (1997) using Splus version 6.0. More specifically, due to the structure of the data (both categorical and continuous variables), we used the Conditional Gaussian Model in the missing data library to perform the imputations. To the best of our knowledge Splus is the only commercial software that permits the use of the Conditional Gaussian Model. Other packages (e.g., SAS Proc MI) support the statistical modeling of only continuous variables. We then removed the imputed values for block design measures from each imputed data set and reimputed them once, treating the imputed values for time to death as fixed, to produce  $mn = 40$  completed data sets in a nested design with  $m = 20$  blocks of size  $n = 2$ . As a final step, any imputed values for the block design measures beyond death were removed, reflecting the fact that they logically cannot exist.

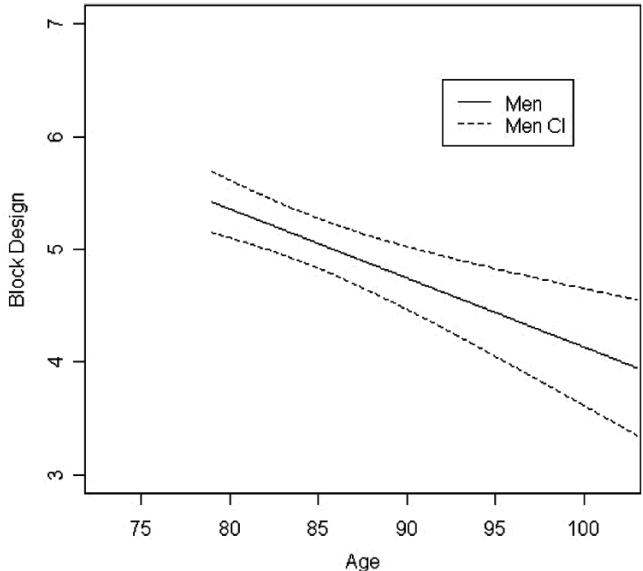
## RESULTS

Two-stage multiple imputations were performed and two descriptive analyses were fit to the imputed data. The first model estimated average population curves for the relationship between age and block design for

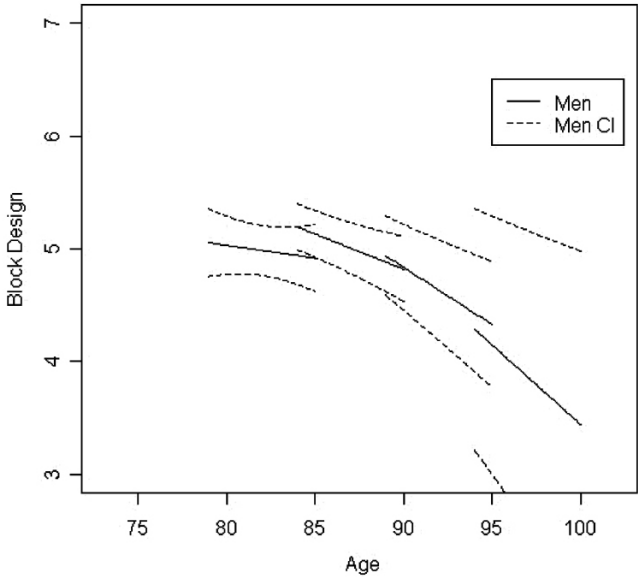
men and women, and included an intercept, age, sex, and the interaction of sex with age. The second model estimated population curves conditional on time to death, and included an intercept, age, time to death, sex, and the interactions Age  $\times$  Time to Death, Age  $\times$  Sex, Sex  $\times$  Time to Death, and Sex  $\times$  Age  $\times$  Time to Death. Quadratic effects were nonsignificant and were thus excluded. Several alternative structures to account for correlations among time points were tested and an autoregressive (AR-1) structure was chosen for both sets of models. Maximum likelihood estimates and standard errors were computed for each imputed data set, and the results were combined by Shen's (2000) rules (see Appendix 1).

Due to the fact that there are several parameter estimates for each curve, and multivariate theory for combining two-stage multiple-imputation estimates is not yet available, we used the point-wise estimates and 95% confidence intervals for producing plots of expected trajectories. For given covariate levels (sex, age, time to death) we estimated block design scores and their variance from the 40 complete data sets. These univariate results were combined using Shen's (2000) rules. For a given sex and time of death we created a grid of values (6 years) for age, which will represent the trajectory for this specific time of death. Results are shown in Figures 2a and b for men and Figures 3a and b for women, in which age was plotted on the  $x$ -axis from 79 to 102 years, with intervals of 0.1 for the marginal curves. The choices for the specific trajectories are arbitrary; the researcher can decide the specific trajectories of interest to be plotted.

Estimated average block design trajectories for men and women, along with point-wise 95% confidence intervals, are shown in Figures 2a and 3a, respectively. Block design performance declined with age, and the confidence intervals about this decline are noticeably narrower at younger ages in which a greater number of individuals contributed to the estimates. The trajectory for women lies above the trajectory for men, and the sex difference increased with age. Similarly, the conditional population curves with time to death included as a covariate for men (Figure 2b) and women (Figure 3b) were computed by using the point-wise combining rules. The conditional curves are plotted for 6 years, the longest possible interval for a given individual (*i.e.*, four time points at 2-year intervals). For example, in the case where an individual entered the study at age 80, with completed data at 80, 82, 84, and died at age 85 (age at death is either known or imputed), the age-based information will be from 79 to 85 years of age. Using the conditional curves we can see the change of slope as the individuals age (both men and women), such that steeper slopes are expected for older individuals. Again, the confidence intervals are noticeably narrower for younger ages reflecting where there is greater sample size.

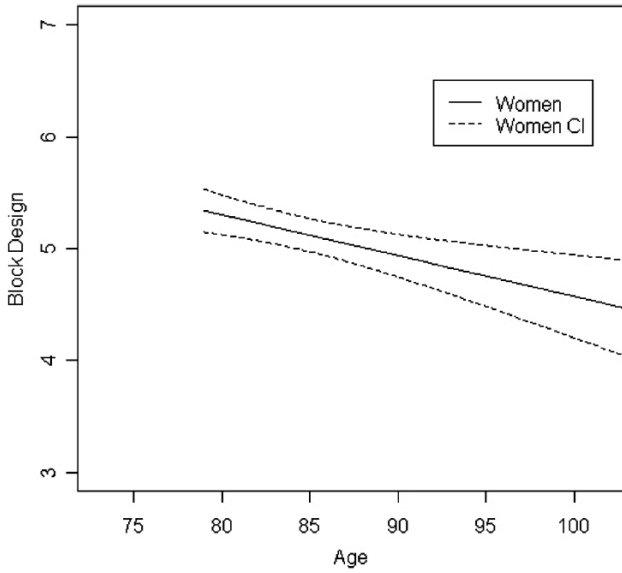


(a)

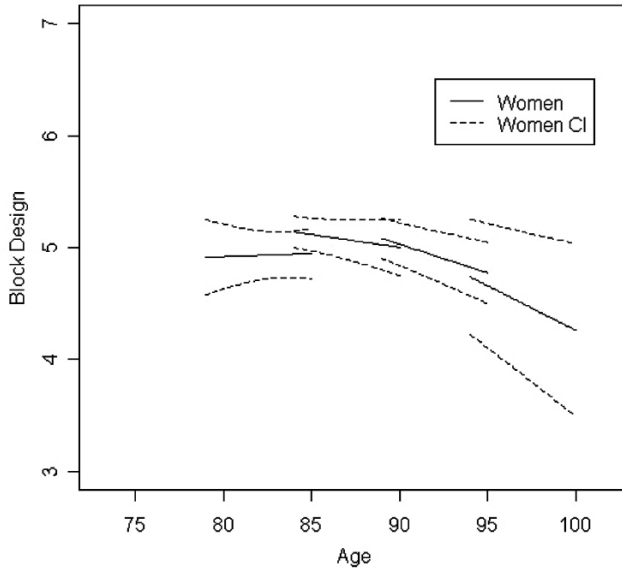


(b)

**Figure 2.** (a) Estimated block design change trajectories and 95% point-wise confidence intervals for men. (b) Estimated block design change trajectories and 95% point-wise confidence intervals conditional on time of death for men.



(a)



(b)

**Figure 3.** (a) Estimated block design change trajectories and 95% point-wise confidence intervals for women. (b) Estimated block design change trajectories and 95% point-wise confidence intervals conditional on time of death for women.

## DISCUSSION

At the population level, causal influences on individual age-related change in cognitive and functional abilities are likely to be highly heterogeneous (*e.g.*, aging, dementia, health). In longitudinal studies on aging, attrition is a function of not only dropout, but also of mortality. These differential causes of missingness are not well addressed by current approaches of analysis that make general assumptions regarding dropout (MI-MAR) and therefore treat missingness due to death as ignorable. In this paper we present an application of a two-stage multiple-imputation (MI) approach in which time to death is imputed in one step, and this information is then used to improve imputed estimates of missing outcomes in a second step. The two-stage MI procedures employed here thus permit the partial distinction of the effects of age (distance from birth) and mortality (distance from death), respectively, and permit a more realistic description of change as a joint function of chronological age and mortality.

When information on time to death is complete, a single-stage MI (conventional MI) procedure or full information maximum likelihood approach is appropriate and will permit conditional estimation as shown here (*e.g.*, Pederson *et al.*, 2003). The two-stage MI procedure is most likely to be useful when date of death or other important covariates are unknown for a portion of the longitudinal sample, because cases with missing covariates are omitted from analysis in most multilevel modeling programs, and interactions among covariates with missing values remain difficult to model in structural equation modeling approaches. In particular, when both age and time to death (*i.e.*, known age at death) are included in the model, the two-stage MI permits computation of the interaction of age and time to death.

A related issue is the extent to which inferences to a single population of aging individuals is possible, given that the population composition is changing over time due to selective mortality. As such, population average or marginal curves are not likely to be reflective of individual outcomes as intended and described what has been referred to as an "immortal" population (DuFouil *et al.*, 2003). We instead recommend the use of predicted trajectories in which inferences are defined as conditional on the probability of surviving and/or remaining in the study. As an illustration we compared estimated marginal curves of predicted declines in visual spatial abilities (as measured by block design scores) in older age with change trajectories conditioned not only on age and sex, but also on time to death. Trajectories from the conditional models are likely to be more accurate than those from marginal models when time to death plays a role in magnitude of decline in older age (*i.e.*, terminal drop phenome-

non), and allows us to better assess and distinguish the multiple factors of age-related decline. For example, marginal trajectories would not provide sufficient information for individual participants relative to age and life span. However, the conditional expected trajectory provides information for similar participants of age 84 with a maximum life span of 90 years. This two-stage MI procedure permits such conditional analysis in both SEM and MLM analysis approaches, and is likely to be useful in other applications in which multiple time-dependent developmental processes underlie individual and population change.

#### APPENDIX 1: SHEN'S RULES

Let  $Q$  denote a scalar quantity to be estimated (e.g., a population mean, regression coefficient or odds ratio), and  $\hat{Q} = \hat{Q}(Y_{obs}, Y_{mis})$  be the estimate for  $Q$  if complete data were available. Also let  $U = U(Y_{obs}, Y_{mis})$  denote its squared standard error. With complete data, we assume tests and intervals based on the normal approximation  $(\hat{Q}-Q)/\sqrt{U} \sim N(0,1)$  would be appropriate. From each of the  $N = mn$  completed data sets, we calculate the point and variance estimates  $(\hat{Q}^{(j,k)}, U^{(j,k)})$ ,  $j = 1 \dots m$ ;  $k = 1 \dots n$ . The overall point estimate is

$$\bar{Q}_{..} = \frac{1}{mn} \sum_{j=1}^m \sum_{k=1}^n \hat{Q}^{(j,k)} = \frac{1}{m} \sum_{j=1}^m \bar{Q}_j.$$

where  $\bar{Q}_j = n^{-1} \sum_k \hat{Q}^{(j,k)}$ . The uncertainty in  $\bar{Q}_{..}$  arises from three components: the estimated complete-data variance

$$\bar{U}_{..} = \frac{1}{mn} \sum_{j=1}^m \sum_{k=1}^n U^{(j,k)}$$

the between-block imputation variance

$$B = \frac{1}{m-1} \sum_{j=1}^m (\bar{Q}_j - \bar{Q}_{..})^2$$

and the within-block imputation variance

$$W = \frac{1}{m} \sum_{j=1}^m \frac{1}{n-1} \sum_{k=1}^n (\hat{Q}^{(j,k)} - \bar{Q}_j)^2.$$

The total variance is

$$T = \bar{U}.. + (1 - 1/n)W + (1 + 1/m)B$$

and inferences are based on the approximation  $T^{-1/2} (Q - \bar{Q}..) \sim t_v$  with degrees of freedom

$$v = \frac{1}{m(n-1)} \left( \frac{(1 - 1/n)W}{T} \right)^2 + \frac{1}{m-1} \left( \frac{(1 + 1/m)B}{T} \right)^2.$$

When  $n = 1$ ,  $W$  becomes undefined and drops out of the total variance and the method reduces to Rubin’s rules for conventional MI.

REFERENCES

American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd edition revised). Washington, DC: Author.

Baltes, P. B. (1968). Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development*, 11, 145-171.

Berg, S. (1987). Intelligence and terminal decline. In G. L. Maddox & E. W. Busse (Eds.), *Aging: The Universal Human Experience* (pp. 411-416). New York: Springer.

Berg, S. (1996). Aging, behavior and terminal decline. In J. Birren & W. Schaie (Eds.), *The Handbook of the Psychology of Aging* (4th ed., pp. 323-336). New York: Academic Press.

Bosworth, H. B. & Siegler, I. C. (2002). Terminal change in cognitive function: An updated review of longitudinal studies. *Experimental Aging Research*, 28, 299-315.

Cooney, T. M., Schaie, K. W., & Willis, S. L. (1988). The relationship between prior functioning on cognitive and personality dimensions and subject attrition in longitudinal research. *Journal of Gerontology*, 43, 12-17.

Deeg, D. J. H., Hofman, A., & van Zonneveld, R. J. (1990). The association between change in cognitive function and longevity in Dutch elderly. *American Journal of Epidemiology*, 132, 973-982.



- Diggle, P. J. & Kenward, M. G. (1994). Informative dropout in longitudinal data analysis (with discussion). *Applied Statistics*, 43, 49-94.
- Diggle, P. J., Liang, K., & Zeger, S. L. (1994). *Analysis of Longitudinal Data*. Oxford: Clarendon Press.
- DuFouil, C., Brayne, C. & Clayton, D. (2004). Analysis of longitudinal studies with death and dropout: A case study. *Statistics in Medicine*, 23, 2,215-2,226.
- Dureman, I. & Sälde, H. (1959). *Psykometriska och Experimental-Psykologiska Metoder för Klinisk Tillämpning*. Uppsala, Sweden: Almqvist & Wiksell.
- O. Harel et al. Harel, O. (2003). *Strategies for Data Analysis with Two Types of Missing Values*. (Tech. Rep. No. 03-60). University Park, PA: The Methodology Center. The Pennsylvania State University.
- Harel, O., Hofer, S. M., & Schafer, J. L. (2003, March). Analysis of longitudinal data with missing values of two qualitatively different types. Paper presented at the annual meeting of the International Biometric Society (Eastern North American Region), Tampa, Florida.
- Harel, O. (2007). Inferences on missing information under multiple imputation and two-stage multiple imputation. *Statistical Methodology*, 4, 75-89.
- Hofer, S. M. & Hoffman, L. (2004). Statistical analysis with incomplete data: A developmental perspective. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling Developmental Processes in Ecological Context*. Mahwah, NJ: LEA.
- Hofer, S. M. & Sliwinski, M. J. (2006). Design and analysis of longitudinal studies of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the Psychology of Aging* (6th ed., pp. 15-37). San Diego: Academic Press.
- Johansson, B. & Berg, S. (1989). The robustness of the terminal decline phenomenon: Longitudinal data from the digit-span memory test. *Journal of Gerontology: Psychological Sciences*, 44, P184-P186.
- Johansson, B., Hofer, S. M., Allaire, J. C., Maldonado-Molina, M., Piccinin, A. M., Berg, S., Pedersen, N., & McClearn, G. E. (2004). Change in memory and cognitive functioning in the oldest-old: The effects of proximity to death in genetically related individuals over a six-year period. *Psychology and Aging*, 19, 145-156.
- Kurland, B. F. (2002). *Analysis of Binary Longitudinal Data with Dropout and Death*. PhD Dissertation. Department of Biostatistics, University of Washington, Seattle, WA.
- Little, R. J. A. & Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. New York: Wiley.
- Little, R. J. A. & Schluchter, M. D. (1985). Maximum likelihood estimation for mixed continuous and categorical data with missing values. *Biometrika*, 72, 497-512.
- McClearn, G. E., Johansson, B., Berg, S., Pedersen, N. L., Ahern, F., Petrill, S. A., & Plomin, R. (1997). Substantial genetic influence on cognitive abilities in twins 80+ years old. *Science*, 276, 1,560-1,563.
- Olkin, I. & Tate, R. F. (1961). Multivariate correlation models with mixed discrete and continuous variables. *The Annals of Mathematical Statistics*, 32, 448-465.
- Pederson, N. L., Ripatti, S., Berg, S., Reynolds, C., Hofer, S. M., Finkel, D. Gatz, M., & Palmgren, J. (2003). The influence of mortality on twin models of change: Addressing missingness through multiple imputation. *Behavior Genetics*, 33(2), 161-169.

- Rabbitt, P. Watson, P., Donlan, C., Bent, N., & McInnes, L. (1994). Subject attrition in a longitudinal study of cognitive performance in community-based elderly people. *Facts and Research in Gerontology*, 29-34.
- Ribaudo, H. J., Thompson, S. G., & Allen-Mersh, T. G. (2000). A joint analysis of quality of life and survival using a random effect selection model. *Statistics in Medicine*, 19, 3,237-3,250.
- Riegel, K. F., Riegel, R. M., & Meyer, G. (1967). A study of the drop-out rates in longitudinal research on aging and the prediction of death. *Journal of Personality and Social Psychology*, 4, 342-348.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 61, 581-592.
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- Rubin, D. B. (2003). Nested multiple imputation of NMES via partially incompatible MCMC. *Statistica Neerlandica*, 57(1), 3-18.
- Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. New York: Chapman and Hall.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147-177.
- Schaie, K. W., Labouvie, G. V., & Barrett, T. J. (1973). Selective attrition effects in a fourteen-year study of adult intelligence. *Journal of Gerontology*, 28, 328-334.
- Shen, Z. J. (2000). *Nested Multiple Imputation*. PhD. Dissertation. Department of Statistics, Harvard University, Cambridge, MA.
- Siegler, I. C. (1975). The terminal drop hypothesis: Factor or artifact. *Experimental Aging Research*, 1, 169-185.
- Siegler, I. C. & Botwinick, J. (1979). A long-term longitudinal study of intellectual ability of older adults: The matter of selective subject attrition. *Journal of Gerontology*, 34, 242-245.
- Simmons, S. F., Johansson, B., Zarit, S. H., Ljungquist, B., Plomin, R., & McClearn, G. E. (1997). Selection bias in samples of older twins? A comparison between octogenarian twins and singletons in Sweden. *Journal of Aging & Health*, 9, 553-567.
- Sliwinski, M. J., Hofer, S. M., & Hall, C. (2003). Correlated and coupled cognitive change in older adults with and without clinical dementia. *Psychology and Aging*, 18, 672-683.
- Small, B. J. & Bäckman, L. (1999). Time to death and cognitive performance. *Current Directions in Psychological Science*, 8, 168-172.
- Smits, C. H. M., Deeg, D. J. H., Kriegsman, D. M. W., & Schmand, B. (1999). Cognitive functioning and health as determinants of mortality in an older population. *American Journal of Epidemiology*, 150, 978-986.
- Streib, G. F. (1966). Participants and drop-outs in a longitudinal study. *Journal of Gerontology*, 21, 200-209.