ASSOCIATION BETWEEN LONGITUDINAL CHANGES IN GAIT SPEED AND SURVIVAL TIME IN AGING MEN:

A JOINT LONGITUDINAL AND TIME-TO-EVENT ANALYSIS OF DATA FROM THE MrOS STUDY

Kyle Damien Hart

A THESIS

Presented to the Department of Public Health & Preventive Medicine

and the Oregon Health & Science University School of Medicine

in partial fulfillment of the requirements for the degree of

Master of Science

April 2014

Department of Public Health & Preventive Medicine
School of Medicine
Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the Master's thesis of
Kyle D Hart
has been approved

Jodi Lapidus, PhD (Mentor/Advisor)

Carrie Nielson, MPH, PhD (Committee Member)

Mara Tableman, PhD (Committee Member)

TABLE OF CONTENTS

List of Tables

List of Figures

1 Acknowledgements

2 Abstract

Slower gait speed has previously been shown to be associated with mortality in aging populations, based on a single measure of gait speed,[1–3] but repeated measures of gait speed may be more informative. Although an extended Cox model can accommodate a time-varying covariate, the Cox framework has certain limitations in context of endogenous or continuous time-varying covariates or time-varying covariates measured with error.

The joint modeling framework uses a smooth, continuous model of the time-varying covariate that may more accurately represent its true, underlying longitudinal trajectory. The framework also incorporates measurement error of the time-varying covariate and uses a likelihood formulation that is consistent with endogenous covariates (i.e., covariates whose future trajectories after the occurrence of the event of interest are altered by the event). The chief benefits of these improvements, substantiated by simulation study,[4] is that joint models are not subject to underestimation of the effect size and standard error, as may occur in extended Cox models. Also, joint models have less stringent assumptions about missing data mechanisms than linear mixed-effects models, so they can correct bias in longitudinal estimates that may result from non-ignorable missingness.

The objective of this thesis is to estimate the association between longitudinal trajectories of gait speed and survival time and to compare estimates of association from separate models (i.e., mixed-effects models and Cox models) and joint models.

A subset of 877 ambulatory, community-dwelling older men from 2 of the 6 sites in the Osteoporotic Fractures in Men (MrOS) study performed a walking test up to 5 times over a median follow-up time of 7 years and were followed for a median of 11 years for mortality. We modeled the hazard ratio (HR) of gait speed 1) as a baseline measure alone, 2) as a time-varying covariate (extended Cox), and 3) as a longitudinal sub-model using linear mixed-effects with cubic natural splines (joint model).

Slower gait speed was associated with mortality in all models. The HR per 0.1 m/s decline in gait speed was 1.08 in the Cox model (95% CI: 1.01 to 1.15); 1.14 (95% CI: 1.05 to 1.22) in the extended Cox model; and 1.25 (95% CI: 1.15 to 1.36) in the best-fitting joint model. Estimates of longitudinal parameters from the linear mixed-effects vs. joint model suggested estimation was not sensitive to missingness assumptions.

As expected, the extended Cox model underestimated the effect of longitudinal gait speed on survival time. Estimates of longitudinal parameters were similar across modeling strategy, suggesting that the longitudinal process was not sensitive to missingness assumptions. Contrary to expectation, standard errors for both longitudinal and event parameters were

very similar across all modeling strategies.

Providers may benefit by considering the increased estimate of the association between gait speed and survival time. Traditional modeling techniques may underestimate the magnitude of this association.

3 Introduction

Life expectancy is a particularly important consideration in treatment plans for older adults, so finding predictors of mortality in older populations is a clinically important endeavor. A 6-meter walking test is a convenient physical test that can be administered in any clinic setting with a 6-meter-long space, such as a hallway, without the need for specialized equipment or training. The general procedure is to measure the time it takes for the patient to walk a 6-meter course at his or her usual walking pace.[5]

Gait speed has previously been shown to be associated with mortality in older populations: A 2009 prospective cohort study examined walking speed and mortality in 3 208 community-dwelling adults over 65 years old and found a hazard ratio for all-cause mortality of 1.44 (95% CI: 1.03 to 1.99), comparing the slowest third to the faster two thirds.[1] For cardiovascular death specifically, the hazard ratio was 2.92 (95% CI: 1.46 to 5.84); there was no significant difference in cancer death. In a recent meta-analysis of the association between physical capability measures and all-cause mortality among 5 cohorts (N=14 692) of older adults, a summary hazard ratio, comparing the slowest-walking quartile to the fastest-walking quartile, was 2.87 (95% CI: 2.22 to 3.72), adjusted for age, sex, and body size.[2] In 2011, Studenski et al published a pooled analysis of 34 485 community-dwelling adults over 65 years old from 9 cohort studies that measured gait speed at baseline and followed patients for mortality for at least 5 years. Estimates of hazard ratios ranged from 1.06 to 1.20 per 0.1 meter-per-second decline in gait speed; the pooled estimate was 1.14 (95% CI: 1.11 to 1.15).[3]

In the discussion of their meta-analysis, Cooper et al point out that new research is needed to "examine the associations between changes in capability with age and mortality, as a steep decline in physical capability may be a better predictor of mortality than is the absolute level at a single point in time.[2]" This reasoning makes good intuitive sense, especially for older patients who are in the phase of life where one can expect to observe changes in physical capability over a reasonably short follow-up period, and where a sudden decline in mobility may be especially predictive of mortality. In that spirit, the current investigation endeavors to measure the association between longitudinal changes in gait speed and mortality and to adjust for this association in estimates of the effects of other predictors of both survival time and the longitudinal trajectory of gait speed.

Survival time and longitudinal trajectories frequently have a non-ignorable relationship. In the case of survival time (or time to some event other than death), there is frequently at least one important covariate that changes over time during follow-up. Accounting for longitudinal changes in time-varying covariates may facilitate better prediction of survival

time than accounting for cross-sectional covariates at baseline alone. On the other hand, when the longitudinal trajectory is the outcome of interest, failing to account for systematic patterns in dropouts may seriously bias estimates of the effects of predictors.[6,7]

In both cases, it may be useful to measure the association between a longitudinal process and a time-to-event process. Traditional methods can be used to model each process separately, but unmeasured association in a separate longitudinal model tends to induce bias.[4] The Cox proportional-hazards model can be extended to accommodate a time-varying covariate, but this framework is not optimal because it does not properly account for longitudinal covariates whose future trajectory has a probabilistic relationship to occurrence of the event of interest (i.e., "endogenous covariates"); it models even continuous time-varying covariates as a step function; and it assumes that longitudinal covariates are measured without error. Joint models for longitudinal and time-to-event data provide a framework for accounting for the effects of endogenous, continuous, longitudinal trajectories, measured with error, on an event process.

In his book on joint modeling, Rizopoulos[8] predicts that, compared to joint models, a Cox proportional-hazards model with a time-varying covariate will underestimate the magnitude of association between the longitudinal and time-to-event process and will underestimate the standard error for longitudinal, event, and association parameters. This can be taken to mean that, in certain circumstances, joint models will provide more realistic estimates than traditional methods.

This thesis explores the emerging method of joint longitudinal and time-to-event modeling by first applying it to a didactic dataset of longitudinal CD4 lymphocyte counts and survival time from a study of acquired immunodeficiency syndrome (AIDS) patients. This dataset has been used repeatedly in the literature on joint modeling,[8–12] so it provides a good didactic opportunity.

Second, the motivating example for the project is to examine the relationship between longitudinal changes in gait speed and survival time in a population of older men. Gait speed has previously been shown to be related to mortality[2,3] and represents a measure of endogenous endurance whose underlying value is continuously changing, so it is a good candidate for joint modeling.

The Osteoporotic Fractures in Men (MrOS) study is a long-duration longitudinal cohort study that examines patterns of osteoporotic fractures and a variety of other clinical indicators in older men.[5] A subset of MrOS participants who participated in two ancillary studies of periodontal health[13] and sleep[14,15] performed a 6-meter walking test up to 5 times over a median follow-up period of 7 years. Data were also collected on a variety of other clinical factors, and participants have, at the time of this writing, been followed for

2

survival over a median time of 11 years. To my knowledge, longitudinal trajectories of gait speed in this cohort have not previously been modeled, and joint modeling has not yet been applied to data from this study.

This paper is organized as follows: In section 4, we describe notation and theory relevant to joint longitudinal and time-to-event models, including a brief review of linear mixed-effects models and Cox proportional-hazards models. In section 5, we describe separate and joint models of longitudinal trajectories of CD4 lymphocyte counts, survival time, and the association between the two outcomes in the didactic AIDS dataset. In section 6, we describe separate and joint models of longitudinal changes in gait speed, survival time, and the association between longitudinal gait speed and survival time in the MrOS cohort. The paper concludes with a discussion of the results of the joint model of gait speed and survival.

4 Theory

4.1 Linear mixed-effects model

Traditional mixed-effects models provide a framework for regression modeling when the crucial assumption of independent observations is violated. This is typical when a study design calls for repeated measures on individual subjects, because one can expect to observe correlation among measurements taken from the same individual, even if those measurements are separated by time.

Let's begin a description of the model by defining some notation:

- $y_i(t)$ is the observed response for individual i at time t.

- $\boldsymbol{\beta}$ is a vector of population-level coefficients corresponding to fixed effects $\mathbf{x}_i$.

- $\mathbf{x}_i(t)$ is a vector of fixed effects for individual i at time t.

- $\mathbf{b}_i$ is a vector of individual-level coefficients corresponding to random effects $\mathbf{z}_i$.

- $\mathbf{z}_i(t)$ is a vector of random effects for individual i at time t.

- $\varepsilon_i(t)$ is the residual error for individual i at time t, conditional on the random effects $\mathbf{b}_i$.

3

- Σ is the covariance matrix whose dimensions correspond to the length of the vector $\mathbf{b}_i$.

- $\sigma^2 I_{n_i}$ is the covariance matrix whose n_i dimensions correspond to the number of longitudinal observations for individual i.

Assuming a normal distribution for the outcome $y_i(t)$, a linear mixed-effects model of form (1) can account for within-individual correlation by introducing in the linear model one or more sets of predictors and corresponding regression coefficients that are allowed to vary randomly across individuals.[16] The model reflects the idea that each individual has his or her own individual-specific response profile over time that is offset from the mean response profile.

$$
\begin{cases}
y_i(t) = \boldsymbol{\beta}\mathbf{x}_i^\top(t) + \mathbf{b}_i\mathbf{z}_i^\top(t) + \varepsilon_i(t), \\
\mathbf{b}_i \sim \mathcal{N}(0, \Sigma), \\
\varepsilon_i \sim \mathcal{N}(0, \sigma^2 I_{n_i})
\end{cases}
\tag{1}
$$

In model (1), the vector of fixed effects $\mathbf{x}_i$, the vector of random effects $\mathbf{z}_i$, and the residual error term ε_i can all be functions of time, written as (t).

A special case of the linear mixed-effects model (a popular choice for correlated observations) is a model allowing individual-level random intercepts and slopes. In the case of time as a covariate, this popular model specification is

$$
\begin{cases}
y_{ij} = \beta_0 + \beta_1 t_{ij} + b_{i0} + b_{i1}t_{ij} + \varepsilon_{ij}, \\
b_{i0},\ b_{i1} \sim \mathcal{N}(0, \Sigma), \\
\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)
\end{cases}
\tag{2}
$$

Model (2) uses a slightly different notation: the subscripts i and j refer to the ith individual and the jth longitudinal observation. Then y_{ij} denotes the outcome for individual i at observation j, and the predictor t_{ij} is the corresponding time.

4

4.1.1 Estimation

The observations $y_i(t)$ are assumed to follow a multivariate normal distribution with mean $\boldsymbol{\beta}\mathbf{x}_i$ and variance-covariance matrix $V_i = \mathbf{z}_i \Sigma \mathbf{z}_i^\top + \sigma^2 I_{n_i}$. So the joint density of all n_i observations of $y_i(t)$ for individual i is

$$
\begin{cases}
p(y_i|\boldsymbol{\theta}) = \dfrac{1}{\sqrt{(2\pi)^{n_i}|V_i|}} \exp\left(\dfrac{-(y_i - \boldsymbol{\beta}\mathbf{x}_i)^\top (y_i - \boldsymbol{\beta}\mathbf{x_i})}{V_i}\right), \\[2ex]
\boldsymbol{\theta}^\top = (\boldsymbol{\beta}^\top, \sigma^2, vech(\Sigma)), \\[2ex]
V_i = \mathbf{z}_i \Sigma \mathbf{z}_i^\top + \sigma^2 I_{n_i}
\end{cases}
\tag{3}
$$

and the log-likelihood for all $y_i(t)$ is

$$
\ell(\boldsymbol{\theta}|y_i) = -\frac{K}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{N}\log|V_i| - \frac{1}{2}\left(\sum_{i=1}^{N}\frac{(y_i - \boldsymbol{\beta}\mathbf{x}_i)^\top(y_i - \boldsymbol{\beta}\mathbf{x_i})}{V_i}\right)
\tag{4}
$$

where $K = \sum_{i=1}^{N} n_i$ is the total number of observations for the N individuals with n_i observations each. Since the parameter $\boldsymbol{\beta}$ appears only in the right-most term, maximizing the log-likelihood with respect to $\boldsymbol{\beta}$ amounts to minimizing

$$
\sum_{i=1}^{N}\frac{(y_i - \boldsymbol{\beta}\mathbf{x}_i)^\top(y_i - \boldsymbol{\beta}\mathbf{x_i})}{V_i},
\tag{5}
$$

and the MLE for $\boldsymbol{\beta}$ is

$$
\hat{\beta} = \left(\sum_{i=1}^{N}(\mathbf{x}_i^\top V_i^{-1}\mathbf{x}_i)\right)^{-1}\left(\sum_{i=1}^{N}(\mathbf{x}_i^\top V_i^{-1}y_i)\right).
\tag{6}
$$

A similar method is employed to estimate the variance parameters, although numerical techniques are necessary, since a closed-form solution usually doesn't exist.

Although the method of maximum likelihood yields asymptotically unbiased estimates of V_i, finite-sample estimates are biased, as a result of using maximum-likelihood estimates for both parameters. The method of restricted maximum likelihood (REML) yields an

unbiased estimate of V_i in finite samples by adding a correction term to the likelihood function (7) to account for the fact that β has been estimated:

$$\ell_{REML}(\boldsymbol{\theta}|y_i) = -\frac{1}{2}\log\left|\sum_{i=1}^{N}\frac{\mathbf{x}_i^\top \mathbf{x}_i}{V_i}\right| - \frac{1}{2}\sum_{i=1}^{N}\log|V_i| - \frac{1}{2}\left(\sum_{i=1}^{N}\frac{(y_i - \beta\mathbf{x}_i)^\top(y_i - \beta\mathbf{x}_i)}{V_i}\right), \quad (7)$$

whose maximization with respect to V_i is again carried out by numerical optimization.[7]

4.1.2   Shape of the longitudinal trajectory

As in other forms of of linear regression, mixed-effects models make use of linear or polynomial shapes to model the shape of the relationship between the outcome variable and the vector of independent variables—of course, in the case of a mixed-effects model of longitudinal data, at least one of the independent variables is time.

A drawback to the use of polynomial shapes is that nature sometimes fails to provide us with natural relationships that are perfectly polynomial. For this reason, it is sometimes necessary to employ more flexible techniques, such as piecewise polynomials, basis splines (b-splines), or natural cubic splines.

The piecewise methods under consideration here are functions that allow a separate linear specification for each of several intervals in the modeling region, separated by K internal boundaries called *knots*. Let the internal knots be called $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_K)$. Then the general form of our piecewise methods is

$$f(x) = \sum_{d=1}^{m} \kappa_d h_d(x) \tag{8}$$

where κ are coefficients for the functions $(h_1, \ldots, h_m)$. The function $h_d(\cdot)$ can be defined as any relevant functional form of X.

Piecewise constant

A simple piecewise approach is a piecewise-constant function, which amounts to a shifting least-squares mean over however many intervals are specified. Consider a region broken into 3 intervals by $K = 2$ internal knots. The piecewise-constant function is as specified in (8), with $m = K + 1$ and with h_d defined as

6

$$\begin{cases} h_1(X) = I(X < \xi_1), \\ h_2(X) = I(\xi_1 \le X < \xi_2), \\ h_3(X) = I(X \ge \xi_2) \end{cases} \tag{9}$$

and the least-squares estimate of $f(x)$ is the mean of Y in the dth interval: $\hat{\kappa} = \overline{Y}_d$.

Figure 1. Illustration of piecewise-constant approach with 2 internal knots ξ_1 and ξ_2.
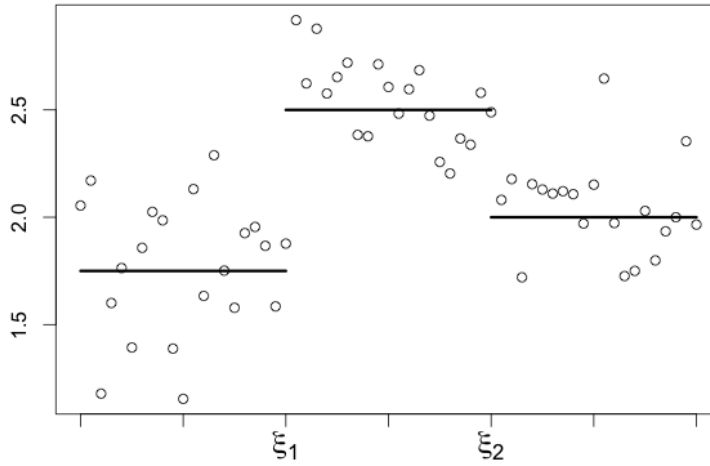
Piecewise polynomial

Constraints can be added to require continuity of the 0th, 1st, 2nd, 3rd, or higher-order derivatives. Enforcing 0th-derivative continuity connects the intervals without any smoothing; enforcing continuity at higher-order derivatives will result in increasingly smooth knots. To produce a set of splines with slopes that are continuous at the knots, we would define h_d as

$$\begin{cases} h_1(X) = 1, \\ h_2(X) = X, \\ h_3(X) = (X - \xi_1)_+, \\ h_4(X) = (X - \xi_2)_+ \end{cases} \tag{10}$$

where the subscript $+$ indicates that we are including only the positive portion of the function. Plugging these functions into (8) produces

7

$$f(X) = \kappa_1 1 + \kappa_2 X + \kappa_3 (X - \xi_1)_+ + \kappa_4 (X - \xi_2)_+, \tag{11}$$

which is a familiar-looking linear function, with special features allowing regions $[-\infty, \xi_1)$, $[\xi_1, \xi_2)$, and $[\xi_2, \infty)$ to have different slopes. In this function, there are $m = 4$ coefficients to accommodate $K = 2$ internal knots; a quadratic model would have $m = 5$ coefficients, as in (12):

$$f(X) = \kappa_1 1 + \kappa_2 X + \kappa_3 X^2 + \kappa_4 (X - \xi_1)_+^2 + \kappa_5 (X - \xi_2)_+^2, \tag{12}$$

A cubic model would have $m = 6$ coefficients, and so on. Generally, if there are K knots, and the function is of order Q, there will be $m = K + Q$ coefficients.

Figure 2. Illustration of a 1$^{\text{st}}$-degree piecewise polynomial with 2 internal knots ξ_1 and ξ_2.
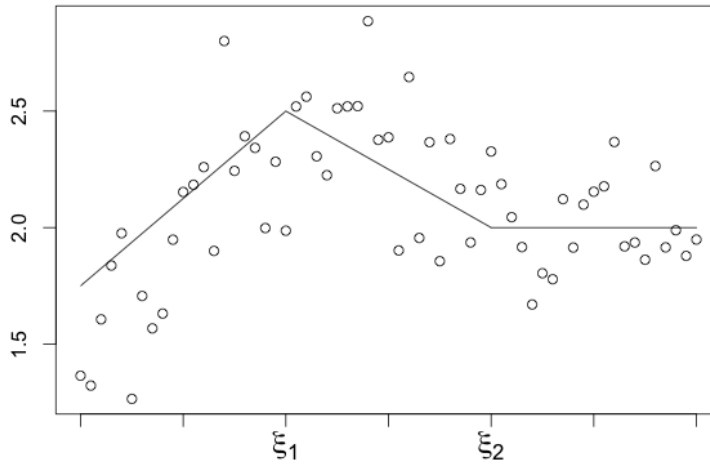
B-splines

To configure the linear space for b-splines, consider a space with Q knots $\tau_1, \ldots \tau_Q$ (with Q corresponding to the order of the b-splines function) to the left of the boundary knot ξ_0; K knots $\tau_{k+Q} = \xi_k$, $k = (1, \ldots, K)$ corresponding to the internal knots $\xi_1, \ldots, \xi_K$; and Q knots $\tau_{K+Q+1}, \ldots, \tau_{K+2Q}$ to the right of the other boundary knot ξ_{K+1}. Then in function (8), h_d is replaced by $B_{iQ}(\cdot)$ for knots $i = (1, \ldots, K+2Q)$. The recursive b-spline function is defined as

$$
\begin{cases}
B_{i,1}(x) = \begin{cases} 1 & \text{if } \tau_i \le x < \tau_{i+1} \\ 0 & \text{otherwise} \end{cases} \\[2ex]
\text{for } i = (1, \ldots, K + 2Q - 1) \\[2ex]
B_{i,q}(x) = \dfrac{x - \tau_i}{\tau_{i+q-1} - \tau_i} B_{i,q-1}(x) + \dfrac{\tau_{i+q} - x}{\tau_{i+q} - \tau_{i+1}} B_{i+1,q-1}(x) \\[2ex]
\text{for } i = (1, \ldots, K + 2Q - q)
\end{cases}
\tag{13}
$$

This function is evaluated by constructing a triangular matrix, beginning with order $q = 1$ and ending with order $q = Q$ and taking the sum of the right-most column of the matrix to construct the b-splines function (14).

$$
f(x) = \sum_{d=1}^{m} \kappa_d B_{iQ}(x)
\tag{14}
$$

Natural cubic splines

B-splines provide great flexibility in modeling non-linear shapes without departing from parametric methods, although they can be somewhat badly behaved at the tails.[17] Natural cubic spines are a modification of (10) with additional constraints of linearity beyond the boundary knots. A natural cubic spline with K knots will have K basis functions (and K coefficients). Each basis function is represented as

$$
\begin{cases}
N_1(X) = 1, \\
N_2(X) = X, \\
N_{k+2}(X) = d_k(X) - d_{K-1}(X), \quad k = (1, \ldots, K), \\[2ex]
d_k(X) = \dfrac{(X - \xi_k)_+^3 - (X - \xi_K)_+^3}{\xi_K - \xi_k}
\end{cases}
\tag{15}
$$

The boundary constraints $(X - \xi_K)_+^3$ in the numerator in $d_k(X)$ drop out when x is inside the boundaries and apply when x is outside the boundaries, causing more reasonable behavior in the tails.

Piecewise methods in mixed-effects models

9

For all piecewise methods, the piecewise function (8) can be inserted into the linear mixed-effects model as a regressor, and, with the functions h_d defined, estimation proceeds as with linear models.[17]

4.1.3 Advantages and limitations of linear mixed-effects models

Key advantages of mixed-effects models include that they can produce estimates of individual subjects' longitudinal trajectories in addition to estimates of the mean response.[8,16] Also, because the count of model parameters does not increase with the number of longitudinal time points in the mixed-effects model, the method tends to produce parsimonious models.[6] Finally, mixed-effects models can accommodate unbalanced study designs, wherein different individuals do not necessarily have the same number of longitudinal time points, although the mechanism underlying missing data must be considered in constructing the model.

In fact, missing-data mechanisms are a key component of longitudinal modeling. To make a discussion of missing-data patterns more clear, let's introduce some definitions:

- Let T_i^* be time to dropout;

- let $R_i(t)$ be an indicator of whether a response for individual i was missing ($R_i(t) = 0$) or observed ($R_i(t) = 1$) at time t;

- let y_i^o be the observed longitudinal observations for individual i;

- let y_i^m be the missing observations for individual i that would have taken place had the dropout event T_i^* not occurred.

There are 3 types of missing-data mechanisms relevant to our discussion: missing completely at random (MCAR); missing at random (MAR); and missing not at random (MNAR).[6,7]

MCAR: If the probability that a response is missing is unrelated to both y_i^o and y_i^m (formally, $P(r|y_i^o, y_i^m) = P(r)$), the mechanism is said to be missing completely at random (MCAR). Under MCAR, limiting the analysis dataset to cases of complete ascertainment will produce unbiased estimates. An example of an MCAR mechanism would be a missing observation that resulted from damage to a source document.

MAR: If the probability that a response is missing depends on y_i^o but not y_i^m (formally, $P(r|y_i^o, y_i^m) = P(r|y_i^o)$), the mechanism is said to be missing at random (MAR). Under MAR, estimates conditioned on observed values y_i^o are unbiased, although they

are sensitive to misspecification of the covariance matrix. Traditional linear mixed-effects models can provide valid parameter estimates under the MAR mechanism, but not the MNAR mechanism. An example of an MAR mechanism would be a missing observation in an interventional trial that resulted from removing a participant because the participant's condition was not sufficiently controlled under the study treatment.

MNAR: If the probability that a response is missing depends on both y_i^o and y_i^m (formally, $P(r|y_i^o, y_i^m) = P(r|y_i^o, y_i^m)$), the mechanism is said to be missing not-at-random (MNAR). An example of an MNAR mechanism would be a missing observation of income that is missing because people with low income are less likely to disclose their income.

Under MNAR, it is necessary to jointly model the missing-data mechanism with the longitudinal trajectory. One method of accomplishing this is the class of shared-parameter models to which joint models belong.[8,18] In the shared-parameter framework, both the longitudinal process and the event process are assumed to depend on shared latent variables. That is, longitudinal measurements Y_i and missingness indicators R_i are assumed independent conditional on random effects b_i. Under this assumption, if the model parameters are maximized conditional on b_i, it is possible to measure the association of this conditional part.

4.1.4 Model diagnostics for mixed-effects models

Conditional residuals (16) plotted against model-fitted values can be used to check the assumption of homoscedasticity in a mixed-effects model. If the spread of points is larger on one side of the plot than on the opposite side, there may be a violation of the homoscedasticity assumption. We hope to observe no clear pattern with relatively constant variance.

$$r_i^{yc}(t) = y_i(t) - \hat{\boldsymbol{\beta}} \mathbf{x}_i^\top(t) - \hat{\mathbf{b}}_i \mathbf{z}_i^\top(t), \tag{16}$$

Conditional residuals can be standardized by scaling them by the standard deviation of the errors. Plotted against theoretical quantiles of a normal distribution, the standardized conditional residuals will form a diagonal line if they are normally distributed.

$$r_i^{ysc}(t) = \frac{y_i(t) - \hat{\boldsymbol{\beta}} \mathbf{x}_i^\top(t) - \hat{\mathbf{b}}_i \mathbf{z}_i^\top(t)}{\hat{\sigma}} \tag{17}$$

Marginal residuals (18) plotted against fitted values can be used to check for misspecification of the design matrix X. A departure from a mean of 0 at any portion of the plot indicates some systematic departure from the fitted values and is an indication of poor fit of the mean response profile. Also, since the variance of the random effects is included in the marginal residuals, a larger spread of points at one end of the plot than on the opposite side may indicate a misspecification of the variance-covariance matrix Σ.

$$r_i^{ym}(t) = y_i(t) - \hat{\beta} \mathbf{x}_i^\top(t) \tag{18}$$

4.2   Time-to-event models

Time-to-event models provide a framework for modeling time to an event of interest given a set of predictors. The defining feature of time-to-event data is the presence of censoring—the inevitable truncation of follow-up time that results when a subset of study participants do not experience the event during their participation in the study. The following notation will be useful in discussing such models:

- T_i^* is the true time to event for individual i.

- C_i is the time of censoring.

- $T_i = \min(T_i^*, C_i)$

- $\delta_i = I(T_i^* \leq C_i)$

where $I()$ is the indicator function (i.e., $\delta_i = 1$ for individuals who experienced the event during the study period, and $\delta_i = 0$ for censored individuals)

The joint likelihood function (19) accounts for censoring by contributing the density function for each non-censored individual and the survivor function (defined as $P(T^* > t)$) for each censored individual.[8]

$$L(\theta|\mathbf{t}) = \prod_{i=1}^{n} \left[p(t_i|\theta)^{\delta_i} S(t_i|\theta)^{(1-\delta_i)} \right] \tag{19}$$

where $S(t|\theta)$ is the survivor function and $p(t|\theta) = h(t|\theta)S(t|\theta)$ is the density of the random variable T.

12

The hazard function $h(t|\theta)$ can be thought of as the instantaneous rate of the event at time t, based on the count of events up to this time, and can be expressed in probability notation as

$$h(t|\mathbf{x}) = \lim_{dt \to 0} \frac{P(t \leq T_i^* < t + dt | T_i^* \geq t)}{dt} \tag{20}$$

or modeled as

$$h(t|\mathbf{x}) = h_0(t)g(\mathbf{x}'\boldsymbol{\gamma}) \tag{21}$$

where $h_0(t)$ is the baseline hazard function, i.e., the hazard when all $(\mathbf{x}_i'\boldsymbol{\gamma}) = 0$, and $g(\cdot)$ is an appropriate link function. In the case of parametric time-to-event models, $h_0(t)$ can be modeled using, for example, an exponential or Weibull distribution. Of course, the natural world rarely presents us with a baseline hazard that aligns perfectly to parametric curves. In cases where it is especially advantageous to accurately model the baseline hazard, additional flexibility can be gained by applying one of the piecewise approaches discussed in section 4.1.2.

If the covariates $\mathbf{x}$ are multiplicatively related to the hazard, then for any points $\mathbf{x_1}$ and $\mathbf{x_2}$, $h_0(t)$ cancels from the hazard ratio (22), which is constant with respect to time, and the model is called a proportional-hazards model.

$$\frac{h(t|\mathbf{x_1})}{h(t|\mathbf{x_2})} = \frac{h_0(t)g(\mathbf{x_1}'\boldsymbol{\gamma})}{h_0(t)g(\mathbf{x_2}'\boldsymbol{\gamma})} = \frac{g(\mathbf{x_1}'\boldsymbol{\gamma})}{g(\mathbf{x_2}'\boldsymbol{\gamma})} \tag{22}$$

When the link function $g(\cdot)$ is the exponential, the model is called a Cox proportional-hazards model. In either case, $h_0(t)$ can be left unspecified, and estimation is carried out by maximizing the partial likelihood function.[19]

4.2.1 Extended Cox model with a time-varying covariate

Since time-to-event datasets deal with the passing of time, they often include one or more covariates whose values at baseline may differ from those at time t. To account for such covariates, the extended Cox model (23) includes a time-varying covariate or vector of covariates $\mathbf{y}_i$ and a corresponding vector of regression coefficients $\boldsymbol{\gamma_2}$:

$$h(t|\mathbf{x}) = h_0(t)\exp(\boldsymbol{\gamma_1}\mathbf{x}_i + \boldsymbol{\gamma_2}\mathbf{y}_i(t)) \tag{23}$$

where $\mathbf{y}_i(t)$ is the value of $\mathbf{y}_i$ at time t.

Time-varying covariates can be endogenous or exogenous. Endogenous covariates are those whose values depend on characteristics or behavior of individual participants, such as exposure to treatment, smoking status, or blood biomarkers; exogenous covariates include external forces acting on the entire sample or subgroups of the sample simultaneously, such as pollution levels or economic conditions.[19]

Rizopoulos[8] formalizes the idea of exogenous covariates by specifying that an exogenous time-varying covariate $y_i(t)$ is associated with the time-varying hazard, but the covariate's future trajectory for individual i is not affected by occurrence of the event for individual i at time s:

$$P[Y_i(t)|Y_i(s), T_i^* \geq s] = P[Y_i(t)|Y_i(s), T_i^* = s], \quad s \leq t \tag{24}$$

Conversely, when the event is death, an endogenous covariate necessarily cannot exist after the event, so the survival function is equal to 1 wherever the covariate $Y_i(t)$ exists:

$$S_i[t|Y_i(t)] = P[T_i^* > t|Y_i(t)] = 1 \tag{25}$$

When death has occurred at time s, $Y_i(t)$ does not exist on $t \geq s$, and the usual relationships

$$\begin{cases} S_i(t|Y_i(t)) = \exp\left[-\int_0^t h_i(s|Y_i(s))\,ds\right], \\ p(t|Y_i(t)) = h_i(s|Y_i(s))S_i(t|Y_i(t)) \end{cases} \tag{26}$$

are fundamentally altered, in that the realities $S_i(\cdot) = 0$, $h_i(\cdot) = 1$, and $p(\cdot) = 0$ are not reflected in (26). Under this circumstance, the log-likelihood (27), which depends on $p(\cdot)$ and $S(\cdot)$, is approximately meaningless.

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \delta_i \log p(T_i|\boldsymbol{\theta}) + (1 - \delta_i)\log S_i(T_i|\boldsymbol{\theta}) \tag{27}$$

Making the situation worse, since the extended Cox model places the time-varying covariate on the right side of the regression equation, it violates the weak-exogeneity assumption

14

of linear regression—that is, it is assumed to be free of measurement error.[20] This problem results in regression coefficients that suffer from overstated precision. A method that accommodates the inclusion of measurement error in the time-varying component would be likely to produce more realistic estimates of standard error.

The extended Cox model treats the time-varying covariate as a step function, assuming that each value is measured without error and carried forward to the next measurement time. In the case of the time-varying covariate of CD4 counts in our didactic dataset, for example, the underlying, unmeasured longitudinal process is a smooth longitudinal curve, changing instantaneously—not just at study visits when CD4 counts were measured (Figure 3).

Figure 3. Graphical representation of a step-function model of a continuous underlying longitudinal covariate. The solid line indicates the step-function estimated by the extended Cox model, and the smooth dotted line indicates the true, continuous underlying longitudinal trajectory. This figure is adapted from Rizopoulos.[8]
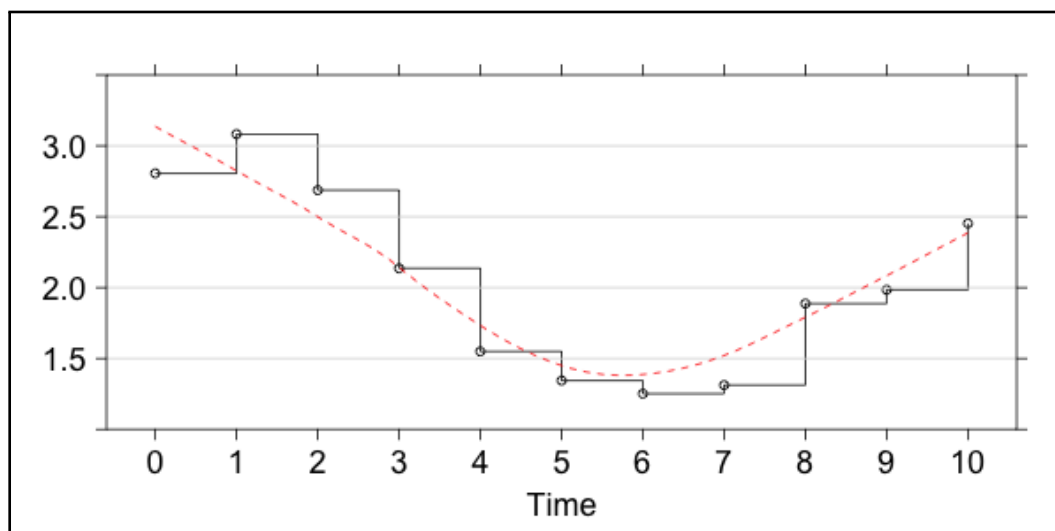
4.2.2   Model diagnostics for proportional-hazards models

Cox-Snell residuals (28) can be useful to diagnose the overall fit of an event model.

$$r_i^{tcs}(t) = \int_0^{T_i} \hat{h}_0(s) \exp(\hat{\gamma}^\top \mathbf{x}_i) \, ds$$
$$= \widehat{H}_0(T_i) \exp(\hat{\gamma}^\top \mathbf{x}_i) \sim \exp(\lambda = 1)$$
(28)

15

Theoretically, if the model fits, its estimated cumulative hazard has an exponential distribution with a hazard rate $\lambda = 1$.[8,19] Then the model fit can be checked by comparing the survival function of an $\exp(\lambda = 1)$ distribution to a Kaplan-Meier estimate of the survival function of the Cox-Snell residuals. If the model fits, the two curves will be similar.

Another option is to plot the cumulative hazard of the Cox-Snell residuals against the residuals themselves. In a theoretical $\exp(\lambda = 1)$ distribution, $H(t) = t$, so the estimated cumulative hazard plotted against the Cox-Snell residuals will form a $45°$ line through the origin if the model fits.[19]

Martingale residuals are a slight modification of Cox-Snell residuals. They are useful for examining the appropriateness of the functional form of individual covariates and identifying thresholds at which covariates might best be categorized. In context of a model where the event of interest is death, Martingale residuals (29) can be interpreted as the difference between the observed number of deaths for individual i at time t and the number of events at t predicted by the mode. departure of observed deaths from the predicted probability of death given the value of x_i at time t. To assess the functional form of an important covariate x_i, we can plot the Martingale residuals against values of x_i and look for a systematic departure from mean 0.

$$r_i^m(t) = \begin{cases} 1 - \int_0^t R_i(s)\hat{h}_0(s)\exp(\hat{\gamma}^\top \mathbf{x}_i)\,ds \\ R_i(s) = \begin{cases} 1 & \text{if individual } i \text{ is at risk at time } s, \\ 0 & \text{otherwise} \end{cases} \\ = R_i(s) - r_i^{tcs} \end{cases} \tag{29}$$

4.3 Joint models

Joint models provide a framework for simultaneously modeling a longitudinal outcome with a time-to-event outcome. A longitudinal sub-model of form (1) provides a continuous estimate of the actual, unmeasured longitudinal trajectory, complete with its own residual term. This model can be incorporated as a sub-model component of a Cox proportional-hazards model (30):

$$\begin{cases} h_i(t|m_i(t), \mathbf{w}_i) = h_0(t)\exp\left(\boldsymbol{\gamma}^{\top}\mathbf{w}_i + \alpha m_i(t)\right), \\ \qquad\quad y_i(t) = m_i(t) + \varepsilon_i(t), \\ \qquad\quad m_i(t) = \beta\mathbf{x}_i^{\top} + \mathbf{b}_i\mathbf{z}_i^{\top}(t), \\ \qquad\quad \mathbf{b}_i, \sim \mathcal{N}(0, \Sigma), \quad \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2) \end{cases} \tag{30}$$

where $h_i(t|m_i(t), \mathbf{w}_i)$ is the hazard function, and $m_i(t)$ represents the modeled estimate of the complete, continuous longitudinal trajectory. The observed longitudinal values $y_i(t)$ are the sums of the modeled trajectory $m_i(t)$ and the measurement error $\varepsilon_i(t)$. The chief proposed advantage of the joint model is that the hazard at time t is adjusted for the value of $m_i(t)$, which is postulated to be a more realistic estimate of the longitudinal path than the Cox model's step function. The association parameter α denotes the association between the longitudinal and event models and can be interpreted in exactly the same way as a regression coefficient from a Cox model—that is, a unit change in the longitudinal outcome predicts a change of $\exp(\alpha)$ in the hazard ratio.[8]

When there is no real association between the longitudinal and event processes ($\alpha = 0$), the joint model is equivalent to the set of separate models. Conversely, when an association does exist, the consequence of ignoring that association is generally an underestimation of the magnitude of the effect of all longitudinal and event model parameters. Henderson et al.[4] conducted a simulation study, using a joint model that assigns association parameters to the longitudinal intercept, slope, and current value at time t. Their study demonstrated that separate models that ignored the association between the longitudinal and event processes consistently underestimated true parameter values, including variance of the random effects, while estimates from joint models were consistently much closer to true values. Standard errors from separate and joint models were similar. The authors applied their model to a didactic dataset from a schizophrenia trial, and comparisons between separate and joint models displayed a similar pattern, indicating that separate models may have underestimated parameters as a consequence of ignoring association.

Of course, as described in section 4.2.1, the Cox model can be extended to accommodate a time-varying covariate on an event process, but in the case of endogenous covariates, this strategy has some serious limitations.

Biomarkers (such as CD4 lymphocyte counts) are especially good candidates as longitudinal covariates for joint models, because, as one can easily imagine, their values are continuous functions of time. The step model shown in Figure 4 seems especially inappropriate for a lymphocyte count, since one can presume that if we were able to monitor this outcome constantly during the study, the count would be changing constantly—not

17

Figure 4. Graphical representaiton of a joint model (concept borrowed from Rizopoulos[8]).
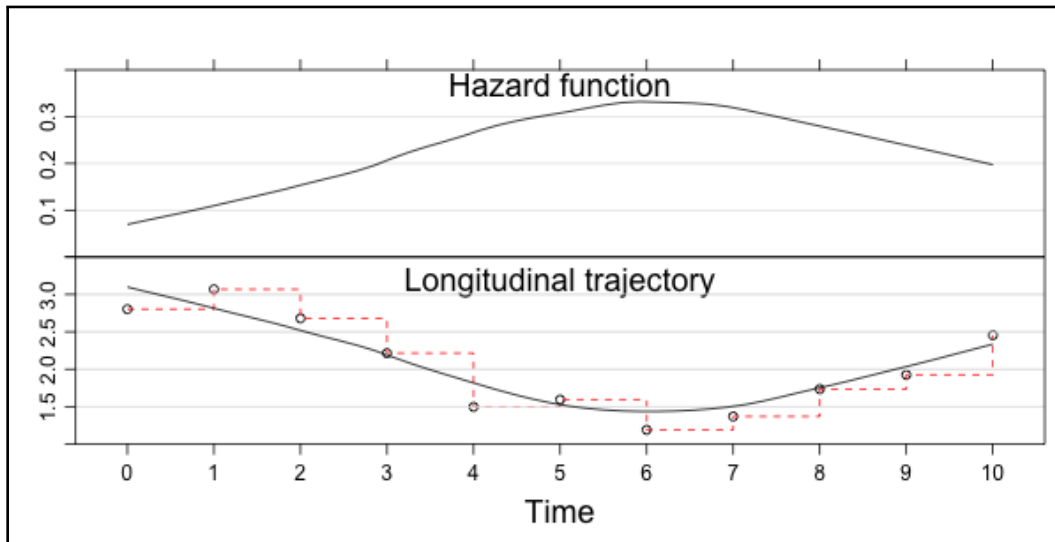
jumping instantly from one value to the next at the time of measurement. A counter-example to lymphocyte counts would be a prescribed drug dose, which is a value we could expect to change abruptly at scheduled clinic visits, coinciding with a dose measurement. In such cases, the step-function model of the extended Cox framework would more closely represent reality. If, however, we wanted to measure the amount of the drug in the partici-pant's blood, rather than the prescribed dose, we would prefer to model it as a continuous function, as in a joint model.

As implemented in the R package JM, joint models can incorporate the longitudinal compo-nent in a variety of ways, including an instantaneous value of the longitudinal covariate at hazard time t; a random intercept and slope; the cumulative area under the curve up to time t; or a lag-function—that is, the value of the longitudinal covariate at some specified time before t. These options provide a great deal of flexibility in accounting for longitudinal changes in a variety of covariates that may differ in how they impact survival time. The JM package facilitates linear mixed-effects longitudinal sub-models, but not generalized lin-ear mixed sub-models, so it is not yet convenient to model discrete (for example, Poisson or binary) longitudinal outcomes, although a new R package based on work by Viviani and Rizopoulos is forthcoming (S Viviani [sara.viviani@uniroma1.it], e-mail, September 3, 2013).

4.3.1 Joint models and mechanisms of missing data

As discussed in section 4.1.3, traditional linear mixed-effects models can provide valid parameter estimates under the MAR mechanism, but not the MNAR mechanism.[7] If missingness is related to both the observed and missing longitudinal trajectories, it is necessary to jointly model the missing-data mechanism with the longitudinal trajectory.[6,7] In a joint model the longitudinal and time-to-event sub-models share the same random effects and so account for a missing-not-at-random (MNAR) mechanism. In fact, the association parameter α measures the effect of drop-out on the longitudinal outcome. If the true underlying mechanism is MCAR, the association parameter will be $\alpha = 0$, and the joint model will be equivalent to the set of separate models.[8]

4.3.2 Estimation

Wulfsohn and Tsiatis proposed the likelihood function

$$\ell(\theta|T_i,\delta_i,y_i) = \sum_{i=1}^{n} \log \int_{-\infty}^{\infty} \left(\prod_{j=1}^{m_i} p(y_{ij}|\mathbf{b}_i,\theta_y) \right) p(\mathbf{b}_i|\theta_b) p(T_i,\delta_i|\mathbf{b}_i,\theta_t,\beta) \, d\mathbf{b}_i \qquad (31)$$

where $\theta = \{\theta_y, \theta_b, \theta_t\}$, $\theta_y = \{\beta, \sigma_e^2\}$, $\theta_b = \{\sigma_b^2\}$, and $\theta_t = \{\gamma, h_0(t), \alpha\}$;

then $p(y_{ij}|\mathbf{b}_i,\theta_y)$ is the joint density for the m_i longitudinal observations for individual i conditioned on the random effects $\mathbf{b}_i$ and the longitudinal outcome parameters $\theta_y = \{\beta, \sigma_e^2\}$;

$p(\mathbf{b}_i|\theta_b)$ is the distribution of the multivariate normally distributed random effects matrix, conditioned on parameters of the random-effects covariance matrix $\theta_b = \{\Sigma\}$;

and $p(T_i,\delta_i|\mathbf{b}_i,\theta_t,\beta)$ is the distribution of the survival process, conditioned on the random effects $\mathbf{b}_i$, the regression coefficients β from the longitudinal model, and the event outcome parameters $\theta_t = \{\gamma, h_0(t), \alpha\}$. Note that α is the association parameter.

The likelihood function (31) is maximized using a 2-stage estimation-maximization (EM) procedure. The E-step computes an expected value of the log-likelihood of the compete model, conditioned on the random effects and current estimates of the parameters. The M-step maximizes the parameter estimates using the computed expected log-likelihood from the E-step. The procedure is repeated until the parameter estimates converge.[4,11]

A key advantage of the Cox proportional-hazards model is the freedom to leave the baseline hazard function unspecified, and much of the theoretical work in joint modeling preserves the unspecified baseline hazard function and estimates parameters by partial maximum likelihood.[4,11,21] Note, however, that the parameter set θ_t includes the baseline hazard function $h_0(t)$, so the log-likelihood can be computed with or without specifying the baseline hazard. When the baseline hazard is left unspecified, the semiparametric likelihood approach necessary to cope with h_0 results in a step function in place of the cumulative hazard $\int h(s)\,ds$, adding a new parameter to the joint model for each step. A profile-likelihood approach is one way to deal with the high-dimensionality of such models (and is the approach used by the R function `jointModel()`, from package JM, when the baseline hazard is left unspecified), but since this leads to an estimator in the M-step of the EM algorithm that depends on h_0, this approach will often underestimate the standard errors for parameters $(\beta, \sigma, \gamma, \alpha)$. As a result, it is advisable to specify a parametric model for h_0.[8]

In the current project, the R function `jointModel()` failed prior to convergence, due to high dimensionality, when the baseline hazard was left unspecified.

4.3.3 Inference

For univariate tests of the joint model parameter estimates $\hat{\theta} = (\hat{\beta}, \hat{\gamma}, \hat{\alpha})$, the differences between maximum-likelihood estimates and their null values are assumed to be normally distributed, so their squares follow a chi-squared distribution, and significance can be tested using the Wald procedure (32).

$$W = \frac{\hat{\theta} - \theta_0}{\widehat{s.e.}(\hat{\theta})} \tag{32}$$

Confidence intervals are also Wald-based and can be generalized to produce confidence intervals for the predicted longitudinal trajectory.

Comparison of nested joint models is possible by standard likelihood-ratio test, and comparison of non-nested models is possible using Akaike's information criterion (AIC) or Bayesian information criterion (BIC). Both AIC and BIC penalize for the total number of both longitudinal and event parameters.

4.3.4 Joint model diagnostics

Graphical diagnostic techniques for separate longitudinal and time-to-event models (see sections 4.1.4 and 4.2.2, respectively) are also available to check the assumptions of each of the 2 sub-models in a joint model, with a key limitation: the mixed-effects model framework assumes a missing-data mechanism of MAR, and this plot of residuals reflects that assumption. In context of a joint model, however, the mixed-effects sub-model assumes a missing-data mechanism of MNAR. We can adjust the plot of residuals to compensate for non-ignorable missingness by imputing the missing longitudinal values. The imputation technique calculates posterior distributions of the joint model parameters and random effects and the draws a sample of a pre-specified size for each individual from the posterior distribution of the missing values y_i^m, given the observed data for each visit time point. Residuals are then calculated from both the observed y_i^o and the imputed y_i^m observations. If the non-imputed plot of residuals shows a systematic deviation from 0 that disappears in a plot with residuals generated from multiply imputed missing values, this is an indication that the missing-data mechanism is MNAR.

Note that, in context of joint modeling, this imputation procedure is intended to be used for generating a plot of residuals—not for inference in the joint model.

If, as in many observational studies, visit times are random (that is, not pre-specified by a protocol, as is often the case in prospective studies), the timing of the visits that would have occurred after the observed event or censoring time T_i is unavailable, so it is necessary to model the timing of these "missing visits." A Weibull model of visit times with a Gamma frailty term (33) is a simple, flexible strategy for modeling event times for each individual.[8]

$$\begin{cases} \lambda(u_{ik}|x_{vi}, w_i) = \lambda_0(u_{ik})w_i \exp(x_{vi}^\top \gamma_v) \\ w_i \sim Gamma(\sigma_w, \sigma_w) \end{cases} \tag{33}$$

In (33), u_i is the kth sequential event time for individual i, w_i is the frailty term, γ_v are the coefficients corresponding to the covariates x_{vi}, and $\frac{1}{\sigma_w}$ is the variance of the frailty terms.

5 Longitudinal measures of CD4 lymphocyte counts and survival time in AIDS patients

To illustrate joint modeling methods, we will use a didactic dataset of longitudinal CD4 lymphocyte counts and survival time from a multi-center, open-label study in which 467 patients with very low CD4 lymphocyte counts ($\leq$ 300 per cubic millimeter) or a diagnosis of acquired immunodeficiency syndrome (AIDS) who had previously failed treatment with zidovudine were randomly assigned to treatment with didanosine (ddI, n=230) or zalcitabine (ddC, n=237). Participants were followed for survival time over a median follow-up period of 16 months, and CD4 counts were measured up to 5 times during study participation.[22] CD4 count is an important biomarker of immune system health, so a decline in this measure is associated with higher risk of infection and death in AIDS patients.[23,24] As such, a secondary goal of the study was to examine the association between the longitudinal process of CD4 counts and survival time. This dataset has been used repeatedly in the literature on joint modeling,[8–12] so it provides a good didactic opportunity.

CD4 counts were right-skewed, reflecting the very low CD4 counts in the study population, and required a square-root transformation for normality. Median square-root CD4 counts declined from 6.1 at baseline to 4.5 at 18 months in an approximately linear trend (Figure 5).

Figure 5. Median $\sqrt{CD4}$ over time among patients diagnosed with AIDS.

22
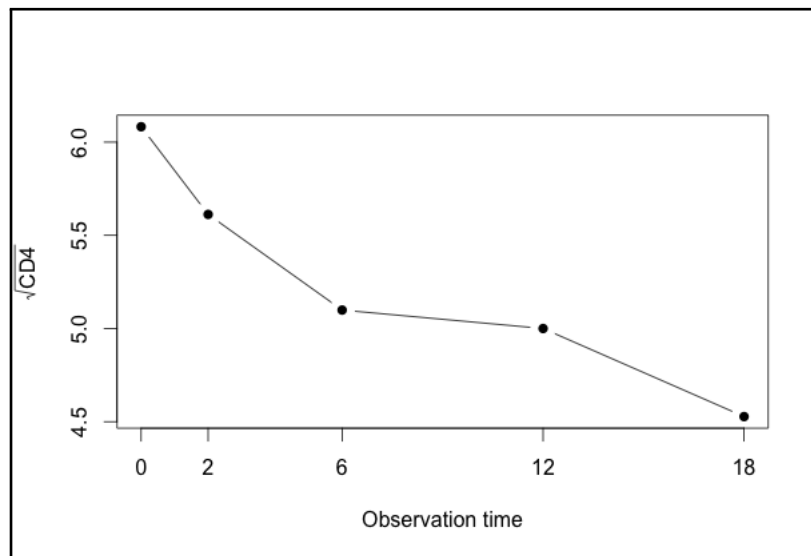
The linear mixed-effects model (34) estimated a decrease of 0.16 units in square-root CD4 count over each passing month (95% CI: 0.12 to 0.20 units); the ddI group's average decline each month did not significantly differ from that of the ddC group. At baseline, the two treatment groups had substantially similar square-root CD4 levels (see Table 1). There was far more variation between individuals in square-root CD4 counts at baseline (SD=4.59) than in change over time (SD=0.17).

$$\begin{cases} y_i(\text{time}) = \beta_0 + \beta_1 \text{ddI}_i + \beta_2 \text{time} + \beta_3 \text{ddI} \times \text{time} + b_{i0} + b_{i1} \text{time} + \varepsilon_i(\text{time}), \\ (b_{i0}, b_{i1}) \sim \mathcal{N}(0, \Sigma), \\ \varepsilon_i(\text{time}) \sim \mathcal{N}(0, \sigma^2) \end{cases} \quad (34)$$

A plot of Kaplan-Meier survival curves (Figure 6) shows that patients in the ddI and ddC groups had very similar survival over the first six months, and then patients in the ddC group had slightly better survival through 18 months of follow-up.

Figure 6. Kaplan-Meier survival curves for patients in the ddI and ddC treatment groups.
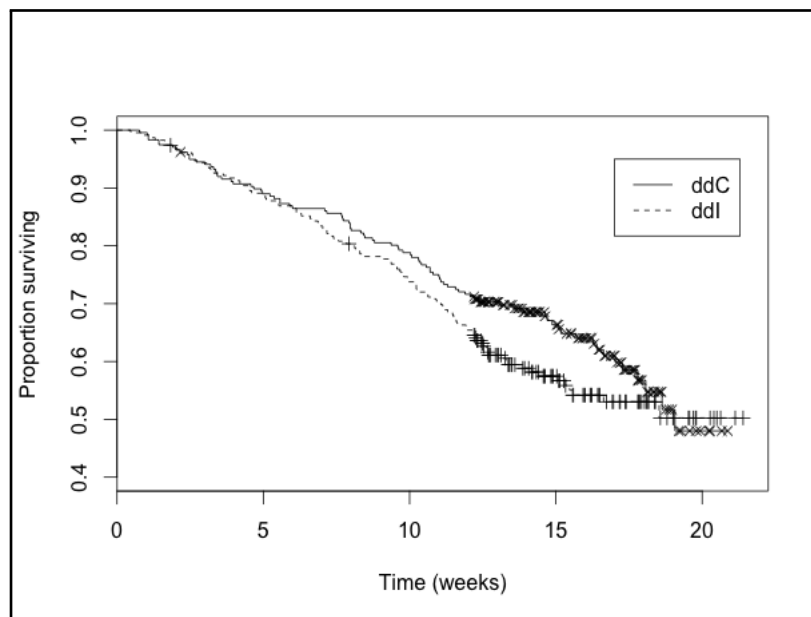
The Cox model (35), with treatment group and a single (baseline) measure of square-root CD4 count, estimates that treatment with ddI is associated with a 30% increase in risk of death compared to treatment with ddC (95% CI: 2% lower to 74% higher) and that each

unit decrease in baseline square-root CD4 count is associated with a 20% higher risk of death (95% CI: 15% to 25%).

$$h_i(\text{time}|\mathbf{x}_i) = h_0(\text{time}) \exp(\gamma_1 \text{ddI}_i + \gamma_2 \text{CD4}_i) \tag{35}$$

Of course, we want to consider the effect of change in square-root CD4 count, rather than a single baseline measure, so we can extend model (35) to account for time-varying square-root CD4 count:

$$h_i(\text{time}|\mathbf{x}_i) = h_0(\text{time}) \exp\left(\gamma_1 \text{ddI}_i + \gamma_2 \text{CD4}_i(\text{time})\right) \tag{36}$$

Model (36) incorporates time-varying square-root CD4 count, which pushes the estimate for treatment with ddI into the realm of statistical significance at the 0.05 level. In this model, treatment with ddI is associated with a 36% increase in risk of death (95% CI: 2% to 82%) compared to treatment with ddC. Each unit decrease in square-root CD4 is associated with a 21% increase in risk of death (95% CI: 16% to 27%).

We can combine the longitudinal model (34) as a sub-model in the event model to produce joint model (37).

$$\begin{cases} h_i(\text{time}|\text{ddI}_i) = h_0(\text{time}) \exp\left(\gamma_1 \text{ddI}_i + \alpha m_i(\text{time})\right), \\ \quad h_0(\text{time}) = 1, \\ \quad y_i(\text{time}) = m_i(\text{time}) + \varepsilon_i(\text{time}), \\ \quad m_i(\text{time}) = \beta_0 + \beta_1 \text{ddI}_i + \beta_2 \text{time} + \beta_3 \text{ddI}_i \times \text{time} + b_{i0} + b_{i1}\text{time} \end{cases} \tag{37}$$

Model (37) estimates that treatment with ddI is associated with a 36% increase in risk of death (95% CI: 2% to 82%) and that each unit decrease in square-root CD4 count is associated with a 21% higher risk of death (95% CI: 16% to 27%). Compared to the linear mixed-effects model (34), the joint model produced a substantially higher estimate of overall baseline square-root CD4 counts, and it reversed the direction of the effect of treatment group on the longitudinal trajectory of CD4 counts. The joint model generally produced slightly narrower confidence intervals, but we should expect the standard errors for these fixed effects to be somewhat underestimated, because the baseline hazard function is unspecified.[8]

In context of joint modeling, it is almost always better to specify the baseline hazard function, as in model (38), which includes a 4th-order b-splines approximated baseline hazard with 5 equally spaced internal knots.

$$
\begin{cases}
h_i(\text{time}|\text{ddI}_i) = h_0(\text{time}) \exp\left(\gamma_1 \text{ddI}_i + \alpha m_i(\text{time})\right), \\[2ex]
\log h_0(\text{time}) = \sum_{d=1}^{9} \kappa_d B_{d4}(\text{time}), \\[2ex]
y_i(\text{time}) = m_i(\text{time}) + \varepsilon_i(\text{time}), \\[2ex]
m_i(\text{time}) = \beta_0 + \beta_1 \text{ddI}_i + \beta_2 \text{time} + \beta_3 \text{ddI}_i \text{time} + b_{i0} + b_{i1} \text{time}
\end{cases}
\tag{38}
$$

In model (38), κ_d are the spline coefficients, and $B_{d4}(time)$ are the summed b- basis functions at time t of degree 4.

By Bayesian information criteria (BIC), model (38) improves the fit of the joint model compared to the joint model with an unspecified baseline hazard (BIC=8760, compared to BIC=9592, respectively). Model (38) estimates that treatment with ddI is associated with a 40% increase in risk of death compared to treatment with ddC (95% CI: 3% to 9%) and that each unit decrease in square-root CD4 count is associated with a 33% higher risk of death (95% CI: 24% to 42%). This model produces longitudinal estimates very similar to those of the linear mixed-effects model (34). Since treatment group is a significant predictor of survival time but not of longitudinal changes in square-root CD4 count, it seems that its effect on survival is not mediated through changes in CD4 count. A likelihood ratio test comparing a model without the effect of treatment group to a model with the effect of treatment group is significant (p=0.04).

As expected, confidence intervals for all estimates are wider than in joint model (37) and than in the extended Cox model (36), and the magnitudes of the event estimates are larger.

Table 1 summarizes parameter estimates from all separate and joint models of square-root CD4 count and survival time.

6 Longitudinal measures of walking gait speed and survival time in older men

The MrOS study is a prospective longitudinal cohort study (n=5995) of ambulatory, community-dwelling men ages 65 and older that examines sequelae of fractures and the relationships between a variety of biomedical and lifestyle factors and risk of fracture. Enrollment took place from March 2000 to April 2002, and follow-up continues at the

Table 1. Parameter estimates from separate and joint models of sqare-root CD4 lymphocyte counts and survival time among patients diagnosed with AIDS.

	Cox[*]	LME/extended Cox[†]	Joint model[‡] BIC = 9592	Joint model[§] BIC = 8760
Longitudinal estimates				
Intercept		6.95 (6.34, 7.56)	7.51 (7.07, 7.95)	6.96 (6.35, 7.57)
ddI		0.48 (-0.39, 1.35)	-0.22 (-0.76, 0.33)	0.49 (-0.38, 1.36)
Time		-0.16 (-0.20, -0.12)	-0.17 (-0.21, -0.13)	-0.19 (-0.23, -0.14)
ddI X time		0.02 (-0.04, 0.08)	0.00 (-0.06, 0.06)	0.01 (-0.05, 0.07)
Event estimates (exponentiated to obtain HRs)				
ddI	1.30 (0.98, 1.74)	1.36 (1.02, 1.82)	1.36 (1.05, 1.77)	1.4 (1.03, 1.9)
$\sqrt{CD4}$ (or α)	1.20 (1.15, 1.25)	1.21 (1.16, 1.27)	1.30 (1.24, 1.37)	1.33 (1.24, 1.42)

[*] Cox proportional-hazards model with baseline CD4 count only.
[†] Linear mixed-effects model (top portion) and Cox proportional-hazards model with time-varying covariate.
[‡] Joint model with unspecified baseline hazard function (`method = "Cox-PH-GH"`).
[§] Joint model with spline-approximated model of baseline hazard function (`method = "spline-PH-aGH"`).

time of this writing. Baseline data included self-reported medical histories, medications, physical activity, diet, and substance use. Participants also submitted to anthropometric, neuromuscular, vision, strength, and cognitive tests, DEXA scans, calcaneal ultrasounds, and vertebral radiographs. As a measure of physical endurance, participants were asked to walk a 6-meter course at their usual walking pace.[5] The current project utilizes data from the February 13, 2013 release of data to the community of MrOS investigators.

A subset of 1210 dentate participants from the Portland, Oregon and Birmingham, Alabama sites in the MrOS study participated in the MrOS Dental ancillary study;[13] a subset of 3135 participants from the greater MrOS cohort participated in the MrOS Sleep ancillary study.[14] Both ancillary studies administered repeated walking tests. Of the n=879 patients who participated in both ancillary studies, 814 (93%) provided at least 4 observations of walking speed. The current project will utilize data from MrOS participants who participated in both ancillary studies and had at least 1 post-baseline follow-up walking test (n=877). An additional 16 participants were excluded for missing baseline covariates or as outliers during longitudinal modeling, resulting in a final analysis dataset of n=861.

Covariates under consideration included those summarized in Table 2. To protect patient privacy, ages for patients older than 90 years were coded as 90. This change affected 4 patients in the sample. Because of the eligibility floor at 65 years old, age is substantially right-skewed, so median is used as a measure of center. Age at the time of each observation was used as the time variable.

Due to underrepresentation of Native Hawaiians, Pacific Islanders, Native Americans, and Alaskan Natives, race and ethnicity were collapsed to a single variable and were recatego-

rized as non-Hispanic white, non-Hispanic African-American, and other. Alcohol use is by self-report of having had 12 or more drinks in the year prior to baseline.

Table 2. Summary of baseline covariate candidates for modeling survival time and longitudinal changes in gait speed among older men from the MrOS cohort.

	N			Minimum	Maximum
Age, years	861	Median = 72.4	IQR = 68 to 76	65	90
BMI	861	Mean = 27.3	SD = 3.5	17.9	39.5

	n (%)		
Study site (N = 861)			
Birmingham	429 (49.8)		
Portland	432 (50.2)		
Race/ethnicity (N = 861)			
Non-Hispanic White	775 (90.0)		
African-American	54 (6.3)		
Asian	18 (2.1)		
Hispanic	5 (0.6)		
Other	9 (1.0)		
Smoking status (N = 861)			
Never smoked	322 (37.4)		
Past smoker	513 (59.6)		
Current smoker	26 (3.0)		
Self-reports 12+ alcoholic drinks in past year (N = 861)			
No	330 (38.3)		
Yes	530 (61.6)		
Don't know	1 (0.1)		
Self-reported overall health (N = 861)			
Good or excellent	773 (89.8)		
Fair, poor, or very poor	88 (10.2)		
Self-reported history of any type of cancer (N = 861)			
No	611 (71.0)		
Yes	250 (29.0)		
Self-reported history of non-skin cancer (N = 861)			
No	743 (86.3)		
Yes	118 (13.7)		
Self-reported history of congestive heart failure (N = 861)			
No	839 (97.4)		
Yes	22 (2.6)		
Self-reported history of diabetes mellitus (N = 861)			
No	781 (90.7)		
Yes	80 (9.3)		

6.1 Longitudinal modeling

Exploratory analyses of longitudinal observations of gait speed revealed a normally distributed outcome (refer to the appendix) and a slightly quadratic shape. The best-fitting model was a natural cubic-splines model with 1 internal knot, although a quadratic model fit almost as well ($BIC = 15824.9$ for the quadratic model, versus $BIC = 15823.5$ for natural-splines model). Participants with different monotonic missing-data patterns had significantly different baseline gait speeds (by ANOVA test, $F = 10.9$, $p < 0.001$), indicating that the missing-data mechanism is informative. In models with all candidate covariates, tests for random intercept and slope were both significant ($p < 0.0001$). Fixed effects were selected using a backwards approach, proceeding by removing the covariate with the largest p-value and continuing to remove covariates until Bayesian information criterion (BIC) was no longer improved. The selection procedure resulted in the selection of age, body mass index (BMI), race/ethnicity, and self-reported good or excellent health.

Each of the selected main effects were evaluated, one at a time, for interaction with age (as time). All interactions made the model fit substantially worse (using BIC), so none were included in the final model (39):

$$\begin{cases} y_i(\text{age}) = \beta_0 + ns(\text{age}) + \beta_1 BMI + \beta_2 \text{race} + \beta_3 \text{health} + b_{i0} + b_{i1}t + \varepsilon_i(\text{age}), \\ ns(\text{age}) = \kappa_0 + \sum_{j=1}^{4} \kappa_j B_j(\text{age}) \\ b_{i0}, \sim \mathcal{N}(0, \Sigma), \\ \varepsilon_i(\text{age}) \sim \mathcal{N}(0, \sigma^2) \end{cases} \tag{39}$$

where time is scaled as age, and $ns(\text{time})$ is a natural cubic splines function with 1 internal knot, producing coefficients $\kappa_0, \ldots, \kappa_4$. Natural cubic splines are basis splines with the basis function constrained to be linear beyond the boundary knots, producing "better-behaved tails".[25] See section 4.1.2 for details. This strategy provides a great degree of flexibility over a simple quadratic shape without resorting to nonparametric methods—this is a crucial feature if we intend to incorporate the longitudinal model into a joint model.

6.2 Time-to-event modeling

In the analysis dataset of 861 participants, 644 (74.8%) participants were living, and 217 (25.2%) were deceased at the time of the last contact. Median baseline age at last follow-up was 82.9 among those still living and 83.6 among the deceased.

All models used age as the time scale. The natural interpretation of survival time in this setting is survival probability as a function of age, rather than as a function of time on study. Also, simulation studies have demonstrated that using time-on-study and adjusting for baseline age yields biased estimates, especially when the cumulative hazard does not have an exponential shape and when there is a statistical association between age and other covariates.[26, 27]

We constructed a Cox proportional-hazards model by considering all covariates that appeared in at least 4 of 7 previously published time-to-event analyses of data from the MrOS study.[28–34] The complete set of candidate covariates included gait speed, self-reported overall health, history of congestive heart failure, history of diabetes mellitus, current or past smoking, and alcohol use. Univariate Kaplan-Meier plots indicated that self-reported overall health, smoking category, and history of diabetes mellitus or congestive heart failure may impact survival time (Figure 7). A Kaplan-Meier plot of based on a dichotomized version of baseline gait speed also indicated that this variable may impact survival time (Figure 8).

A backwards stepwise selection procedure based on Bayesian information criterion (BIC) selected model (40) with the effects of time-varying gait speed and self-reported overall health. For comparison, we also constructed model (41), which includes a single measure of gait speed (at baseline) instead of time-varying gait speed.

$$h_i(\text{age}|\mathbf{x}_i) = h_0(\text{age}) \exp(\gamma_1 \text{health}_i + \gamma_2 \text{gait}_i(\text{age})) \tag{40}$$

$$h_i(\text{age}|\mathbf{x}_i) = h_0(\text{age}) \exp(\gamma_1 \text{health}_i + \gamma_2 [\text{baselinegait}]_i) \tag{41}$$

Kaplan-Meier curves show differential survival curves by each of the selected categorical covariates (Figure 7). A dichotomized version of baseline gait speed (Figure 8) shows a less dramatic difference in survival curves, but participants whose baseline gait speed was below the mean appear to have lower survival probability between ages 83 and 92.

Including gait speed as a time-varying covariate (model (40)) produces a substantially higher estimate of the effect of gait speed than in model (41) (HR of 1.08 for model (41),

Figure 7. Kaplan-Meier curves by self-reported overall health category, smoking category, and diagnoses of diabetes mellitus and congestive heart failure.
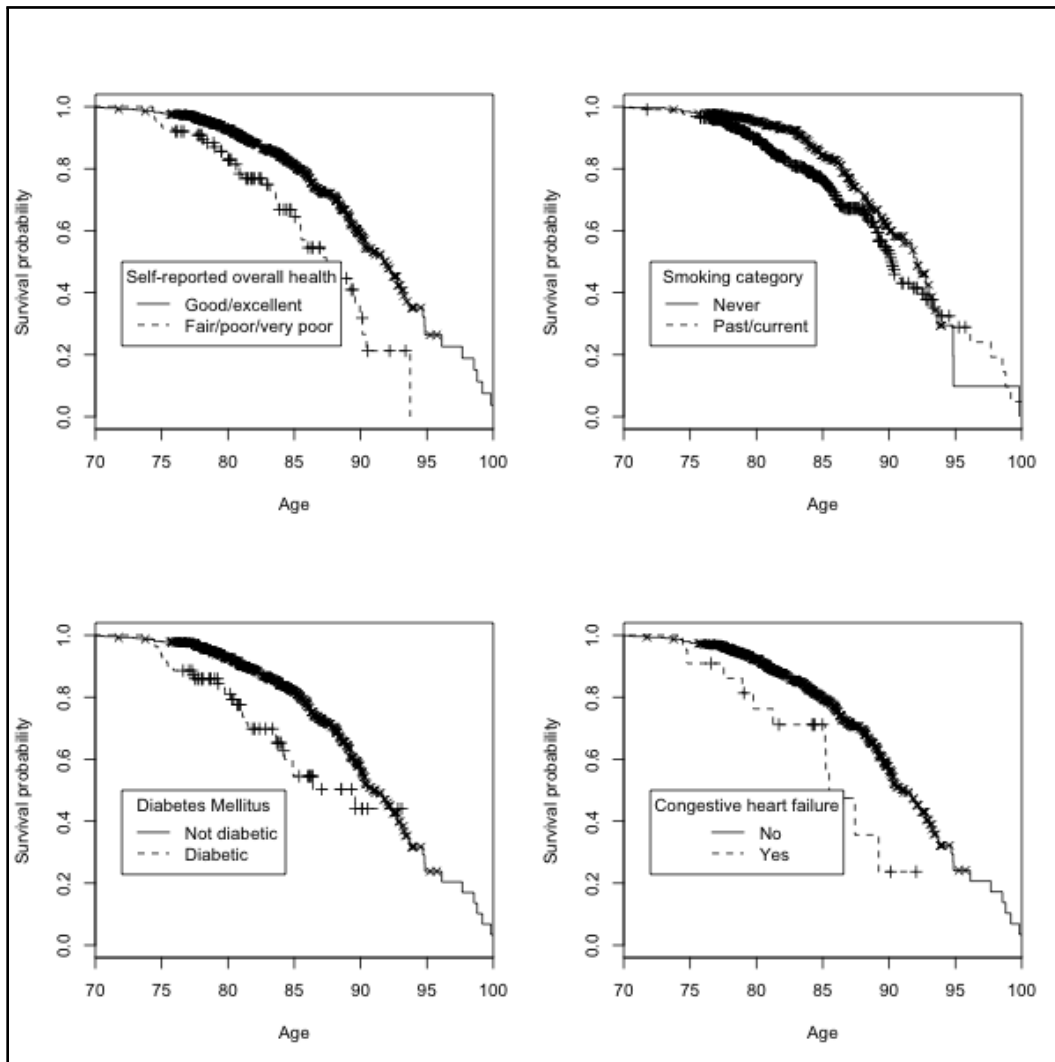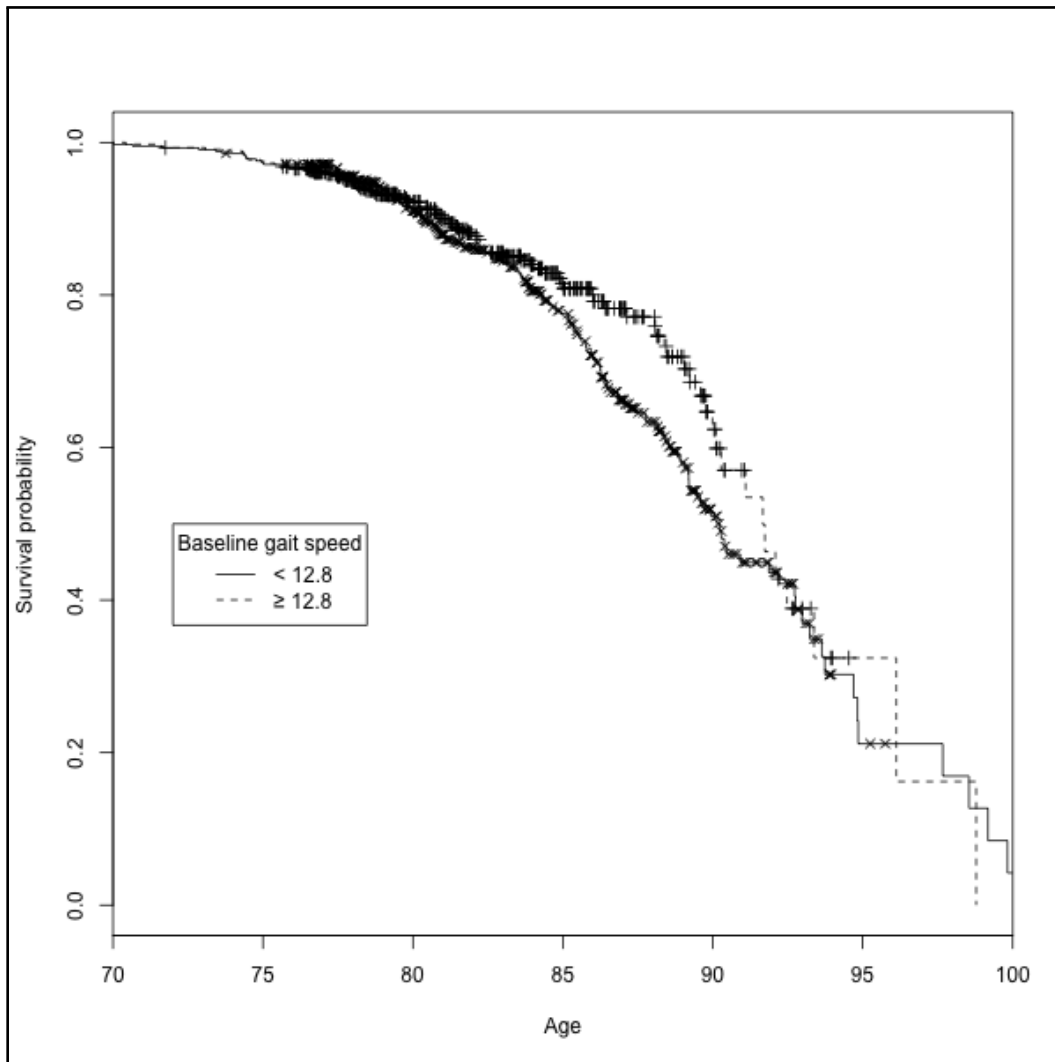
Figure 8. Kaplan-Meier curves by baseline gait speed (dichotomized to less-than versus greater-than-or-equal-to the mean.

compared to 1.14 for model (40)). Estimates of standard error are similar in the 2 models (0.033 for the baseline-only model, versus 0.037 for the time-varying model).

6.3 Joint modeling

The joint model accounts for time-varying changes in gait speed by including model (39) as a longitudinal sub-model. The complete joint model with an unspecified baseline hazard (42) can be expected to produce results similar to those produced by model (40). Unfortunately, this model does not converge under the R function `jointModel()`.

$$
\begin{cases}
h(\text{age}|\mathbf{x}_i) = h_0(\text{age}) \exp\left(\gamma_1 \text{health}_i + \alpha m_i(\text{age})\right), \\
h_0(\text{age}) = 1, \\
y_i(\text{age}) = m_i(\text{age}) + \varepsilon_i(\text{age}), \\
m_i(\text{age}) = \beta_0 + ns(\text{age}) + \beta_1 \text{BMI} + \beta_2 \text{race} + \beta_3 \text{health} + b_{i0} + b_{i1} \text{age}, \\
ns(\text{age}) = \kappa_{i1} N_1(\text{age}) + \kappa_{i2} N_2(\text{age}) \\
\quad b_{i0}, \sim \mathcal{N}(0, \Sigma) \quad \varepsilon_i(\text{age}) \sim \mathcal{N}(0, \sigma^2)
\end{cases}
\tag{42}
$$

A better approach is to use a parametric method to model the baseline hazard, as Rizopoulos suggests.[8] A simple option is to use a Weibull baseline hazard, but this parametric model does not offer the same level of flexibility as piecewise or b-splines-based approaches. A piecewise-constant model with 6 internal knots (43),

$$
h_0(\text{age}) = \sum_{d=1}^{6} \kappa_q I(\xi_{d-1} < \text{age} \leq \xi_d),
\tag{43}
$$

where each κ_d is a constant hazard on interval (ξ_{d-1}, ξ_d) provides the flexibility to specify a constant hazard that changes as a step function of time. Another flexible alternative is a linear (second-order) b-splines model with a single internal knot:

$$
\log h_0(\text{age}) = \sum_{d=1}^{3} \kappa_d B_d(\text{age}, Q = 2),
\tag{44}
$$

where $B_d(\cdot)$ is the b-splines function defined in (13). This model yields 3 coefficients and provides allows different slopes over each of the 2 splines while enforcing continuity at the internal knot and linearity beyond the boundaries.

By BIC, the joint model with a b-splines approximated baseline hazard is the best fitting (BIC = 17433, compared to 17548 and 17594 for the joint models with piecewise-constant and Weibull baseline hazards, respectively).

6.4 Results

Table 3 summarizes parameter estimates and confidence intervals from both Cox models and all 3 joint models. As expected, gait speed declines over time in all sub-groups. Figure 10 displays predicted longitudinal trajectories by race and self-reported health category, with BMI fixed at the mean 27.32. Estimates of the longitudinal parameters were similar in the linear mixed-effects model and all 3 joint models, indicating that the longitudinal estimates are not particularly sensitive to modeling strategy or, by extension to assumptions about missing-data patterns. Contrary to expectation, standard errors and confidence intervals for longitudinal estimates were also very similar across modeling strategy.

In table 3, $\hat{\lambda}$ denotes the Weibull shape parameter for the joint model with a Weibull-shaped baseline hazard. In the joint model with a piecewise-constant baseline hazard, the parameters $(\hat{\kappa}_1, \ldots, \hat{\kappa}_6)$ correspond to the estimates of the mean constant hazard in each of 6 equally spaced intervals from age 65 to age 90. In the joint model with a b-splines-approximated baseline hazard, the parameters κ_1 and κ_2 correspond to the overall hazard intercept and slope, and the parameter κ_3 corresponds to the shift in slope at age 81.

Marginal survival curves were similar among the 3 models (see Figure 9). The jaggedness of the models with piecewise and b-splines baseline hazards reflects the underlying piecewise changes in baseline hazard. The small uptick on the leftmost side of the model with a b-splines baseline hazard results from the left tail of the b-splines function and is typical odd behavior in the tails of b-splines.[8, 17]

The modeled longitudinal trajectory of gait speed, stratified by race and by self-reported overall health category, is an identical shape in all strata, because there were no interactions among race, self-reported health, or age. Nevertheless, there are clear differences in baseline gait speed among the strata (Figure 10).

Gait speed was significantly associated with mortality regardless of the modeling strategy. The best-fitting joint model predicts a 25% increase in risk of death for each 0.1 m/sec decline in gait speed. Estimates of predictors of gait speed were not sensitive to modeling strategy (or, by extension, to assumptions about missing data). Compared to the joint model, the extended Cox model underestimated the effect of longitudinal gait speed and overestimated the effect of self-reported health on survival time (see Table 3). Contrary to

Figure 9. Marginal survival curves from the 3 joint models. Internal knots in the piecewise-constant model are indicated by vertical dotted lines. The single internal knot in the b-splines approximated baseline hazard is at age 81.
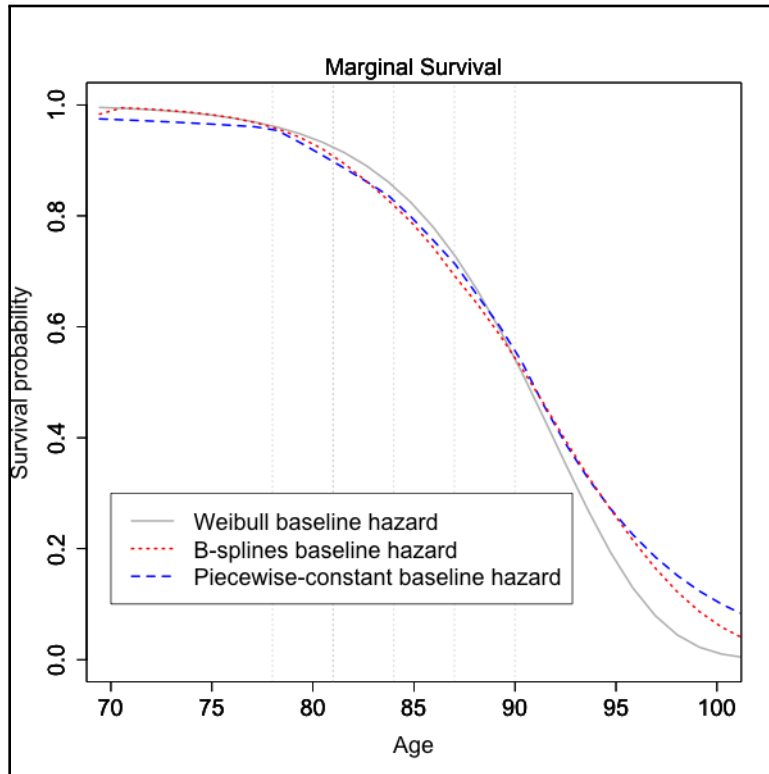
expectation, standard errors were similar in the separate and joint models.

These results add to existing literature on the association between gait speed and mortality and offer a method to avoid the attenuation of effect size that results from measurement error when the underlying longitudinal process is better modeled as a continuous curve than as an extended Cox-style step function.

Table 3. Longitudinal and event parameter estimates from separate and joint models of longitudinal gait speed and survival time among older men from the MrOS study.

	Cox*	LME/extended Cox†	Joint (Weibull)‡		Joint (piecewise)§		Joint (b-splines)‖
BIC (Joint models only)			17594		17548		17433
Event estimates (exponentiated to obtain HRs)							
Fair/Poor/Very Poor health	1.90 (1.31, 2.75)	1.80 (1.18, 2.73)	1.74 (1.17, 2.59)		0.96 (0.64, 1.42)		1.43 (0.97, 2.13)
Gait speed	1.08 (1.01,1.15)	1.14 (1.05,1.22)	1.16 (1.06, 1.26)		1.53 (1.41, 1.65)		1.25 (1.15, 1.36)
Baseline hazard parameters							
			17.10 (15.24, 19.17)	λ	0.37 (0.12, 1.19)	κ_1	0.00 (0.00, 0.00)
					2.65 (1.05, 6.67)	κ_2	0.33 (0.12, 0.87)
					2.45 (1.01, 5.97)	κ_3	2.77 (1.15, 6.69)
					3.66 (1.56, 8.56)	κ_4	
					4.72 (2.04, 10.93)	κ_5	
					7.58 (3.34, 17.20)	κ_6	
Longitudinal estimates							
Intercept		16.40 (15.49, 17.31)	16.45 (15.53, 17.36)		16.75 (15.93, 17.57)		16.40 (15.58, 17.22)
ns(age,2)1		-5.16 (-5.70, -4.61)	-5.27 (-5.82, -4.72)		-5.92 (-6.36, -5.48)		-5.28 (-5.74, -4.81)
ns(age,2)2		-6.15 (-6.92, -5.37)	-6.25 (-7.05, -5.46)		-6.11 (-6.76, -5.47)		-6.35 (-7.00, -5.69)
Body mass index		-0.09 (-0.12, -0.05)	-0.09 (-0.12, -0.05)		-0.09 (-0.12, -0.06)		-0.09 (-0.11, -0.06)
Race: African-American		-1.32 (-1.77, -0.87)	-1.33 (-1.78, -0.88)		-1.27 (-1.69, -0.85)		-1.21 (-1.63, -0.79)
Race: Other		-0.77 (-1.36, -0.18)	-0.77 (-1.36, -0.18)		-0.80 (-1.33, -0.27)		-0.81 (-1.34, -0.29)
Fair/Poor/Very Poor health		-1.49 (-1.86, -1.12)	-1.49 (-1.87, -1.12)		-1.42 (-1.76, -1.09)		-1.44 (-1.77, -1.10)

* Cox proportional-hazards model with baseline CD4 count only.
† Linear mixed-effects model (top portion) and Cox proportional-hazards model with time-varying covariate (bottom portion).
‡ Joint model with Weibull baseline hazard function (method = "weibull-PH-aGH").
§ Joint model with piecewise-constant baseline hazard function (method = "piecewise-PH-aGH").
‖ Joint model with b-splines-approximated baseline hazard function (method = "splines-PH-aGH").

Figure 10. Predicted longitudinal trajectories of gait speed by race and self-reported health covariate values. BMI is fixed at the mean 27.32. The single internal knot for the natural cubic splines is indicated by a dotted vertical line at age 75.
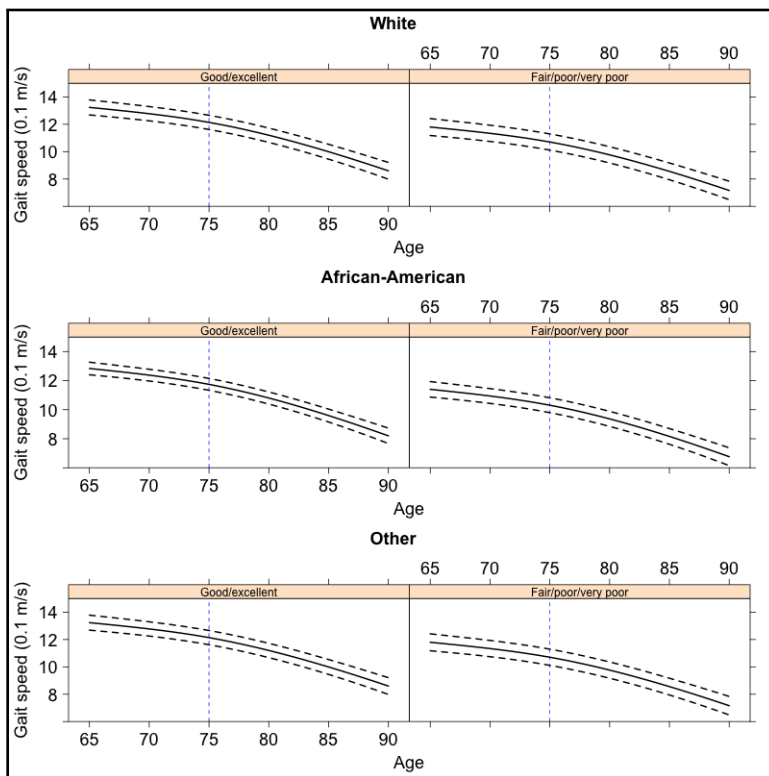
The `JM` package is capable of producing subject-specific survival curves. In case of a prospective study, it can dynamically update survival curves at each longitudinal observation. For the current retrospective study, it can be informative to examine the top and bottom 10 survival curves (Figure 11) to get a sense of how their gait speeds differ at each observation. Table 4 shows that those individuals with the highest survival probability had substantially faster gait speeds. Absolute differences from first to last observation in gait speeds were very similar, but if we consider this change as a proportion of baseline gait speed, those with lower survival probability had a much larger decline.

Figure 11. Subject-specific survival curves, with the highest and lowest 12 curves from each model isolated. Table 4 shows the mean gait speeds at each longitudinal measurement for these selected individuals.
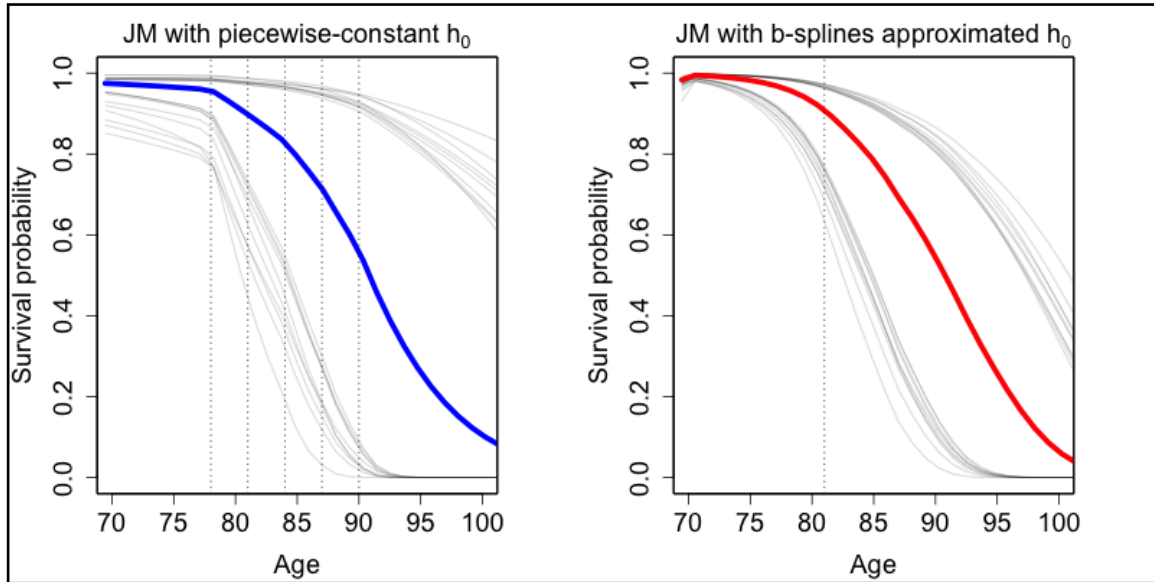

Table 4. Mean gait speed at each visit among individuals with highest and lowest survival curves (see Figure 11).

Visit	Lowest curves	Highest curves
Baseline	8.72	17.88
Follow-up 1	8.55	17.50
Follow-up 2	7.80	15.68
Follow-up 3	7.68	16.43
Follow-up 4	7.64	16.76
Net change	-1.08	-1.12

6.5   Additional modeling: Time-varying self-reported overall health

As an exploratory analysis, we constructed an additional extended Cox model that included a time-varying version of the self-reported overall health variable (collected at baseline (visit 1), the dental visit (visit 2), and the 2 main study follow-up visits (visits 4 and 5). We also constructed a joint model, with time-varying self-reported health in the event sub-model and gait speed in the longitudinal sub-model (45).

$$
\begin{cases}
h(\text{age}|\mathbf{x}_i) = h_0(\text{age}) \exp\left(\gamma_1 \text{health}(age)_i + \alpha m_i(\text{age})\right), \\[2mm]
h_0(\text{age}) = \log h_0(\text{age}) = \displaystyle\sum_{d=1}^{3} \kappa_d B_d(\text{age}, Q = 2), \\[2mm]
y_i(\text{age}) = m_i(\text{age}) + \varepsilon_i(\text{age}), \\
m_i(\text{age}) = \beta_0 + ns(\text{age}) + \beta_1 \text{BMI} + \beta_2 \text{race} + \beta_3 \text{health} + b_{i0} + b_{i1}\text{age}, \\
ns(\text{age}) = \kappa_{i1} N_1(\text{age}) + \kappa_{i2} N_2(\text{age}) \\
\quad b_{i0}, \sim \mathcal{N}(0, \Sigma) \quad \varepsilon_i(\text{age}) \sim \mathcal{N}(0, \sigma^2)
\end{cases}
\tag{45}
$$

By including updated information about self-reported overall health at the time of estimating the hazard, we arrive at substantially attenuated estimates of the association between longitudinal gait speed and mortality (see Table 5). In other words, model (45) reveals time-varying self-reported health to be a confounder of the association between longitudinal gait speed and mortality.

It is not terribly surprising to see that the effect of this very general question overwhelms a specific physical function measure such as gait speed. Since self-reported overall health is a very non-specific measure, there are a number of possibilities for the causal relationships at play here. For example, perhaps some unmeasured health event, captured by a change in the self-reported overall health variable, is causing a slower gait speed, and both of these variables are impacting survival probability; or perhaps slower gait speed is causing patients to report poorer overall health, and both of these variables are impacting survival probability. A limitation of the models reported here is that they do not include updated information on age-related diagnoses or other health events that may be partially captured by changes in self-reported overall health. As a result, there is bound to be residual unmeasured confounding.

Self-reported overall health can be expected to overlap with the effects of a number of confounders–especially new diagnoses of age-related disease–that remain unmeasured in this model, so it may be of limited clinical utility compared to the clinical information available to a treating provider.

Table 5. Longitudinal and event parameter estimates from separate and joint models of longitudinal gait speed and survival time among older men from the MrOS study.

	Extended Cox* Mixed-effects model‡	Extended Cox†	Joint model§	Joint model‖
Form of self-reported health	single measure	time-varying	single measure	time-varying
BIC (joint models)			17433	17281
Event estimates (exponentiated to obtain HRs)				
Fair/Poor/Very Poor health	1.80 (1.18, 2.73)	2.56 (1.79, 3.67)	1.43 (0.97, 2.13)	2.96 (2.18, 4.04)
Gait speed	1.14 (1.05,1.22)	1.11 (1.03, 1.19)	1.25 (1.15, 1.36)	1.18 (1.09, 1.28)
Longitudinal estimates				
Intercept	16.40 (15.49, 17.31)		16.40 (15.58, 17.22)	14.93 (14.02, 15.85)
ns(age,2)1	-5.16 (-5.70, -4.61)		-5.28 (-5.74, -4.81)	-5.22 (-5.69, -4.75)
ns(age,2)2	-6.15 (-6.92, -5.37)		-6.35 (-7.00, -5.69)	-6.23 (-6.89, -5.57)
Body mass index	-0.09 (-0.12, -0.05)		-0.09 (-0.11, -0.06)	-0.08 (-0.11, -0.06)
Race: African-American	-1.32 (-1.77, -0.87)		-1.21 (-1.63, -0.79)	-1.22 (-1.64, -0.8)
Race: Other	-0.77 (-1.36, -0.18)		-0.81 (-1.34, -0.29)	-0.81 (-1.34, -0.29)
Fair/Poor/Very Poor health	-1.49 (-1.86, -1.12)		-1.44 (-1.77, -1.10)	1.43 (1.09, 1.76)

* Cox model with time-varying gait speed and a single measure of self-reported health (top portion).
† Cox model with time-varying gait speed and time-varying self-reported health (top portion).
‡ Mixed-effects model of longitudinal gait speed (bottom portion).
§ Joint model with baseline-only self-reported health (b-splines-approximated baseline hazard).
‖ Joint model with time-varying self-reported health (b-splines-approximated baseline hazard).

6.6 Discussion

This thesis investigated the utility of using joint models toward three ends:

- Improving estimates of the association between a longitudinal process and a time-to-event process;

- improving estimates of other longitudinal and time-to-event parameters;

- and accounting for non-ignoreable (MNAR) missing-data patterns.

In this investigation, the longitudinal covariate, gait speed, was a continuous, endogenous outcome, measured with error, and as such was a good candidate for joint modeling. As expected, the extended Cox model underestimated the effect of gait speed on survival time. The results indicate that the association between a longitudinal decline in gait speed is more important than the association between a single measure of gait speed in this population of older men. Providers may benefit during treatment planning by considering this higher estimate of association. This work implies that traditional proportional-hazards methods may underestimate the magnitude of association between endogenous or continuous covariates and time to an event.

6.6.1 Future directions

Future research along this line in the MrOS study would do well to consider other functional forms of the longitudinal gait speed covariate—especially a joint model that considers the effect of the rate of change of gait speed at hazard time t, rather than the current gait speed at time t, as in this investigation. It might also be particularly illuminating to examine the effects of a variety of other longitudinal covariates, especially counts of falls or fractures—both of which are longitudinal variables collected in the MrOS datasets—although these analyses will depend on the future release of software for producing generalized linear mixed joint models for discrete longitudinal outcomes, as with Dr. Viviani's forthcoming R package. It may also be useful to consider a vector of longitudinal covariates, including longitudinal changes in diagnoses of chronic diseases such as diabetes mellitus or congestive heart failure, or longitudinal changes in other health markers such as smoking status.

Since the MrOS study collects a number of discrete measures that change over time, the MrOS community may particularly benefit from emerging methods in generalized linear

mixed joint models that can account for longitudinal changes in discrete variables.[35] An R package accommodating this sort of model is expected to be published to the Comprehensive R Archive Network (CRAN) shortly.

6.6.2 Summary

Under certain circumstances, researchers may benefit from using the joint modeling framework when their primary research objective involves either a longitudinal or a time-to-event outcome:

Longitudinal outcome: If the longitudinal data suffer from non-ignorable missingness, and data are available for the event responsible for the missingness, a joint model will correct the bias that would otherwise result in a mixed-effects model. If it is unclear whether non-ignorable missingness is present, a joint model provides a good sensitivity analysis. If, as in the analysis of longitudinal gait speed here, the longitudinal estimates are not sensitive to modeling strategy, one can conclude that missingness is not important to estimates of longitudinal parameters.

Event outcome: If there are important covariates that are endogenous, naturally continuous, or measured with error, a joint model will correct underestimation of effect that is likely to result from modeling time to the event under the proportional hazards framework alone.

References

[1] Dumurgier J, Elbaz A, Ducimetière P, Tavernier B, Alpérovitch A, Tzourio C. Slow walking speed and cardiovascular death in well functioning older adults: prospective cohort study. *BMJ: British Medical Journal*. 2009;339 .

[2] Cooper R, Kuh D, Hardy R, et al. Objectively measured physical capability levels and mortality: systematic review and meta-analysis. *BMJ: British Medical Journal*. 2010; 341 .

[3] Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, et al. Gait speed and survival in older adults. *JAMA: the journal of the American Medical Association*. 2011; 305(1):50–58 .

[4] Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics*. 2000;1(4):465–480 .

[5] Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, et al. Design and baseline characteristics of the osteoporotic fractures in men (mros) study—a large observational study of the determinants of fracture in older men. *Contemporary clinical trials*. 2005;26(5):569–585 .

[6] Fitzmaurice G VGMG Davidian M. *Longitudinal Data Analysis*. 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742: Chapman & Hall/CRC. 2009.

[7] Fitzmaurice GM, Laird NM, Ware JH. *Applied longitudinal analysis*. Hoboken, NJ: John Wiley & Sons, 2 edition. 2011.

[8] Rizopoulos D. *Joint Models for Longitudinal and Time-to-event Data With Applications in R*. Boca Raton, FL: CRC Press. 2012.

[9] Guo X, Carlin BP. Separate and joint modeling of longitudinal and event time data using standard computer packages. *The American Statistician*. 2004;58(1):16–24 .

[10] Tsiatis A, Degruttola V, Wulfsohn M. Modeling the relationship of survival to longitudinal data measured with error. applications to survival and cd4 counts in patients with aids. *Journal of the American Statistical Association*. 1995;90(429):27–37 .

[11] Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics*. 1997;330–339 .

12 Pawitan Y, Self S. Modeling disease marker processes in aids. *Journal of the American Statistical Association*. 1993;88(423):719–726 .

13 Phipps KR, Chan BK, Jennings-Holt M, Geurs NC, Reddy MS, Lewis CE, et al. Periodontal health of older men: the mros dental study. *Gerodontology*. 2009;26(2):122–129 .

14 Blackwell T, Yaffe K, Ancoli-Israel S, Redline S, Ensrud KE, Stefanick ML, et al. Association of sleep characteristics and cognition in older community-dwelling men: the mros sleep study. *Sleep*. 2011;34(10):1347–1356 .

15 Lung NH, Institute B. Outcomes of sleep disorders in older men. 2009. URL http://clinicaltrials.gov/show/NCT00070681.

16 Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics*. 1982; 963–974 .

17 Hastie T, Tibshirani R, Friedman J, Hastie T, Friedman J, Tibshirani R. *The elements of statistical learning*, volume 2. Springer. 2009.

18 Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica*. 2004;14(3):809–834 .

19 Tableman M, Kim JS. *Survival analysis using S: analysis of time-to-event data*. CRC Press. 2003.

20 Kleinbaum K, Nizam M. *Applied Regression Analysis and Other Multivariable Methods*. Belmont, CA: Thompson, 4 edition. 2008.

21 Hsieh F, Tseng YK, Wang JL. Joint modeling of survival and longitudinal data: likelihood approach revisited. *Biometrics*. 2006;62(4):1037–1043 .

22 Abrams DI, Goldman AI, Launer C, Korvick JA, Neaton JD, Crane LR, et al. A comparative trial of didanosine or zalcitabine after treatment with zidovudine in patients with human immunodeficiency virus infection. *New England Journal of Medicine*. 1994; 330(10):657–662 .

23 Hoffmann CJ, Schomaker M, Fox MP, Mutevedzi P, Giddy J, Prozesky H, et al. Cd4 count slope and mortality in hiv-infected patients on antiretroviral therapy: Multicohort analysis from south africa. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. 2013;63(1):34–41 .

[24] Mills EJ, Bakanda C, Birungi J, Mwesigwa R, Chan K, Ford N, et al. Mortality by baseline cd4 cell count among hiv patients initiating antiretroviral therapy: evidence from a large cohort in uganda. *Aids*. 2011;25(6):851–855 .

[25] Chambers JM, Hastie T, et al. *Statistical models in S*. Chapman & Hall London. 1992.

[26] Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. *Am J Epidemiol*. 1997;145(1):72–80 .

[27] Thiébaut A, Bénichou J. Choice of time-scale in cox's model analysis of epidemiologic cohort data: a simulation study. *Statistics in medicine*. 2004;23(24):3803–3820 .

[28] Cawthon PM, Parimi N, Barrett-Connor E, Laughlin GA, Ensrud KE, Hoffman AR, et al. Serum 25-hydroxyvitamin d, parathyroid hormone, and mortality in older men. *Journal of Clinical Endocrinology & Metabolism*. 2010;95(10):4625–4634 .

[29] Dominguez JR, Kestenbaum B, Chonchol M, Block G, Laughlin GA, Lewis CE, et al. Relationships between serum and urine phosphorus with all-cause and cardiovascular mortality: the osteoporotic fractures in men (mros) study. *American Journal of Kidney Diseases*. 2012; .

[30] Fink H, Kuskowski M, Taylor B, Schousboe J, Orwoll E, Ensrud K. Association of parkinson's disease with accelerated bone loss, fractures and mortality in older men: the osteoporotic fractures in men (mros) study. *Osteoporosis International*. 2008; 19(9):1277–1282 .

[31] Johansson H, Odén A, Kanis J, McCloskey E, Lorentzon M, Ljunggren Ö, et al. Low bone mineral density is associated with increased mortality in elderly men: Mros sweden. *Osteoporosis international*. 2011;22(5):1411–1418 .

[32] Lee CG, Boyko EJ, Nielson CM, Stefanick ML, Bauer DC, Hoffman AR, et al. Mortality risk in older men associated with changes in weight, lean mass, and fat mass. *Journal of the American Geriatrics Society*. 2011;59(2):233–240 .

[33] Paudel ML, Taylor BC, Ancoli-Israel S, Blackwell T, Stone KL, Tranah G, et al. Rest/activity rhythms and mortality rates in older men: Mros sleep study. *Chronobiology international*. 2010;27(2):363–377 .

[34] Waring AC, Harrison S, Samuels MH, Ensrud KE, LeBlanc ES, Hoffman AR, et al. Thyroid function and mortality in older men: a prospective study. *Journal of Clinical Endocrinology & Metabolism*. 2012;97(3):862–870 .

[35] Viviani S, Alfó M, Rizopoulos D. Generalized linear mixed joint model for longitudinal and survival outcomes. *Statistics and Computing*. 2013;1–11 .

[36] Chan BK, Marshall LM, Winters KM, Faulkner KA, Schwartz AV, Orwoll ES. Incident fall risk and physical activity and physical performance among older men the osteoporotic fractures in men study. *American journal of epidemiology*. 2007;165(6):696–703 .

7 Appendix

7.1 Descriptive analysis of CD4 counts and survival time

A histogram of CD4 counts makes it clear at a glance why Rizopoulos[8] used a square-root transformation before modeling changes in the biomarker under an assumption of normality (see figure 1a). Although a square-root transformation doesn't quite achieve symmetry, it certainly moves the mode away from the floor of zero (figure 1b).

Figure 12. Histograms for CD4 and $\sqrt{CD4}$
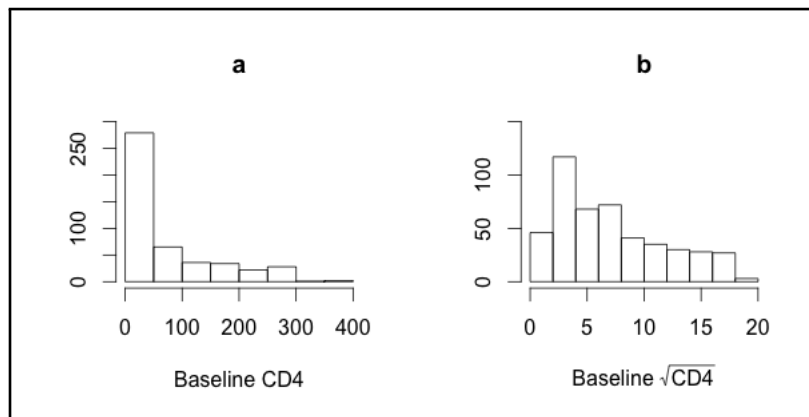
A plot of mean square-root CD4 counts (Figure 13) shows a general downward trend over time, although there are increases at months 2 and 12.

A series of box plots of square-root CD4 count by observation time (Figure 14) indicates that the distributions are somewhat skewed, so to evaluate a linear trend over time, it may be more appropriate to examine median square-root CD4 counts over time.

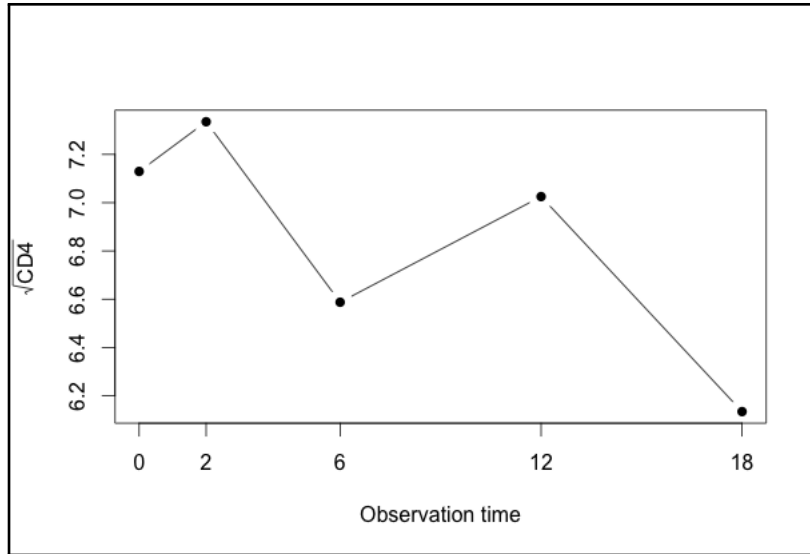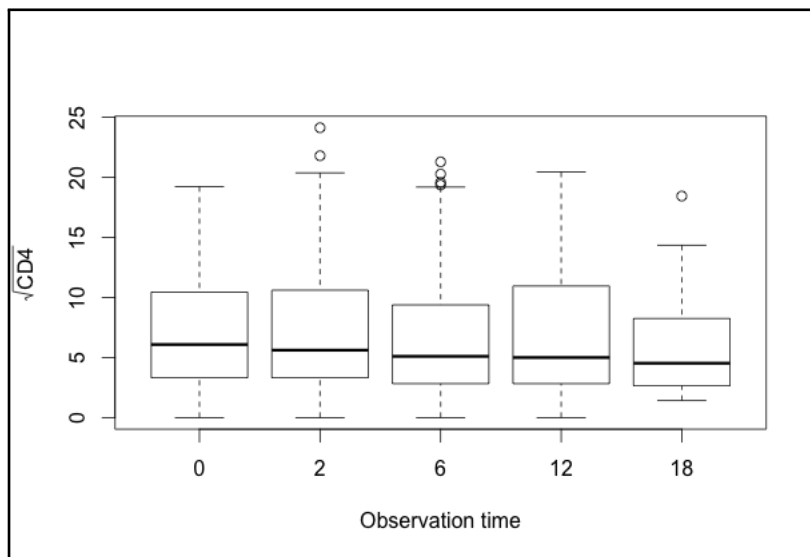**Figure 13.** Mean $\sqrt{\text{CD4}}$ over time

Figure 14. Box plots of $\sqrt{\text{CD4}}$ by observation time
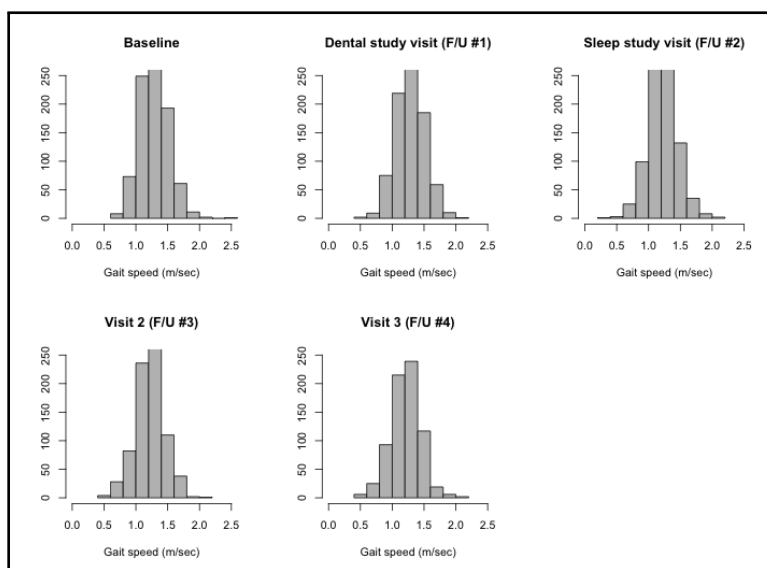

47

7.2 Longitudinal modeling of gait speed from the MrOS study

7.2.1 Distributions of gait speed

Histograms (Figure 15) and QQ-plots (Figure 16) of gait speed at each visit support the assumption of normality: distributions are unimodal and symmetric and almost perfectly aligned with normal quantiles.

Figure 15. Histograms for gait speed at each visit

7.2.2 Longitudinal patterns of change in gait speed

The longitudinal outcome of interest is gait speed, in which MrOS participants walked a 6-m course at their normal walking pace, and the fastest speed (meters per second) was recorded for analysis.[36] Previous research has shown a strong relationship between gait speed and survival.[2,3]

Table 6 summarizes the counts of observations of gait speed and the number of follow-up days at each visit. Together with the spaghetti plots in figure 18, showing longitudinal patterns of gait speed over time for unique random subsets of 50 individuals, this summary indicates a slight decline in mean gait speed over time.

Spaghetti plots do not demonstrate a strong overall longitudinal pattern. Some individual
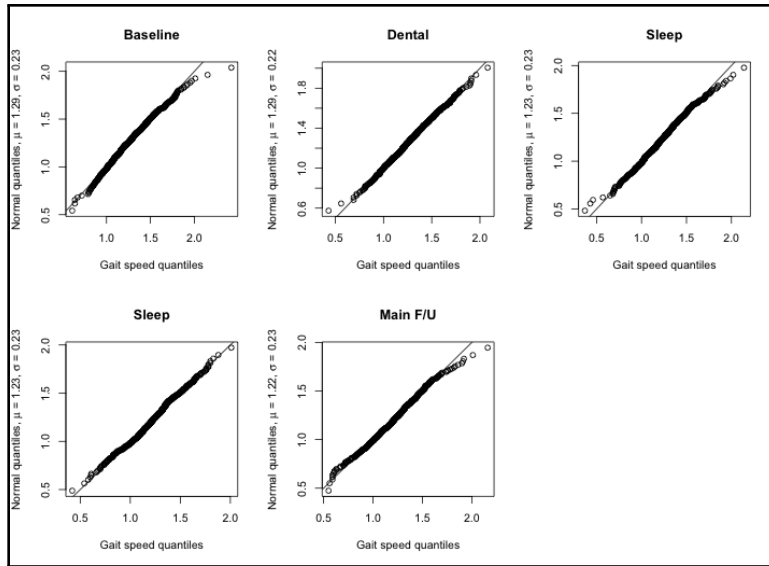
Figure 16. Q-Q plots for gait speed at each visit

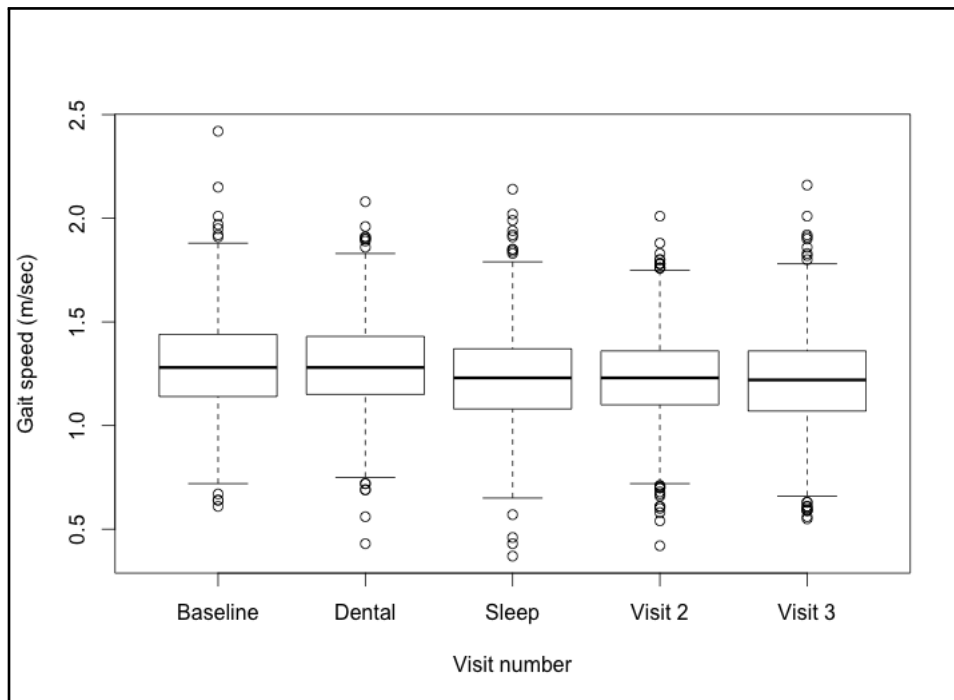**Figure 17.** Box plots for gait speed at each visit

Table 6

	Baseline	Dental visit	Sleep visit	Visit 2	Visit 3
n	861	857	853	779	708
Mean gait speed* (SD)	12.91 (2.24)	12.88 (2.19)	12.29 (2.28)	12.27 (2.24)	12.11 (2.30)
Median age at follow-up (IQR)	72.0 (8.0)	73.9 (8.4)	75.3 (8.3)	76.3 (8.2)	77.9 (7.8)

profiles exhibit a linear trend in gait speed, but others deviate from a linear trend. Differences in baseline measures of narrow-walk pace seem to determine some of the differences between individuals in follow-up measures.

The overall, whole-sample profile (Figure 19) indicates a small but detectable decline in gait speed over the 5 visits, spanning on average 6.75 years. A simple linear regression model predicts a mean baseline gait speed of 1.29 m/sec that declines by 0.01 m/sec each year. Of course, this estimate and the profile plot in Figure 19 should be interpreted with caution, because they do not account for the within-individual correlation that should be expected from repeated measures.

Table 7. Follow-up time (in years) at each study visit.

Visit	N	Mean follow-up time (SD)	Mean time between visits (SD)
Baseline	877	N/A	N/A
Dental visit	870	1.86 (0.40)	1.86 (0.40)
Sleep visit	866	3.39 (0.53)	1.54 (0.35)
Visit 2	792	4.58 (0.36)	1.21 (0.32)
Visit 3	720	6.75 (0.29)	2.19 (0.34)

Figure 18. Spaghetti plots of longitudinal changes in gait speed for 9 subsets of 50 randomly selected individuals
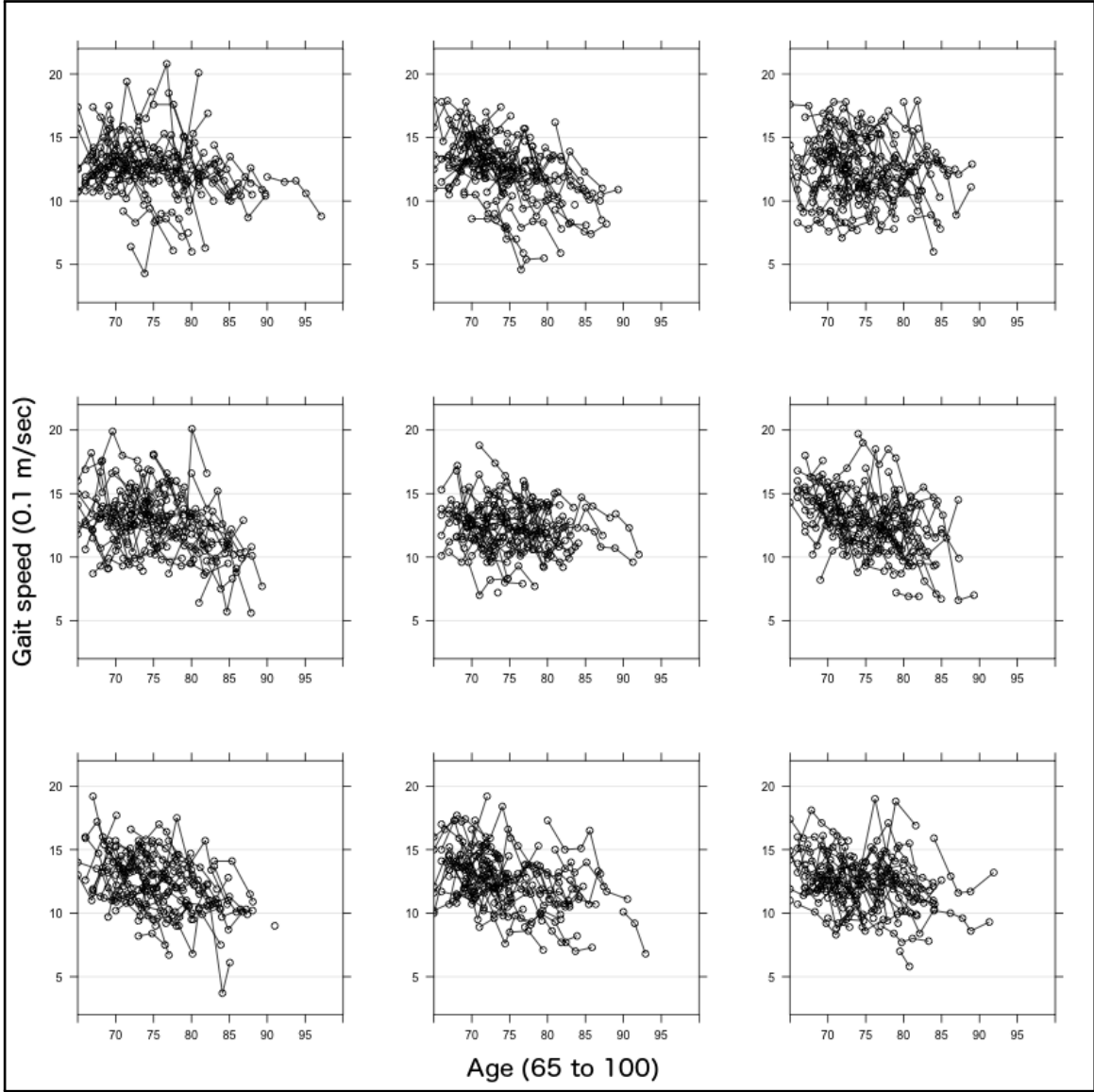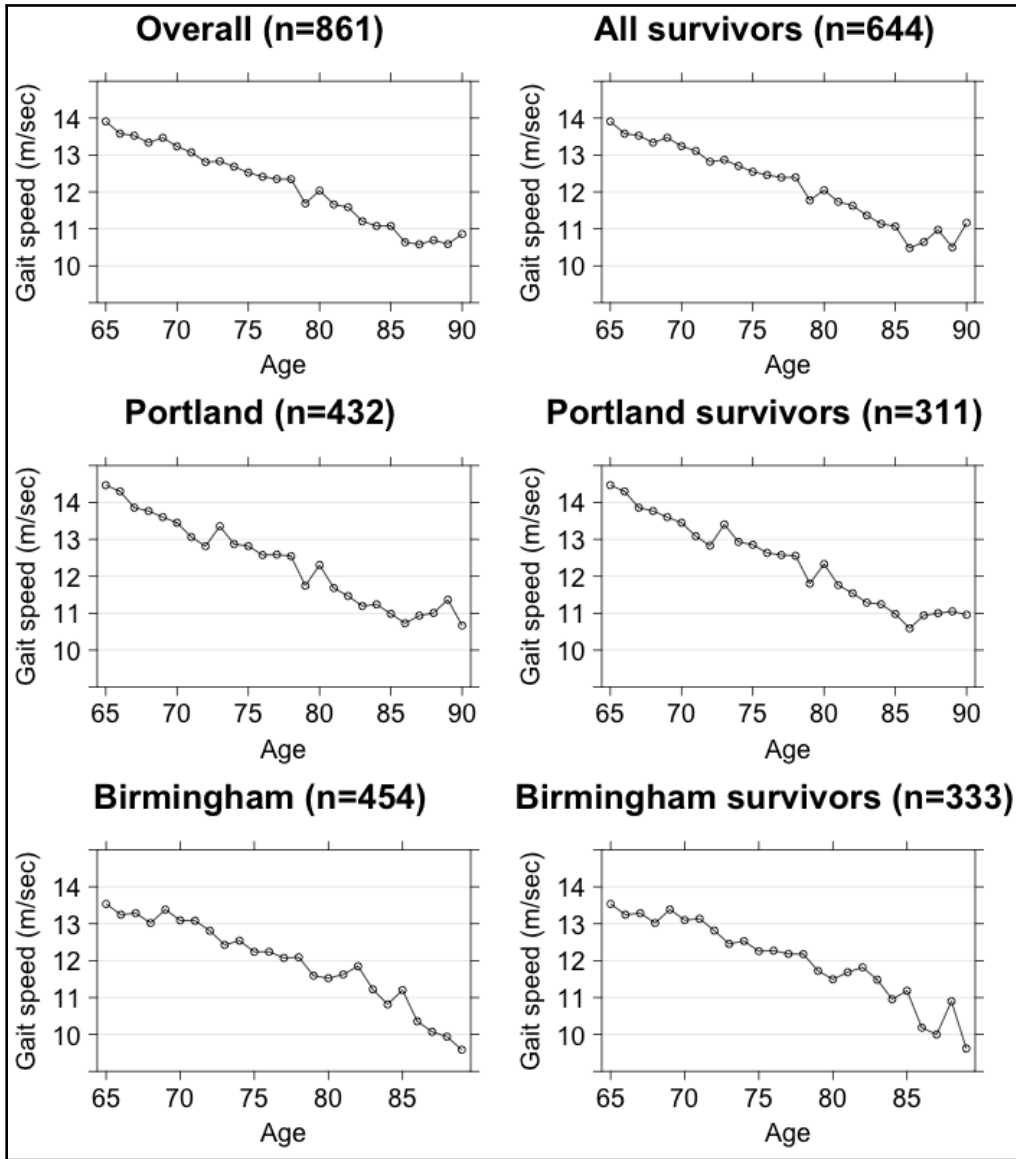

51

Figure 19. Trend in mean gait speed over time for overall sample and stratified by site and survival status at last follow-up.

7.2.3 Missing-data patterns

Of the 861 patients in the sample, 681 (79.1%) had complete ascertainment of gait speed data for all 5 visits; 152 (17.7%) had monotone missing-data patterns (i.e., they had no missing data or had non-intermittent missing data), and 28 (3.3%) had intermittent missing data. A one-way ANOVA test of intercepts indicates a significant difference in intercepts ($F = 10.9$, $p < 0.001$). It will, therefore, be important to account for differential patterns of missing data in the longitudinal analysis.

A much larger proportion of the patients with monotone missing data died than did those with complete ascertainment (62.5% versus 16.4%, $\chi^2 = 138.7$, $p < 0.0001$), so dropouts resulting from death likely account for the difference in longitudinal trajectory of gait speed among different missing-data patterns. Longitudinal profiles categorized by deceased status at last contact show a baseline difference in gait speed (Figure 20).

Figure 20. Gait speed profiles by deceased status.
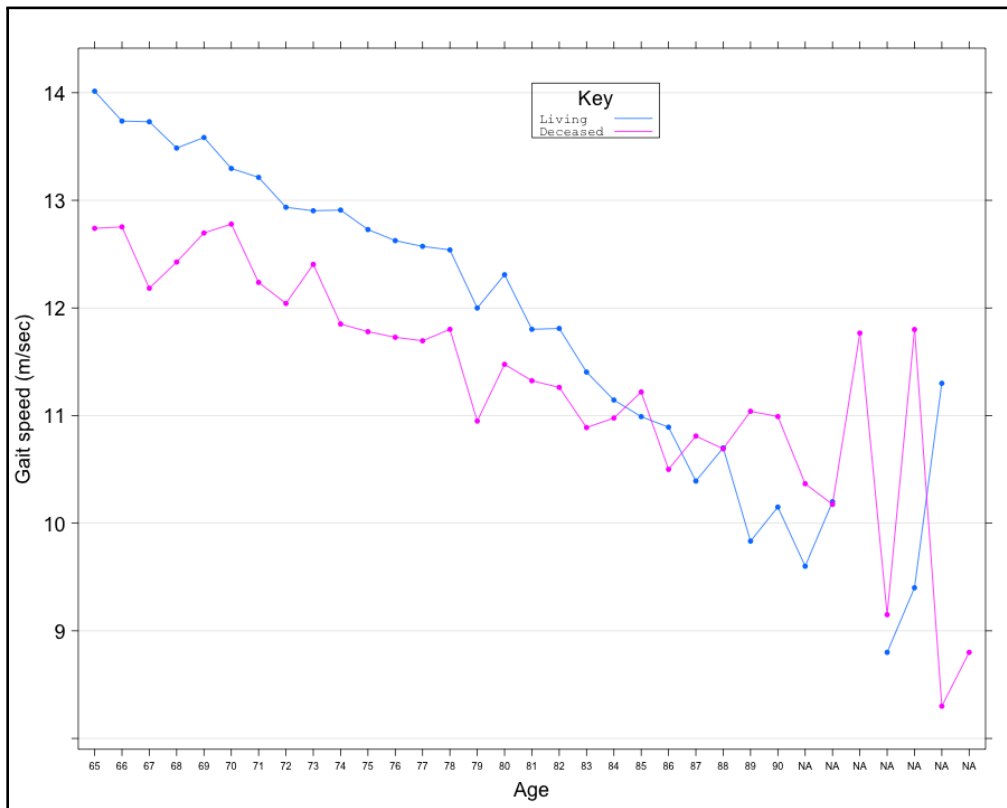
7.2.4 Correlation structure

The baseline measure of gait speed is less correlated with subsequent visits than the subsequent visits are with each other (Figure 21). Otherwise, correlation does not appear to degrade over time.

Figure 21. Scatterplot and correlation matrix for gait speed by follow-up visit (missing values omitted)
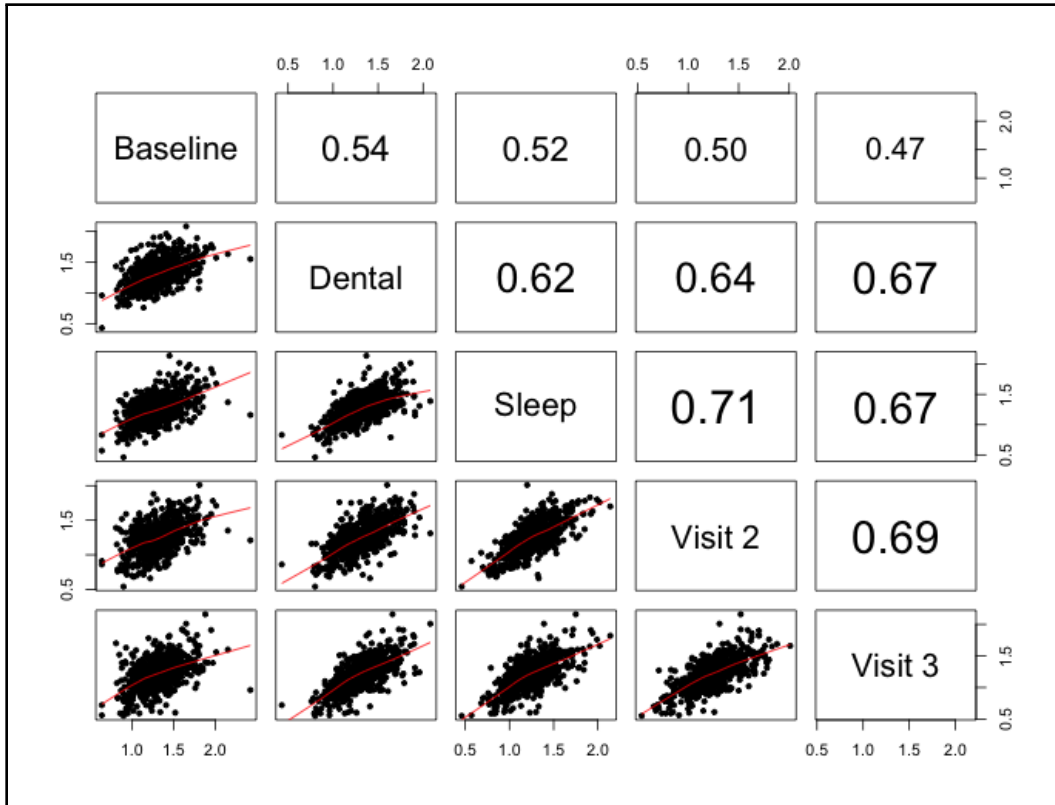
7.2.5 Model selection for longitudinal gait speed

Shape of time

Comparison of a simple linear model to quadratic, cubic, and natural-splines models indicated that higher-order terms and splines models do not improve the fit over a linear model (confirmed by LRT comparing the linear model to the quadratic model ($\chi_1^2 = 1.3$, p=0.26).

Repeating the comparisons with a random intercept in the model indicate that a quadratic

model fits better than a linear model ($\chi^2_1 = 23.4$, $p < 0.001$, $BIC = 15824.88$). A natural-splines model with 1 interior knot fits slightly better than a quadratic model ($BIC = 15823.50$).

Random effects

We compared a model with all candidate covariates with no random effects to the same model with a random intercept only and to the same model with a random intercept and slope, using likelihood ratio tests (summarized in Table 8). Models including polynomial random effects of time failed to converge.

Table 8. Likelihood ratio test comparing a linear model to a mixed-effects model with a random intercept and and mixed-effects model with a random intercept and slope.

Model	df	AIC	BIC	Log Likelihood	Test	-2 (LR)	p-value
Linear model (1)	16	17242.75	17343.68	-8605.37			
Random intercept (2)	17	15659.82	15767.06	-7812.908	(1) vs (2)	1584.9	<0.0001
Random intercept & slope (3)	19	15637.18	15757.04	-7799.592	(2) vs (3)	26.632	<0.0001

The covariance structure obtained from the random intercept and slope model indicates that there is much more variation in baseline values of gait speed than in change over time in gait speed (SD of intercept = 8.2, SD of slope = 0.11). Nevertheless, the likelihood ratio tests indicate that accounting for between-individual differences in slope improves the fit of the model.

Fixed effects

Since several racial and ethnic groups were underrepresented, we recategorized race as "white", "African-American", or "other."

Model selection was a backwards approach, proceeding by removing the covariate with the largest p-value and continuing to remove covariates until Bayesian information criterion (BIC) was no longer reduced. The selection procedure, summarized in Table 9, resulted in the selection of age, body mass index (BMI), race/ethnicity, and self-reported good or excellent health.

Each of the selected main effects were evaluated, one at a time, for interaction with age (as time). All interactions made the model fit substantially worse (by BIC), so none were included in the final model (46).

55

Table 9. Selection of fixed effects.

Fixed effects	BIC
[ns(age,2)] + [BMI] + [smoke] + [race/ethnicity] + [drank in past year] + [good/excellent health] + [non-skin cancer] + [congestive heart failure] + [diabetes] + [BMD]	15784.25
[ns(age,2)] + [study site] + [BMI] + [smoke] + [race/ethnicity] + [good/excellent health] + [non-skin cancer] + [congestive heart failure] + [diabetes] + [BMD]	15769.87
[ns(age,2)] + [study site] + [BMI] + [smoke] + [race/ethnicity] + [good/excellent health] + [non-skin cancer] + [diabetes] + [BMD]	15762.65
[ns(age,2)] + [study site] + [BMI] + [smoke] + [race/ethnicity] + [good/excellent health] + [diabetes] + [BMD]	15754.54
[ns(age,2)] + [study site] + [BMI] + [race/ethnicity] + [good/excellent health] + [diabetes] + [BMD]	15740.04
[ns(age,2)] + [BMI] + [race/ethnicity] + [good/excellent health] + [diabetes] + [BMD]	15733.52
[ns(age,2)] + [BMI] + [race/ethnicity] + [good/excellent health] + [diabetes]	15730.43
[ns(age,2)] + [BMI] + [race/ethnicity] + [good/excellent health]	**15725.96**
[ns(age,2)] + [BMI] + [good/excellent health]	15744.79

$$
\begin{cases}
y_i(t) = \beta_0 + ns(t) + \beta_1 BMI + \beta_2 race + \beta_3 health + b_{i0} + b_{i1}t + \varepsilon_i(t), \\
ns(t) = \sum_{d=1}^{3} \kappa_j N_d(t) \\
b_{i0}, \sim \mathcal{N}(0,\Sigma), \\
\varepsilon_i(t) \sim \mathcal{N}(0,\sigma^2)
\end{cases}
\tag{46}
$$

7.2.6 Diagnostics

A plot of conditional residuals versus fitted values (Figure 22) is not particularly helpful for establishing fit or homoscedasticity, because it fails to account for correlation within individuals; nevertheless, it identifies 12 outliers to be removed from the analysis to obtain the analysis dataset of n=861.

Figure 22. Conditional residuals versus fitted values for random intercept model of narrow-walk pace.
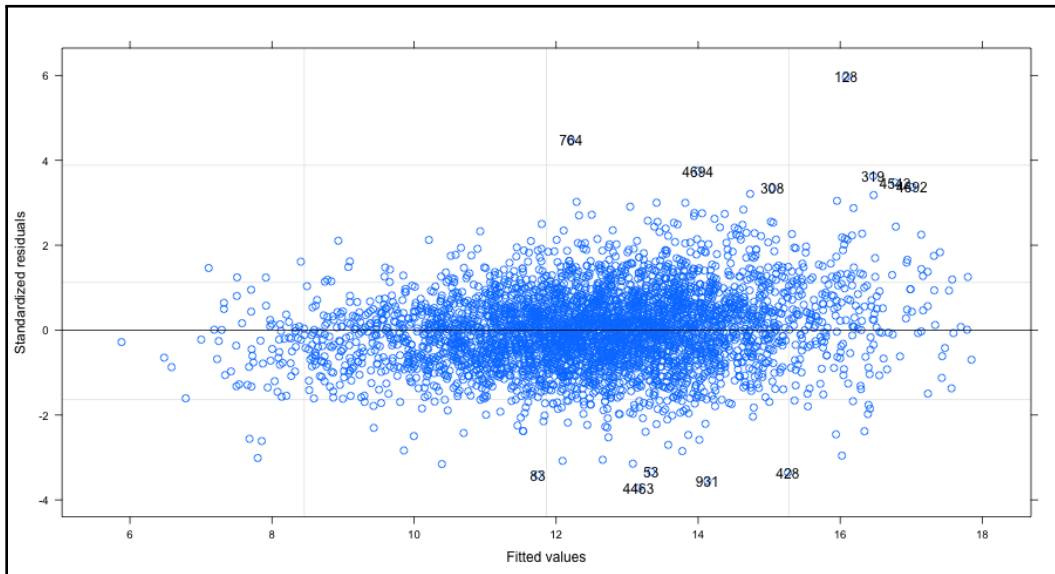
7.3 Time-to-event modeling of survival time from the MrOS study

The cumulative hazard 23 has an approximately exponential shape, indicating that the shape of the hazard does not prohibit using time-on-study as the time scale; nevertheless, many of the covariates under consideration are diseases of aging and can be assumed to be associated with age, so it may be better to model age as time.

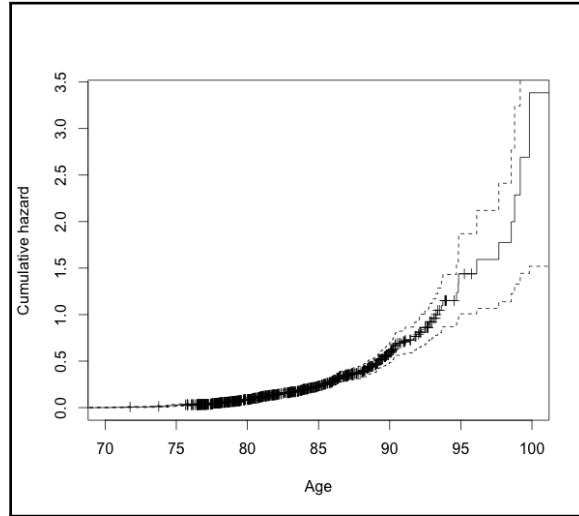**Figure 23.** Cumulative hazard, using age as time scale.

Figure 24. Overall Kaplan-Meier curve for survival among MrOS participants.
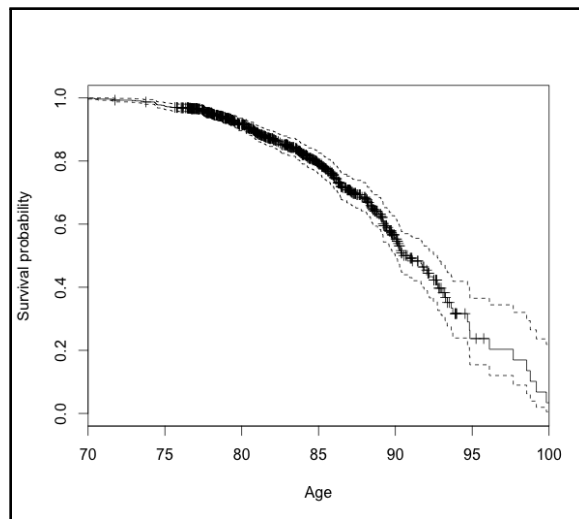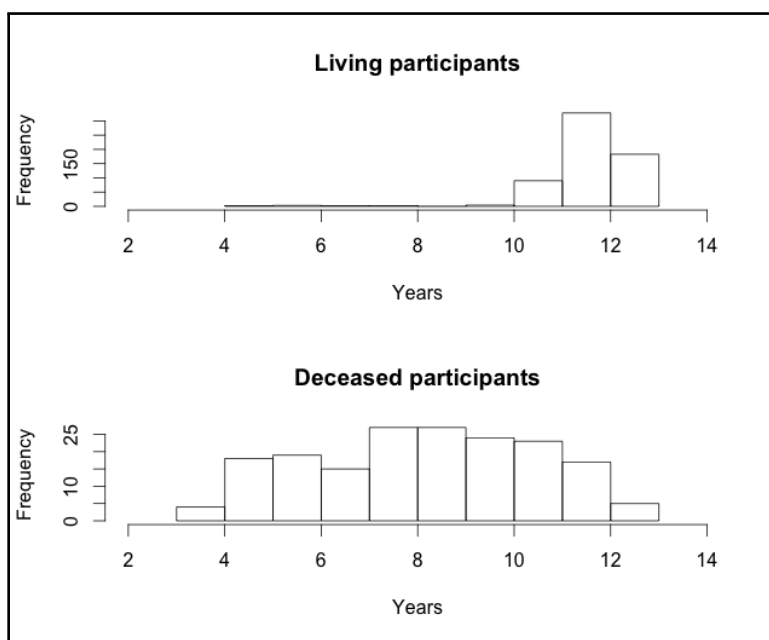

58

Figure 25. Histograms for follow-up time among living and deceased participants

7.4 Joint modeling of longitudinal changes in gait speed and survival time from the MrOS study

7.4.1 B-splines baseline hazard function

The baseline hazard function defined in (44) is a locally defined 2nd-order b-splines function. What follows here is a demonstration of the algebra for calculating the local shape for an arbitrarily selected value of the x variable, *age*. This process finishes by summing the right-most column of calculations to obtain the local definition of the b-splines function.
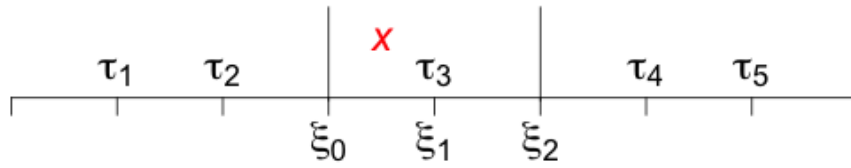
For $\tau_2 \leq X < \tau_3, \quad \tau_2 = 0.30, \quad \tau_3 = \xi_1 = 81.00$:

$$
\begin{pmatrix}
B_{11} = 0 & B_{12} = \dfrac{x - \tau_1}{\tau_2 - \tau_1} 0 + \dfrac{\tau_3 - x}{\tau_3 - \tau_2} 1 = \dfrac{\tau_3 - x}{\tau_3 - \tau_2} \\[2ex]
B_{21} = 1 & B_{22} = \dfrac{x - \tau_2}{\tau_3 - \tau_2} 1 + \dfrac{\tau_4 - x}{\tau_4 - \tau_3} 0 = \dfrac{x - \tau_2}{\tau_3 - \tau_2} \\[2ex]
B_{31} = 0 & B_{32} = \dfrac{x - \tau_3}{\tau_4 - \tau_3} 0 + \dfrac{\tau_5 - x}{\tau_5 - \tau_4} 0 = 0 \\[2ex]
B_{41} = 0 &
\end{pmatrix}
$$

$$
\Rightarrow \log h_0(\text{age})_{\tau_2 \leq X < \tau_3} = \kappa_1 \frac{\tau_3 - x}{\tau_3 - \tau_2} + \kappa_2 \frac{x - \tau_2}{\tau_3 - \tau_2} + \kappa_3 0
$$

7.4.2 Standard errors from the joint model

Rizopoulos predicts that joint models will generally yield larger standard errors compared to traditional methods. In the current investigation, however, this loss of stated precision did not occur (see Table 10).

Table 10. Standard errors from separate and joint models.

	Cox*	LME/extended Cox†	Joint (Weibull)‡	Joint (piecewise)§	Joint (splines)‖
Longitudinal estimates					
Intercept		0.47	0.47	0.42	0.42
ns(age,2)1		0.28	0.28	0.22	0.24
ns(age,2)2		0.4	0.4	0.33	0.33
Body mass index		0.02	0.02	0.01	0.01
Race: African-American		0.23	0.23	0.21	0.21
Race: Other		0.3	0.3	0.27	0.27
Fair/Poor/Very Poor health		0.19	0.19	0.17	0.17
Event estimates					
Fair/Poor/Very Poor health	0.19	0.21	0.2	0.2	0.2
Gait speed	0.03	0.04	0.05	0.04	0.04
		0.04 *shape*	0.06 *shape*	κ_1 0.59	κ_1 2.32
				κ_2 0.47	κ_2 0.5
				κ_3 0.45	κ_3 0.45
				κ_4 0.43	
				κ_5 0.43	
				κ_6 0.42	

* Cox proportional-hazards model with baseline CD4 count only.
† Linear mixed-effects model (top portion) and Cox proportional-hazards model with time-varying covariate (bottom portion).
‡ Joint model with Weibull baseline hazard function (method = "weibull-PH-aGH").
§ Joint model with piecewise-constant model of baseline hazard function (method = "piecewise-PH-aGH").
‖ Joint model with piecewise-constant model of baseline hazard function (method = "splines-PH-aGH").

7.4.3 Joint model diagnostics

**Diagnostics for the longitudinal sub-model**

A plot of conditional residuals versus model-fitted values (Figure 26) shows some systematic departure from the expected mean 0 for all 3 joint models, indicating that the specification of the hierarchical model can be improved. However, we should note that these plots are more appropriate for a traditional mixed-effects model than for a joint model, since they do not show the event sub-model's accounting for a significant NMAR mechanism. The deviation from 0 on the left side of the plot may be attenuated if we multiply imputed the missing observations at the low end of gait speed, as described in section 4.3.4. Unfortunately, this method is not available in the `jointModel` package when visit times are random and data are left-censored (as is the case in our model, because we use age the time scale).

Figure 26. Diagnostic residuals plot for the longitudinal sub-model: Conditional residuals versus model-fitted values.
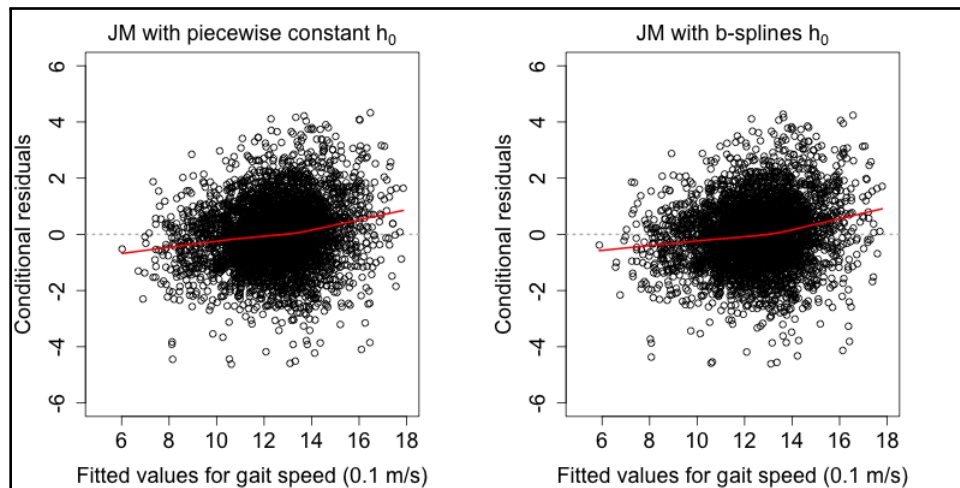
A plot of standardized conditional residuals versus theoretical normal quantiles (Figure 27) indicates slight deviation from normality in the tails, but the data within 2 standard deviations appear normal.

Figure 27. Diagnostic residuals plot for the longitudinal sub-model: Standardized conditional residuals versus theoretical normal quantiles.
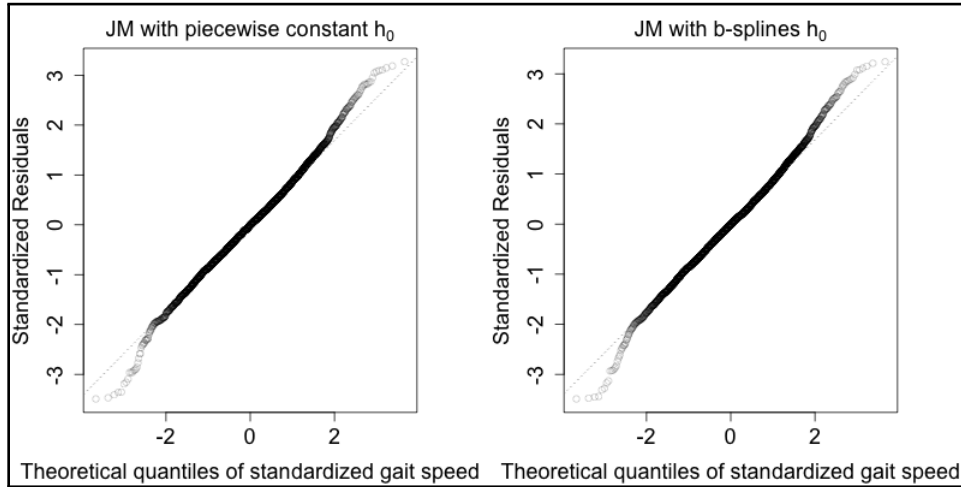
Diagnostics for the event sub-model

Figure 28. Martingale residuals for the models with piecewise constant and b-splines baseline hazards. At low values of walking speed, there is some systematic deviation from 0, indicating that there are fewer deaths in this range of walking speed than the model predicts. This may be an artifact of sparcity of data at these very slow walking speeds, or an indication that model fit can be improved. At moderate and fast walking speeds, the predictions are much closer to observed deaths. For reference, average human walking speed is 1.4 m/s (14 on the scale below).
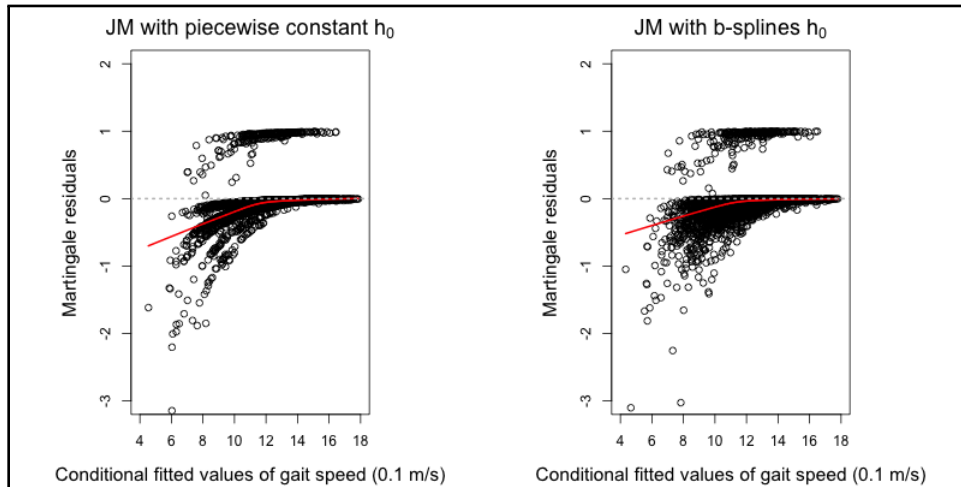
Figure 29. Survival Function of Cox-Snell Residuals, with a superimposed unit exponential curve (in gray). Ideally, the survival function will match the unit exponential curve. These plots indicate that the model with a b-splines baseline hazard is a better fit.
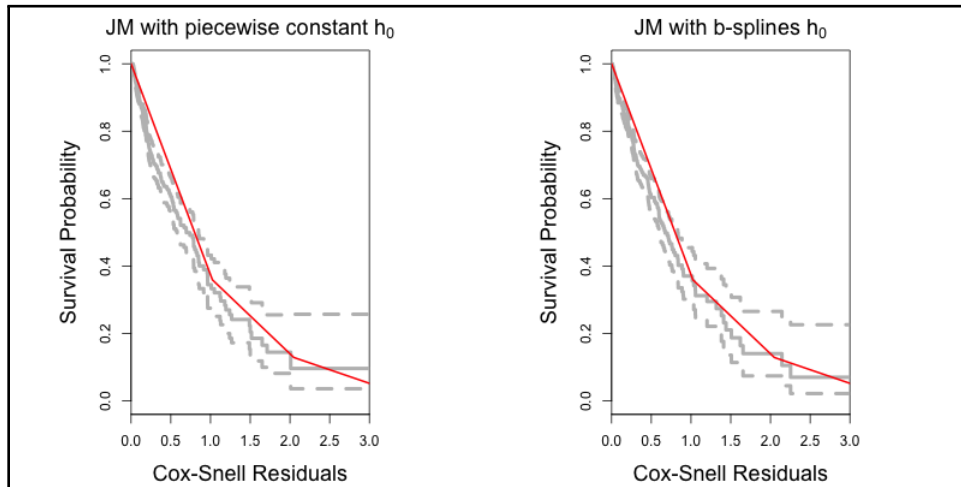

Figure 30. Cumulative hazard of the Cox-Snell residuals plotted against the Cox-Snell residuals themselves. If $\hat{H}_i(r_i^{cs}) \sim \exp(\lambda = 1)$, the plot will be a 45° line through the origin. These plots indicate a reasonable fit for all 3 joint models, with the model with a b-splines baseline hazard showing the best fit.