

**JOINT MODELLING OF THE RELATIONSHIP BETWEEN
SLEEP, DISEASE AND MORTALITY, EXCLUSIVELY
IN A COHORT OF OLDER AUSTRALIAN WOMEN
(AGED 70-75 YEARS AT BASELINE)**

**LUCY LEIGH¹, IRENE LENA HUDSON^{1,2}
and JULIE ELLEN BYLES¹**

¹Research Centre for Generational, Health and Ageing
Faculty of Health and Medicine
HMRI Building
The University of Newcastle
Callaghan, NSW, 2308
Australia

²School of Mathematical and Physical Sciences
Faculty of Science and Information Technology
The University of Newcastle
Callaghan, NSW, 2308
Australia
e-mail: irenelena.hudson@gmail.com

Abstract

This paper explores the relationship between sleep, disease, and survival, utilising shared random effects joint modelling to account for informative dropout, in a cohort of very old women. Joint modelling simultaneously models the longitudinal sleep trajectory and its effect on survival. A series of nested

2010 Mathematics Subject Classification: 62-XX.

Keywords and phrases: ALSWH, joint models, longitudinal sleep trajectory, shared parameter models, sleep difficulty, survival.

Received October 29, 2016

joint models are implemented to separate out the effects on survival of sleep and disease, the latter of which is known to have a causal role in mortality, and other factors of interest, such as self-rated health, physical functioning, mental health, vitality, baseline age, BMI, area of residence, marital status, and education. Dynamic survival predictions are also created and compared for a small sample of subjects.

1. Introduction

Sleep quality deteriorates with age, due to both normal age-related sleep changes and as a result of ill health [1-3], medications, and other factors. The incidence of disturbed sleep is strongly increased in older people [4, 5], with some studies reporting as much as 42% of older adults experiencing difficulty initiating and maintaining sleep [1, 6-9]. While there have been a number of studies investigating the relationship between sleep, mortality and disease, few have examined sleep in very old populations [10-12], particularly those living in the community, and it may be inappropriate to extrapolate results from sleep research on adults aged less than 75 years to those aged 75 and over, given that sleep behaviour changes with age. For instance, some studies have reported a protective effect of sleep apnoea in adults aged over 65 [13].

The current evidence of a relationship between sleep duration and mortality is conflicting. Some systematic reviews [14, 15] have found that a U-shaped relationship exists, with both long and short sleep duration (> 8-9 hours a night, and < 7 hours a night, respectively) associated with increased risk of mortality. However, a more recent systematic review [16] found that the literature does not support a U-shaped relationship. Furthermore, a study of adults aged 32-86 years found that the U-shaped relationship was present only for the elderly (aged 60-86 years) and that it was unlikely to contribute towards mortality, but rather is a consequence of medical conditions, such as diabetes, body weight, hypertension, general health and cancer, and age related sleep changes [17].

The strong associations between socio-demographic factors, health related factors, and sleep duration, may explain the reported associations between sleep duration and mortality [16]. Older people are more likely

to report long or short sleep duration, and unmarried people are more likely to report short, or long *and* short sleep duration [16]. Insufficient sleep and sleep disorders have been associated with a number of chronic diseases [18-22], such as diabetes [23, 24], hypertension [25, 26], cardiovascular disease [27, 28], stroke [29], arthritis [30], depression [30, 31], cognitive function decline [32-34], and obesity [33, 35]. There is a strong bidirectional relationship between sleep disorders and serious medical problems in adults; many diseases being more likely to develop in individuals with sleep disorders, and individuals with sleep disorders at greater risk of developing these diseases [36]. Nearly all of the aforementioned socio-demographic and health related factors are also correlated with mortality, and so the reported associations between sleep and mortality need to be interpreted with care [16]. Socio-demographic factors are likely to lie upstream of sleep duration, whereas health factors may be either upstream or downstream of sleep duration on the causal pathway [16]. ‘Ambiguities about where on the causal pathway the health conditions lie make it impossible, given the present state of knowledge, to distinguish confounders and mediators’ [16].

Possible mechanisms by which sleep difficulty may lead to increased risk of mortality include adverse effects on endocrinologic, immunologic or metabolic effects [16, 37], induction of chronic, low-grade inflammation [16, 38], increases in cortisol secretion and growth hormone metabolism [16, 39, 40], or changes in immune function, depression, or underlying disease processes [16, 41]. Kurina et al. [16] also state that reverse causation, in which subjects ‘with health conditions putting them at increased risk of mortality have either lengthened or shortened sleep’ is an alternative explanation for the observed associations between sleep duration and mortality.

Lack of sleep, or poor sleep quality, also has societal impacts for older adults. Poor sleep can have detrimental consequences on many aspects of vitality and resilience required for successful ageing [42]. Sleep disturbance is associated with increased risk of nursing home placement, with as high as 50% of admissions in long-term care directly attributable to elderly sleep disturbances [42]. Sleep difficulty also exacts a toll on

family support by disrupting caregivers' sleep [42]. The risk of falls is increased in older adults, and this risk should be taken into account when prescribing sleep medications, which also increase the risk of falls [43, 44]. Sleep disturbances have been shown to be associated also with greater frailty in the elderly [45]. Inability to sleep can also lead to delayed response time, which is particularly important as it can affect driving ability as well as increase the risk of falls [46]. Finally, insomnia has been linked to increased consumption of healthcare resources [46]. In America, hundreds of billions of dollars are spent each year on doctor visits, hospital services, prescriptions, and over-the-counter-medications, associated with chronic disorders of sleep and wakefulness [47]. A recent report found that the economic impact of sleep disorders costs Australia \$5.1 billion per year, spread over health care costs, costs associated with medical conditions attributable to sleep disorders, and non-medical costs resulting from sleep loss related accidents [48]. Early detection and intervention for sleep disturbance in older adults can help reduce the financial burden associated with sleep related accidents, depression, and illness, and promote better quality lives [49].

This study investigates the relationship between sleep, disease and mortality, exclusively in a cohort of older women (aged 70-75 years at baseline) [50, 51], using data from the 1921-1926 cohort of the Australian Longitudinal Study on Women's Health (ALSWH; www.alswh.org.au). Rather than posit a causative relationship, we aim to identify whether an association between sleep and mortality exists. Many studies that have investigated sleep and mortality focus on sleep duration. We aim to explore whether a similar relationship holds for sleep difficulty, measured longitudinally using a simple 5 item questionnaire developed by Hunt et al. [52]. Sterniczuk et al. [12] reported that 'sleep disturbance might have prognostic utility in predicting future increases in frailty and decline in health' [12]. A short sleep difficulty assessment, such as the one utilised in this paper, has the benefit of reducing responder burden, which would be beneficial both to the clinical staff in busy hospital or nursing home settings, as well as to very old or frail respondents.

Recent analysis [53] of this ALSWH cohort identified unobserved groups (latent classes) of women who were similar in their sleep patterns by Latent Class Analysis (LCA) [54], and related these classes to mortality via distal survival regression techniques [55-57]. Leigh et al. [53] hypothesized that increased sleep difficulty would be associated with mortality. Four latent classes of persistent sleep difficulty were identified, corresponding to women who reported no sleep difficulty ('untroubled'), those who reported great sleep difficulty ('troubled' sleepers), those who reported taking a long time to fall asleep and those who woke in the early hours of the morning. More than 65% of women were classified into a latent class that was characterised by some level of sleep difficulty. These classes were significant indicators for survival, with worse survival for the troubled sleepers compared to untroubled sleepers, and better outcomes for the early wakers and troubled falling asleep classes. This relationship may largely be due to the effects of disease and other factors in survival, which correlate with poor sleep. A differential effect of morbidity across sleep classes was also found. The main effect of disease on mortality was that increased number of diseases was associated with greater hazard of death. However, once an interaction with sleep was included in the modelling of Leigh et al. [53] only having 3 or more diseases was associated with a greater hazard of death, and within the troubled sleepers class the effect of disease on mortality was nonsignificant. Sleep difficulty class was also shown to be associated with a number of specific diseases. Leigh et al. [53] established that having any of arthritis, heart disease, asthma, bronchitis/emphysema, diabetes, hypertension, or osteoporosis, was associated with greater odds of belonging to their troubled sleeping class, early wakers, or trouble falling asleep class, when compared to the untroubled sleepers. Stroke and cancer were not significantly associated with sleep difficulty. Supplementary Material 2 gives the details of these results using a series of multinomial logistic regressions of the latent sleep classes on disease.

One drawback of the previous analysis of Leigh et al. [53] was that the LCA relied on a Missing at Random (MAR) assumption [54], which assumed that dropout was uninformative. Given the advanced age of the cohort, there exists however the possibility that dropout is informative, in which case a Not Missing at Random (NMAR) [58] assumption should be utilised. The use of a MAR assumption instead of a NMAR assumption in the previous study may account for the lack of a significant effect of disease on mortality for women who report the greatest sleep difficulty. Furthermore, the protective effect of the two mild sleep difficulty classes (early wakers or trouble falling asleep) could be biased by NMAR dropout; in that these classes may consist of survivors, women who therefore have greater opportunity to report some sleep difficulty given they live longer.

Therefore in the current study, we investigate simultaneously the evolution of the longitudinal trajectory of the sleep outcome and time-to-death process, as well as the distribution of time to-death conditional on intermediate longitudinal measurements by following the Joint Modelling (JM) with shared-parameter models developed by Rizopoulos [59]. We aim to test for any association between longitudinal sleep trajectory and mortality in a cohort of very old women, using shared random effects joint modelling to account for informative dropout. We also aim to test for a possible interaction between disease and sleep on survival. We hypothesise that greater sleeping difficulty will be associated with worse survival outcomes, and that the effect may be different across the number of reported diseases.

Joint modelling has been utilised widely in the medical literature, including applications to heart failure studies [60, 61], HIV/AIDS research [62], mental health [63, 64], liver transplant studies [65, 66], stem cell transplant [67], cystic fibrosis [68], and in clinical trials [69, 70], however to the authors' knowledge JM has not been used to date to investigate sleep and mortality.

2. Data

Data analysed were part of the ALSWH; a nationally-representative, prospective study of over 40,000 participants, that commenced in 1996. Cohorts of women, born in 1973-78, 1946-51, and 1921-26 were sampled from the Medicare Australia database [71] and invited to complete the baseline postal survey. Since then women have been re-surveyed on a three-yearly basis. Further details on the establishment of the cohorts, and follow-up, have been published elsewhere [72, 73]. This study presents data collected on the 1921-1926 cohort who completed the baseline survey in 1996 (Survey 1) when aged 70-75 years, and who have since completed sleep questionnaires for at least one of Survey 2 (1999, 73-78 years), Survey 3 (2002, 76-81 years), and Survey 6 (2011, 85-90 years) ($n = 10,606$).

The longitudinal sleep outcome was the Nottingham Health Profile (NHP) sleep subscale score, a composite measure of sleep quality. The NHP was measured at three time points (surveys 2, 3 and 6). The sleep subscale consists of five yes/no items which are combined to produce a single score; Q1. Do you wake in the early hours of the morning? ; Q2. Do you lie awake most of the night?; Q3. Do you take a long time to get to sleep?; Q4. Do you sleep badly at night?; Q5. Do you take sleep medication? The following weighted sum, as outlined in the NHP instrument scoring instructions [74], is used to create the single sleep subscale score: $12.57 * Q1 + 27.26 * Q2 + 16.10 * Q3 + 21.70 * Q4 + 22.37 * Q5$, giving a total score out of 100, with larger values indicating greater sleep difficulty. Details of the validity of the NHP sleep subscale were reported previously [53], and are included in Supplementary Material 1.

A number of covariates were also measured at each survey. Body Mass Index (BMI) was calculated from self-reported weight and height, classified according to the World Health Organisation categories [75]. QoL measures of mental health, vitality, and physical function were available at each survey using the Short Form (36) Health Survey (SF-36) subscales [76]. The SF-36 provides subscale scores out of 100, with higher scores indicating better mental health, vitality and physical function. At

each survey, women were also questioned about diagnosed medical conditions including diabetes, arthritis, heart disease, hypertension, asthma, bronchitis/emphysema, stroke, osteoporosis, and cancer. These were combined into a single disease count variable, categorised into 0 diseases, 1-2 diseases, and 3 or more diseases. Area of residence was determined from geocoded address data and categorised using the Accessibility/Remoteness Index of Australia (ARIA +) [77]. Marital status was recorded at each survey, and women were categorised as either married/de facto, separated/divorced, widows, or never married. Highest level of education was ascertained at baseline only, with women categorised as having no education, a school or higher school certificate, having a trade or diploma, or higher education. Self-rated health was also queried at each survey, categorized into good/very good/excellent or fair/poor. Baseline age was included in the analytic modelling. Deaths were ascertained from the National Death Index [78]. Rates of missingness on each of the covariates ranged from less than 1%, to a maximum of around 20% for BMI. A basic cohort descriptive is given in Supplementary Material 2.

3. Methodology

3.1. Joint modelling

Joint modelling simultaneously models the longitudinal sleep outcome and the survival outcome; implemented using the R package ‘JM’ [79]. There are various types of joint models in the literature, which generally fall under the banner of either selection models, pattern mixture models, or random-effects models. The difference between these is shown in Equations (1)-(3) [80]:

$$\text{Selection models: } [W, Y, b] = [b][W|b][Y|W], \quad (1)$$

$$\text{Pattern mixture models: } [W, Y, b] = [b][Y|b][W|Y], \quad (2)$$

$$\text{Random effects models: } [W, Y, b] = [b][W|b][Y|b], \quad (3)$$

where W denotes the vector of longitudinal repeated measures, Y the event times, and b the latent random effects that link the longitudinal and survival processes, and $[X, Y, b]$ is the joint distribution [80].

The JM used in this research was the random effects model, due to its popularity in health sciences [60-66]. For simplicity, the random effects model will be referred to as JM for the remainder of the paper. The JM consists of two sub-models, the survival and longitudinal submodels, and shared random effects are assumed to account for the association between the longitudinal and survival outcome [59, 80, 81]. For this reason, random effects models are also known as shared-parameter models.

The formulation of the JM, as used in the R package ‘JM’ and described by Rizopoulos [82], is as follows.

For subjects $i = 1, \dots, n$, let Y_i denote the event time and C_i the censoring time, let $Z_i(y)$ be the q -dimensional vector of exogenous (possibly time-varying) covariates, and $X_i(y)$ the longitudinal process at time $y \geq 0$. For each subject, we observe an event time or censoring time, i.e., $V_i = \min(Y_i, C_i)$. d_i is a dichotomous indicator variable; equal to 1 if the event occurs and 0 otherwise. The covariate $X_i(y)$ is observed only intermittently, at times $y_{ij} \leq V_i$, $j = 1, \dots, m_i$, for subject i . Additionally, due to error, we may not directly observe $X_i(y)$, and instead observe only $w_i = \{w_i(y_{i1}), \dots, w_i(y_{im_i})\}^T$, which follows the observed longitudinal process $w_i(y)$.

The joint model is comprised of two sub-models, one for the ‘true’ longitudinal process $X_i(y)$, and the other for the event time. This model is presented graphically in Figure 1, where Y denotes the time-to event process, W the observed longitudinal process, with coefficients γ and β , respectively; b represents the random effects, which links the two sub-models, and accounts for all association between the longitudinal and survival processes [80].

An advantage of JM is that it allows the longitudinal outcome to be included as an endogenous predictor of the survival outcome. An endogenous (or ‘internal’) covariate is a ‘time measurement taken on the individual’, and requires the survival of the individual for its existence, whereas an exogenous (or external) covariate is not directly involved with the failure mechanism [83]. The observed trajectory of an endogenous covariate for a subject also provides information on the survival or failure time of that subject. Examples of endogenous covariates include biomarkers, clinical parameters or questionnaire data; the measurement of which depends on the subject being alive. Conversely, covariates that are measured in advance and remained fixed throughout the study (for example, seasons or time of day) [83], and covariates that are pre-defined, such as a treatment strategy where dose is adjusted according to predetermined criteria [82], are exogeneous.

Endogeneous covariates are typically measured with error, due to variation induced by the subject [81, 82]. Their complete path until any time y is not fully observed [82] – rather, it is only known at specific occasions, when the individual presented to the study to provide measurements for follow-up, or returned survey data [82].

Finally, because the occurrence of the survival event effectively censors the longitudinal process, this may induce informative censoring, a type of NMAR missingness [84, 85]. Traditional survival analysis models assume that time-varying covariates are predictable processes, in the sense that the value of the covariate is known at an instant just prior to time y (i.e., a time-continuous process), and are also measured without error, and have fully specified paths [81, 82, 86]. The construction of the likelihoods for these models assumes that the behaviour of the covariate is piecewise-constant, i.e., that the covariate value changes at the observation time, and remains constant in between observation times [82]. This assumes that the value of the covariate is known for any time at which an individual is under observation [86]. The assumption of

piecewise-constant hazards is unrealistic for many covariates, such as biomarkers [82]. Traditional survival analytic approaches, therefore, are only appropriate for exogeneous time-varying covariates [82]. JM overcomes these drawbacks by modelling the longitudinal outcome as an endogeneous predictor.

3.2. Survival sub-model

A proportional hazards model is generally postulated to estimate the strength of the association between $x_i(y)$ and the risk of the event [59, 82, 84],

$$\begin{aligned}
 h_i(y | X_i(y), Z_i(y)) &= \lim_{dy \rightarrow 0} \Pr\{y \leq Y_i \leq y + dy | Y_i \geq y, X_i(y), Z_i(y)\} / dy \\
 &= h_0(y) \exp\{\gamma^T Z_i(y) + \beta x_i(y)\}, y > 0.
 \end{aligned}
 \tag{4}$$

$X_i(y) = \{x_i(s), 0 \leq s \leq y\}$ denotes the history of the *true* (unobserved) longitudinal process up to time point y , and $h_0(y)$ the baseline risk [82, 84]. β quantifies the effect of the underlying longitudinal outcome to the risk for an event [82]. $Z_i(y)$ may contain either baseline or timevarying exogeneous covariates, with effect size γ .

The method of B-splines was chosen in this study, given the need to include time-varying predictors in the survival model (in addition to those modelled by the longitudinal sub-model). Currently this is the only option available for time-varying covariates in the R package ‘JM’. The description of this baseline hazard formulation, as set out by Rizopoulos [82], is as follows. The log of the baseline hazard function is expanded into B-spline basis functions for cubic splines:

$$\log h_0(y) = \kappa_0 + \sum_{d=1}^m \kappa_d B_d(y, q),
 \tag{5}$$

where $\kappa^T = (\kappa_0, \kappa_1, \dots, \kappa_m)$ are the spline coefficients, q is the degree of the B-splines basis functions $B(\cdot)$, and $m = \ddot{m} + q - 1$, where \ddot{m} is the number of interior knots [82].

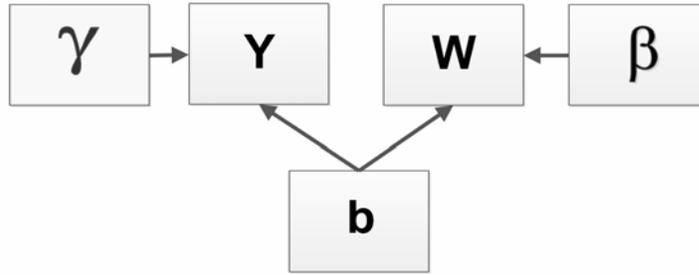


Figure 1. Shared random effects joint model (sourced from McCrink et al. [80]).

3.3. Longitudinal sub-model

Recall that we denote the true value of the underlying longitudinal process at time y as $x_i(y)$, but the process is actually observed intermittently, possibly with error, at set times y_{ij} , denoted by $w_i(y_{ij})$. This measurement error is accounted for using the first line of Equation (6). We then estimate $x_i(y)$ in order to construct the complete longitudinal history $X_i(y)$ for each subject [82]. This can be achieved in a number of ways. Rizopoulos describes a linear mixed-effects model with normally distributed longitudinal outcomes as follows [82]:

$$\begin{cases} w_i(y) = x_i(y) + \varepsilon_i(y), \\ x_i(y) = m_i^T(y)\alpha + n_i^T(y)b_i, \\ b_i \sim \mathcal{N}(0, D), \varepsilon_i \sim \mathcal{N}(0, \sigma^2). \end{cases} \quad (6)$$

The design vectors $m_i(y)$ for the fixed effects α and $n_i(y)$ for the random effects b_i , as well as the error terms, are all time-dependent [82]. The error terms are also assumed mutually independent, normally distributed with zero mean and variance σ^2 , and independent of the random effects [82]. The first line of Equation (6) accounts for measurement error in the observed longitudinal response. Together with

the random effects, the use of the time structure in the definition of $m_i(y)$ and $n_i(y)$, reconstructs the complete path of the longitudinal process $X_i(y)$ [82].

3.4. Maximum likelihood estimation

Maximum likelihood estimates are derived as the modes of the log-likelihood function corresponding to the joint distribution of the observed outcomes $\{Y_i, d_i, w_i\}$ [82]. It is assumed that a vector of time-independent random effects b_i underlies both the longitudinal and survival processes [82]. These random effects account for the association between the longitudinal and survival outcomes, as well as the correlation between the repeated measures [59, 82]. This corresponds to an assumption of conditional independence, that is,

$$p(Y_i, d_i, w_i | b_i; \theta) = p(Y_i, d_i | b_i; \theta)p(w_i | b_i; \theta), \quad (7)$$

$$P(w_i | b_i; \theta) = \prod_j p(w_i(y_{ij}) | b_i; \theta). \quad (8)$$

The $\theta = (\theta_y^T, \theta_w^T, \theta_b^T)$ denotes the full set of parameters, with θ_y indicating the parameters for the survival outcome, θ_w the parameters for the longitudinal outcome, and θ_b the parameters of the random-effects covariance matrix, with w_i the $n_i \times 1$ vector of longitudinal response for subject i [59, 82]. Also assumed is that, given the observed history, the censoring mechanism and observation times are independent of the event times and longitudinal measurements [82]. That is, the censoring mechanism (whether a subject withdraws from a study), and the observation of the longitudinal process (i.e., whether a subject turns up for an appointment or returns a survey), depends only on their observed past history of longitudinal and baseline covariates [82]. The log-likelihood contribution for the i -th subject is then [59, 80, 82]:

$$\begin{aligned}
& \log p(Y_i, d_i, w_i; \theta) \\
&= \log \int p(Y_i, d_i, w_i; b_i; \theta) db_i \\
&= \log \int p(Y_i, d_i | b_i; \theta_y, \alpha) \left[\prod_j p\{w_i(y_{ij}) | b_i; \theta_w\} \right] p(b_i, \theta_b) db_i. \quad (9)
\end{aligned}$$

The conditional density for the survival part of the JM is written as [59]

$$p(Y_i, d_i | b_i; \theta_y, \alpha) = h_i(Y_i | X_i(Y_i); \theta_y, \alpha)^{d_i} S_i(Y_i | X_i(Y_i); \theta_y, \alpha), \quad (10)$$

where $h_i(\cdot)$ is defined in Equation (4), and

$$\begin{aligned}
S_i(Y | X_i(Y), z_i; \theta_y, \alpha) &= \Pr(Y_i^* > y | X_i(Y), z_i; \theta_y, \alpha) \\
&= \exp\left\{-\int_0^y h_i(s | X_i(s); \theta_y, \alpha) ds\right\}, \quad (11)
\end{aligned}$$

where $p\{w_i(y_{ij}) | b_i; \theta_w\}$ is the univariate normal density for the longitudinal responses, and $p(b_i, \theta_b)$ is the multivariate normal density for the random effects [59]. $h_0(\cdot)$ from Equation (4) is any positive function of time (e.g., piecewise constant or B-splines, or a known parametric distribution) [82].

The joint density for the longitudinal responses, with the random effects, takes the form,

$$\begin{aligned}
p(w_i | b_i; \theta) p(b_i; \theta) &= \prod_j p\{w_i(y_{ij}) | b_i; \theta_w\} p(b_i, \theta_b) \\
&= (2\pi\sigma^2)^{-n_1/2} \exp\{-\|w_i - M_i\beta - N_i b_i\|^2 / 2\sigma^2\} \\
&\quad \times (2\pi)^{-q_b/2} \det(D)^{-1/2} \exp(-b_i^T D^{-1} b_i / 2), \quad (12)
\end{aligned}$$

where q_b denotes the dimensionality of the random-effects vector, and $\|m\| = \{\sum_i m_i^2\}^{1/2}$ the Euclidean vector norm [59, 82]. Due to the intractable nature of the survival and random effects integrals in the

joint likelihood, numerical integration using the Gauss-Hermite quadrature rule [87] is utilised to approximate the above integrals [59, 80]. Details of an EM algorithm which maximises the log likelihood can be found in Rizopoulos [82].

3.5. Interaction effects

An interaction between the longitudinal outcome and other exogeneous covariates can also be accommodated in the JM. Such a model assumes that the effect of the longitudinal outcome behaves differently across different subgroups of the subjects [82]. The survival model is expanded to

$$h_i(y) = h_0(y) \exp\{\gamma^T Z_{i1}(y) + \beta[Z_{i2}(y) \times x_i(y)]\}, \quad (13)$$

where Z_{i1} contains direct effects on the event risk, and Z_{i2} denote the interaction terms that expand the association of $x_i(y)$ in different subgroups [82].

3.6. Dynamic survival predictions

Survival predictions for specific subjects at different follow-up times can also be estimated [82, 88, 89]. The dynamic survival model, as formulated by Rizopoulos [87], is described below. A joint model is fitted based on a random sample $D_n = \{Y_i, d_i, w_i; i = 1, \dots, n\}$, with the aim of predicting survival probabilities for a new subject i with longitudinal measurements $W_i(y) = \{w_i(s); 0 \leq s < y\}$, and covariates z_i . The observed longitudinal covariate $w_i(y)$ is endogeneous, and as such contains information directly related to the failure mechanism – i.e., a subject, providing longitudinal measurements up to time y , implies survival up to this point [82]. Therefore, JM can predict the conditional probability of surviving time $u > y$, given a subject has already survived up to time y , $\pi_i(u | y)$ [82]. That is,

$$\pi_i(u | y) = \Pr(Y_i^* \geq u | Y_i^* > y, W_i(y)z_i, D_n; \theta^*), \quad y > 0, \quad (14)$$

where θ^* denotes the true parameter values [82, 89]. An important characteristic of such predictions from the JM is that they are dynamic in nature, in that as time progresses and additional information is recorded for a subject, these predictions can be updated to incorporate such new information.

Estimation of the subject-specific, dynamic, conditional survival probabilities utilises the conditional independence assumptions outlined in Subsection 3.4 above. Following the formulation by Rizopoulos [82], in which conditioning on covariates z_i is assumed but omitted from notation, Equation (14) can be rewritten as follows:

$$\begin{aligned} \Pr(Y_i^* \geq u | Y_i^* > y, W_i(y); \theta) &= \int \Pr(Y_i^* \geq u | Y_i^* > y, W_i(y), b_i; \theta) \\ &\quad \times p(b_i | Y_i^* > y, W_i(y); \theta) db_i \\ &= \int \Pr(Y_i^* \geq u | Y_i^* > y, b_i; \theta) p(b_i | Y_i^* > y, W_i(y); \theta) db_i \\ &= \int \frac{S_i\{u | X_i(u, b_i, \theta)\}}{S_i\{y | X_i(y, b_i, \theta)\}} p(b_i | Y_i^* > y, W_i(y); \theta) db_i, \quad (15) \end{aligned}$$

where $S_i(\cdot)$ is the survival function defined earlier, and the longitudinal history $X_i(\cdot)$ is a function of both the random effects and the parameters [82]. Further details regarding the derivation of empirical Bayes estimates for $\pi_i(u | y)$ and their estimation using Monte Carlo simulations can be found in Proust-Lima and Taylor [88], Rizopoulos [90] and Rizopoulos [82].

3.7. Missing covariates

Since JM requires complete data on the longitudinal covariates, we used multiple imputation (MI) [91] to handle the missing data on the covariates. Missing data rates for the cohort are given in Supplementary Material 2. Due to the variety of covariate types (continuous,

dichotomous, and unordered categorical), multiple imputation using chained equations ('mice') was utilized [91]. Graham et al. [92] recommend 20 imputed data sets for imputing data with 10-30% missing information. Bodner [93] and White et al. [94] also recommend that the number of imputations should be similar to the percentage of cases that are incomplete.

3.8. Statistical analyses

A set of four nested joint models were created to investigate the impact of sleep, disease and other covariates on survival. Time was centred at survey 2 ($y = 0$ years), with survey 3 at $y = 3$ years, and survey 6 at $y = 12$ years. Survival data was censored at the end of 2013 ($y = 14.77$ years).

In the first model, only time was included as a predictor of the longitudinal sleep outcome (an 'empty' model), and the longitudinal sleep trajectory was the sole predictor for survival. In Model 2, categorical disease count was included as a predictor in both the longitudinal and survival parts of the JM. In Model 3, the survival part of Model 2 was extended to include an interaction between sleep and disease; the longitudinal part of the model remained unchanged. Finally, Model 4 extends Model 3 by including all other covariates of interest (BMI, self-rated health, physical functioning, mental health, vitality, marital status, area of residence, baseline age, and education) as time-varying predictors in both the longitudinal and survival parts of the JM (excluding age and education, which were included as baseline predictors).

MI was not performed for Model 1, since no additional covariates with missing data were included in this model. In Models 2 and 3, 10 imputed data sets were analysed and combined, since the amount of missing on the disease covariate was around 3% across all surveys. Finally, 20 imputed data sets were used to impute all missing covariates for Model 4, due to the higher levels of missing on BMI.

4. Results

4.1. The longitudinal sub-models

To determine the suitability of a linear mixed model to model change in the NHP sleep subscale scores over time, the intra class correlation coefficient (ICC) [80, 95] was first estimated by building a null mixed model using the R package ‘nlme’ [96]. The ICC is calculated as the ratio of the subject random effect variance over the total variance, and allows assessment of whether or not a random effect is present in the data [95]. Approximately 53% of the variation in NHP sleep subscale scores was found to be due to the within-individual correlation, and as such was suitable for random effects modelling.

Table 1 displays the results for the longitudinal parts of Models 1, 2, and 4. Model 3 is not reported, since the longitudinal results for Models 2 and 3 were almost identical (to the second decimal place), not surprising given these models differ only in their survival sub-model. Greater sleep difficulty was associated with time, with 0.54 unit increases (95%CI (0.48, 0.60), $p < 0.01$) and 0.35 unit (95%CI (0.28, 0.41), $p < 0.01$) increases in the NHP per year estimated from Models 1 and 2, respectively. However, after adjusting for all other covariates this effect was non-significant (Model 4; Table 1).

Table 1. Point estimates and 95% CIs for the longitudinal sub-model results (bold values are significant at alpha = 0.05)

	Model 1	Model 2	Model 4
Intercept	29.11 (9.06, 49.16) <i>p</i> = 0.0044	26.17 (6.52, 45.83) <i>p</i> < 0.01	76.91 (58.43, 95.39) <i>p</i> < 0.01
Survey (years)	0.54 (0.48, 0.60) <i>p</i> < 0.01	0.35 (0.28, 0.41) <i>p</i> < 0.01	0.06 (− 0.01, 0.128) <i>p</i> = 0.09
Baseline age	− 0.12 (− 0.40, 0.16) <i>p</i> = 0.39	− 0.13 (− 0.40, 0.14) <i>p</i> = 0.21	− 0.39 (− 0.64, − 0.13) <i>p</i> < 0.01
0 disease		Ref	Ref
1-2 disease		3.36 (2.65, 4.07) <i>p</i> < 0.01	1.58 (0.88, 2.28) <i>p</i> < 0.01
3+ disease		8.93 (8.00, 9.86) <i>p</i> < 0.01	4.78 (3.84, 5.71) <i>p</i> < 0.01
SF36:Physical Functioning			− 0.03 (− 0.05, − 0.17) <i>p</i> < 0.01
SF36:Mental Health			− 0.23 (− 0.26, − 0.21) <i>p</i> < 0.01
SF36:Vitality			− 0.14 (− 0.16, − 0.12) <i>p</i> < 0.01
BMI: Underweight			− 1.24 (− 2.76, 0.27) <i>p</i> = 0.11
BMI: Normal weight			Ref
BMI: Overweight			0.19 (− 0.52, 0.92) <i>p</i> = 0.99
BMI: Obese			− 0.47 (− 1.55, 0.62) <i>p</i> = 0.4
SRH: Fair/Poor			Ref
SRH: Good/Very Good/Excellent			− 2.05 (− 2.78, − 1.32) <i>p</i> < 0.01
MS: Married/De facto			Ref
MS: Separated/Divorced			1.68 (0.09, 3.28) <i>p</i> = 0.04
MS: Widow			2.20 (1.52, 2.88) <i>p</i> < 0.01

Table 1. (Continued)

MS: Never married	- 1.10 (- 3.36, 1.17) <i>p</i> = 0.34
ARIA: Major City	Ref
ARIA: Inner Regional	- 0.20 (- 0.96, 0.56) <i>p</i> = 0.61
ARIA: Outer Regional	- 0.56 (- 1.56, 0.44) <i>p</i> = 0.27
ARIA: Remote/Very Remote	- 1.37 (- 3.82, 1.09) <i>p</i> = 0.28
EDU: None	Ref
EDU: School/High school certificate	- 1.01 (- 1.86, - 0.17) <i>p</i> = 0.02
EDU: Trade/Diploma	- 1.53 (- 2.81, - 0.25) <i>p</i> = 0.02
EDU: Higher education	- 0.81 (- 2.82, 1.21) <i>p</i> = 0.43

Disease count was strongly associated with longitudinal sleep difficulty in both Models 2 and 4. In Model 2, having 1-2 diseases was associated with a 3.36 unit increase (95%CI (2.65, 4.07), $p < 0.01$) in NHP sleep subscale (increased sleep difficulty), compared to subjects with zero diseases. The effect of disease was larger for subjects reporting 3 or more diseases, associated with an 8.93 unit increase in NHP sleep subscale (95%CI (8.0, 9.86), $p < 0.01$), compared to women with 0 diseases. After adjusting for all other predictors, the association between disease and sleep was still significant, but the effect sizes were reduced to 1.58 (95%CI (0.88, 2.28), $p < 0.01$) and 4.78 (95%CI (3.83, 5.71), $p < 0.01$) for 1-2 diseases and 3 or more diseases, respectively, compared to 0 diseases.

For every year increase in age at baseline, NHP Sleep Score decreased by - 0.39 points (95% CI (- 0.64, - 0.13), $p < 0.01$). BMI and area of residence were not significant predictors of sleep difficulty. Women who were separated/divorced or widowed scored 1.68 (95%CI (0.09, 3.28), $p = 0.04$) units and 2.20 (95%CI (1.52, 2.88), $p < 0.01$) units

higher respectively on the NHP sleep subscale, compared to married/de facto women. The effect was not significant for women who had never married. Education and self-rated health were both significantly associated with improved sleep. Women who reported their health as good/very good/excellent scored -2.05 (95%CI $(-2.78, -1.32)$, $p < 0.01$) lower than women who rated their health as fair/poor. Women whose highest level of education was SC/HSC or trade/diploma also scored better on the NHP (-1.01 (95%CI $(-1.86, -0.17)$, $p = 0.02$); -1.53 (95%CI $(-2.81, -0.25)$, $p = 0.02$) respectively), but the effect was not significant for women with higher education. Improved physical functioning, mental health and vitality were all associated with improved sleep, corresponding to -0.03 (95%CI $(-0.05, -0.017)$, $p < 0.01$), -0.23 (95%CI $(-0.26, -0.21)$, $p < 0.01$) and -0.14 (95%CI $(-0.16, -0.12)$, $p < 0.01$) unit decreases in sleep difficulty respectively.

Trajectory plots comparing the longitudinal sleep trajectories across the latent sleep classes of earlier work [53] are included in Supplementary Material 2.

4.3. The Survival sub-model

From 10,606 women, there were 5186 deaths. Table 2 displays the results for the survival parts of Models 1-4. Model 1 is the effect of sleep on survival, unadjusted for any other covariates. A unit increase in the NHP sleep subscale (i.e., increasing sleep difficulty) was associated with a slightly higher hazard of death (HR = 1.005 (95%CI (1.003, 1.006), $p < 0.01$). However, as more covariates were added to the model, the size of this effect decreased. In Model 4, which adjusts for all covariates and an interaction between sleep and diseases, the effect of sleep on survival was non-significant.

Table 2. Hazard ratios and 95% CIs for survival sub-model results (bold values are significant at alpha = 0.05)

	Model 1	Model 2	Model 3	Model 4
NHP Sleep Subscale	1.005 (1.003,1.006) <i>p</i> < 0.01	1.003 (1.002,1.005) <i>p</i> < 0.01	1.006 (1.002,1.011) <i>p</i> < 0.01	0.999 (0.994, 1.003) <i>p</i> = 0.99
Baseline age	1.14 (1.12, 1.16) <i>p</i> < 0.01	1.10 (1.08, 1.12) <i>p</i> < 0.01	1.10 (1.08, 1.12) <i>p</i> < 0.01	1.12 (1.10, 1.15) <i>p</i> < 0.01
0 diseases		Ref	Ref	Ref
1-2 diseases		0.95 (0.88,1.02) <i>p</i> = 0.18	1.0 (0.88,1.14) <i>p</i> = 0.99	0.86 (0.76, 0.97) <i>p</i> = 0.01
3+ diseases		1.27 (1.17,1.39) <i>p</i> < 0.01	1.42 (1.22,1.65) <i>p</i> < 0.01	0.91 (0.79, 1.06) <i>p</i> = 0.22
NHP*0 diseases			Ref	Ref
NHP*1-2 diseases			0.998 (0.992,1.003) <i>p</i> = 0.35	0.998 (0.993, 1.003) <i>p</i> = 0.74
NHP*3+ diseases			0.996 (0.99,1.001) <i>p</i> = 0.12	0.999 (0.993, 1.004) <i>p</i> = 0.74
SF36:Physical functioning				0.991 (0.99, 0.993) <i>p</i> < 0.01
SF36:Mental health				0.997 (0.995, 0.999) <i>p</i> < 0.01
SF36:Vitality				0.999 (0.997, 1.001) <i>p</i> = 0.32
SRH: Fair/Poor				
SRH: Good/Very Good/Excellent				0.68 (0.63, 0.74) <i>p</i> < 0.01
BMI: Underweight				1.67 (1.48, 1.88) <i>p</i> < 0.01
BMI: Normal weight				Ref

Table 2. (Continued)

BMI:	0.83 (0.77, 0.89)
Overweight	$p < 0.01$
BMI: Obese	0.84 (0.76, 0.92)
	$p < 0.01$
EDU: None	
EDU:	0.93 (0.87, 0.99)
School/High school certificate	$p = 0.02$
EDU: Trade/ Diploma	0.82 (0.74, 0.90)
	$p < 0.01$
EDU: Higher education	0.74 (0.63, 0.88)
	$p < 0.01$
ARIA:	Ref
Major City	
ARIA: Inner Regional	1.08 (1.02, 1.15)
	$p = 0.01$
ARIA: Outer Regional	1.06 (0.98, 1.15)
	$p = 0.12$
ARIA: Remote/Very remote	1.33 (1.10, 1.60)
	$p < 0.01$
MS:	Ref
Married/De facto	
MS: Separated/ Divorced	1.09 (0.96, 1.24)
	$p = 0.21$
MS: Widowed	1.0 (0.94, 1.08)
	$p = 0.91$
MS: Never married	1.31 (1.12, 1.54)
	$p < 0.01$

Model 2 also included disease as a predictor for survival, in addition to sleep. Women reporting 1-2 diseases were at no greater risk compared to women reporting 0 diseases, but women reporting 3 or more diseases

had a HR = 1.27 (95%CI (1.17, 1.39), $p < 0.01$) compared to women with 0 diseases. This effect was strengthened slightly after adjusting for an interaction between sleep and disease (Model 3); but once all other covariates were included, having 3 or more diseases was no longer a significant predictor of death. However having 1-2 diseases was now protective, HR = 0.86 (95%CI (0.76, 0.97), $p = 0.01$). The interaction effect between sleep and disease was not significant in either Model 3 or 4.

Improved physical function and mental health were associated with a decrease in the hazard of death (HR = 0.991, 95%CI (0.99, 0.993), $p < 0.01$; and HR = 0.997, 95%CI (0.995, 0.999), $p < 0.01$, respectively). Also associated with decreased hazard of death was self-rated health, with women who rated their health as good/very good/excellent having a HR = 0.68 (95%CI (0.63, 0.74), $p < 0.01$), compared to women who rated their own health as fair/poor. Compared to women with no education, educated women were at a lower risk of death, with HR = 0.93 (95%CI (0.87, 0.99), $p = 0.02$); HR = 0.82 (95%CI (0.74, 0.90), $p < 0.01$) and HR = 0.74 (95%CI (0.63, 0.88), $p < 0.01$) for school/higher school certificate, trade/diploma and higher education levels, respectively. Women who were overweight or obese also had lower hazard of death, with HR = 0.83 (95%CI (0.77, 0.89), $p < 0.01$) and HR = 0.84 (95%CI (0.76, 0.92), $p < 0.01$), respectively, compared to normal weight women. However, women who were underweight had a greater hazard of death compared to women of normal weight, with HR = 1.67 (95%CI (1.48, 1.88), $p < 0.01$).

Also associated with greater hazard of death was area of residence. Compared to women in major cities, women in inner regional areas had a HR = 1.08 (95%CI (1.02, 1.15), $p = 0.01$), and women in remote or very remote areas had HR = 1.33 (95%CI (1.10, 1.60), $p < 0.01$). The effect was non-significant for women in outer regional areas. Baseline age was a significant predictor of mortality, with HR = 1.12 (95%CI (1.10, 1.15), $p < 0.01$). Finally, women who were never married had HR = 1.31 (95%CI (1.12, 1.54), $p < 0.01$) compared to married/de facto women, but the effect was not significant for widows or women who were separated or divorced. Vitality was not a significant predictor of mortality.

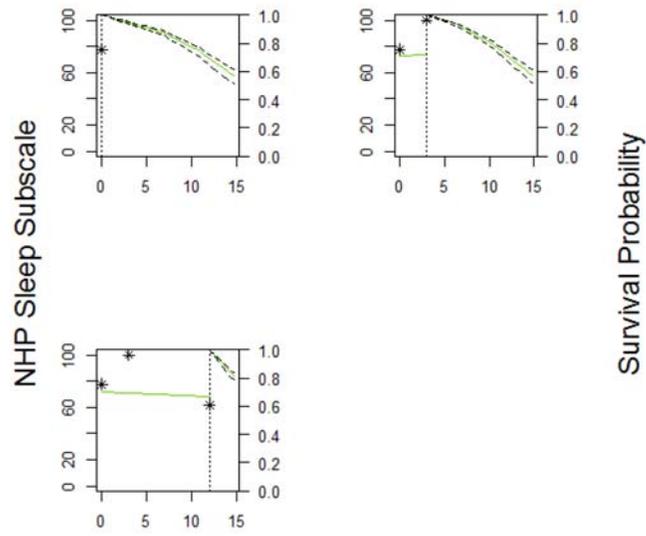
4.4. Dynamic survival predictions

Dynamic survival predictions based on Model 1 were created for a select sample of individuals' longitudinal profiles. To compare these results to the previous sleep analysis on the cohort [53], subjects were selected so as to include one subject from each of the four latent sleep classes identified by Leigh et al. [53], and included also a range of dropout times.

Model 1 was chosen to create the predictions, as these results represent the effect of sleep on survival adjusted only for age, and so are most directly comparable to the latent sleep classes obtained from the previous analysis [53]. These predictions are displayed for women who had data for all three longitudinal time points (Figure 2), surveys 2 and 3 only (Figure 3), and survey 2 only (Figure 4). The accompanying tables containing specific survival probabilities for each subgraph are included in Supplementary Material 2 (see Tables 2-4).

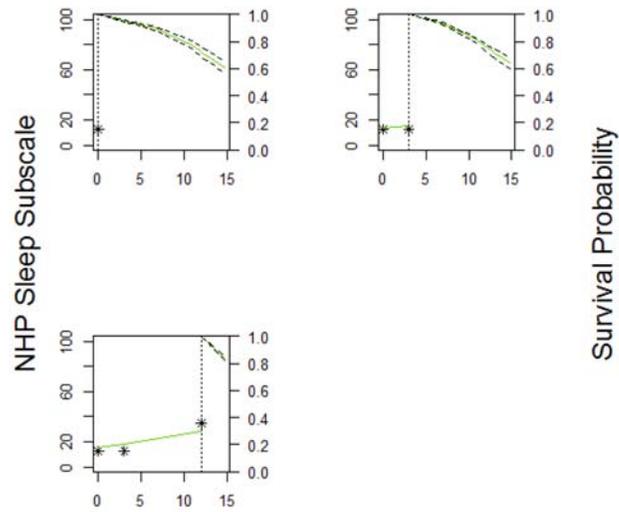
For women with sleep data for surveys 2, 3, and 6 (Figure 2), little discernible difference exists in the survival predictions across the four latent sleep classes. The results are similar for women who had sleep data for surveys 2 and 3 (Figure 3), although in this case the woman from the troubled sleeping class (top left corner) appears to have slightly lower predicted survival at censoring than the women from the remaining three classes. Similarly for women who only answered survey 2 (Figure 4), the woman from the troubled sleeping class appears to have slightly lower predicted survival at censoring than the women from the other three classes.

Survival Predictions



(a)

Survival Predictions



(b)

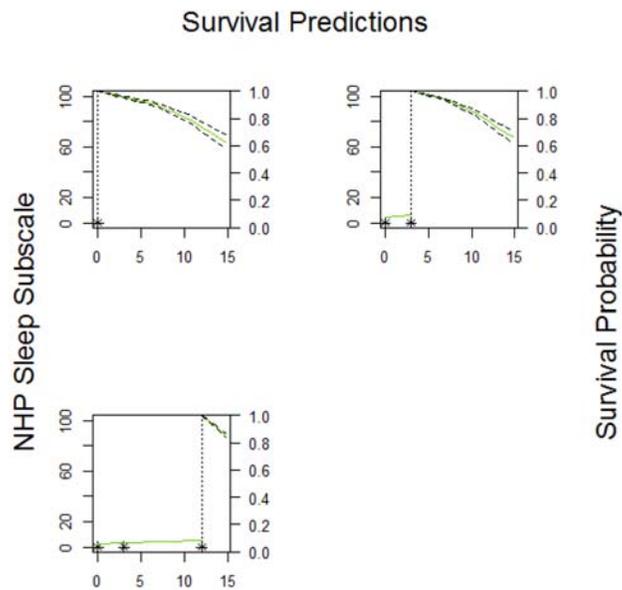
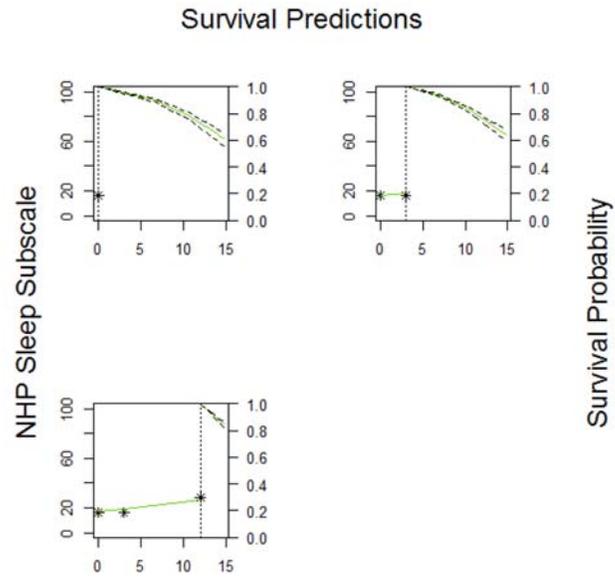
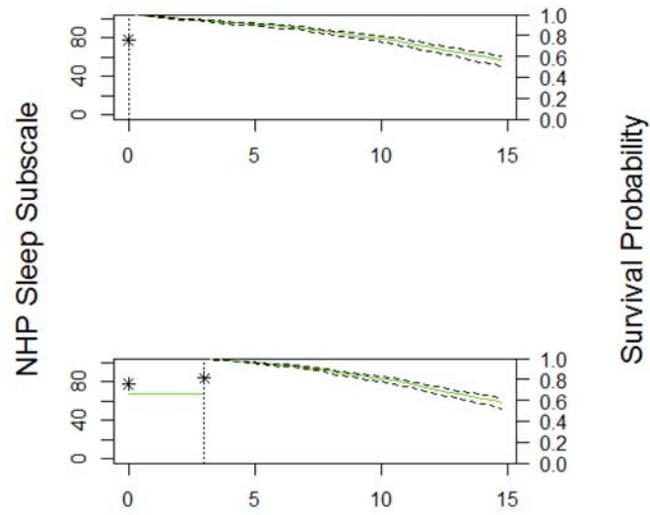


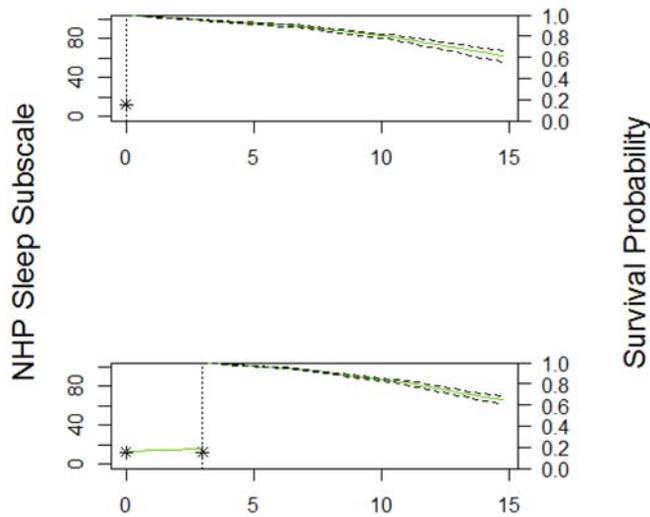
Figure 2. Dynamic survival predictions for subjects surviving until survey 6, Model 1. Participants belonged to (a) ‘troubled sleepers’, (b) ‘early wakers’, (c) ‘trouble falling asleep’, and (d) ‘untroubled sleepers’ classes [53].

Survival Predictions



(a)

Survival Predictions



(b)

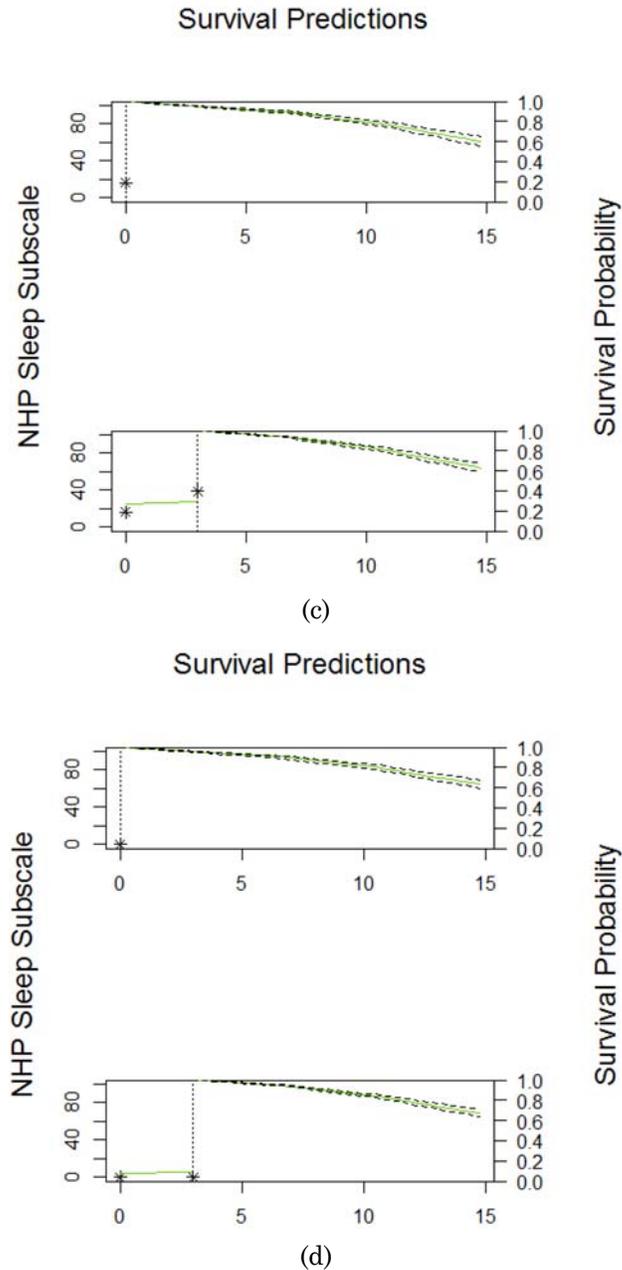
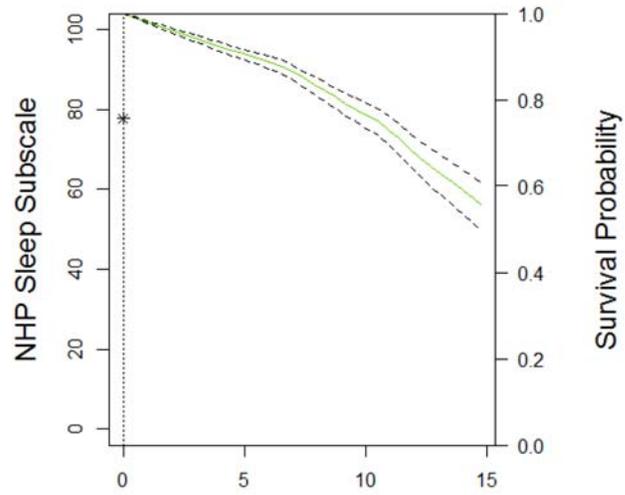


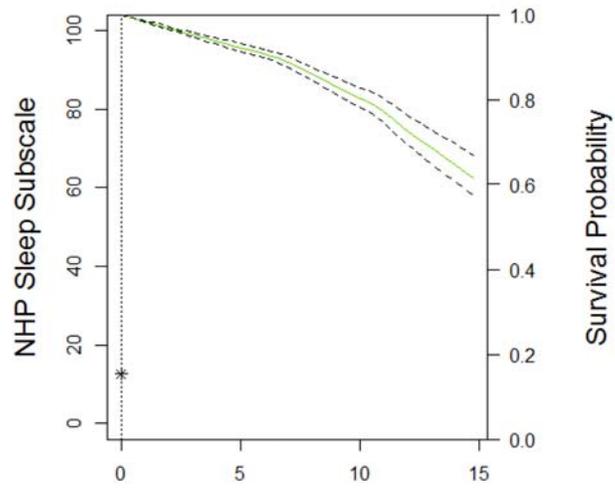
Figure 3. Dynamic survival predictions for subjects surviving until survey 3, Model 1. Participants belonged to (a) ‘troubled sleepers’, (b) ‘early wakers’, (c) ‘trouble falling asleep’, and (d) ‘untroubled sleepers’ class [53].

Survival Predictions



(a)

Survival Predictions



(b)

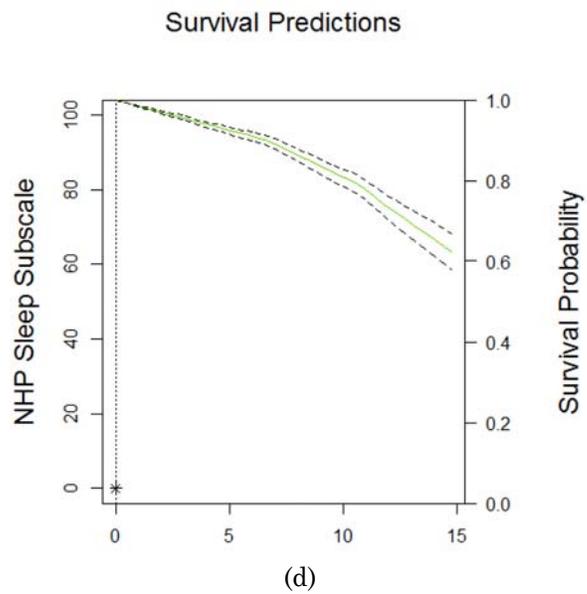
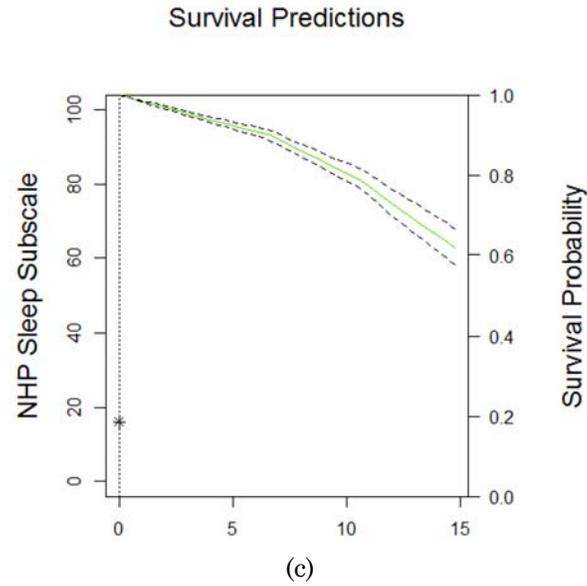


Figure 4. Dynamic survival predictions for subjects surviving until survey 2, Model 1. Participants belonged to (a) ‘troubled sleepers’, (b) ‘early wakers’, (c) ‘trouble falling asleep’, and (d) ‘untroubled sleepers’ class in previous analysis [53].

5. Discussion

Longitudinal trajectories of sleep difficulty were strongly associated with disease count, self-rated health, quality of life measures (physical function, vitality and mental health), as well as baseline age, marital status and highest level of education. This supports the current literature that reports associations between poor sleep and a variety of illnesses and health complaints. However surprisingly, given the correlation between obesity and sleep apnoea, BMI was not a significant predictor for sleep difficulty.

Our first hypothesis was that greater sleep difficulty would be associated with lower survival. This was supported by the results from Model 1, with increased sleep difficulty associated with greater hazard of death. Dropout was assumed to be NMAR, and unlike a previous analysis of the data [53] (which assumed MAR dropout), after adjusting for multiple health-related and demographic covariates, the effect of sleep on survival was no longer significant. However, this may be due to the fact that sleeping difficulty may not play a causative role in poorer survival, but rather it is correlated with other factors such as disease which do have a causal role. This correlation is supported and demonstrated by the significant impacts of disease, self-rated health, quality of life, and other demographics, on sleep in the longitudinal part of the JM. Use of sleep medications was also factored into the NHP sleep difficulty score, so it is unlikely that use of sleep medications is confounding the results by masking sleep difficulty.

Our second aim was to investigate the interaction effect between longitudinal sleep and disease on survival. We hypothesized that disease would be a significant predictor for survival, and that an interaction effect between disease and sleep difficulty would exist. This relationship was evident in the results of the models with only disease and sleep (Models 2 and 3), where having 3 or more diseases was associated with greater hazard of death. However surprisingly, after adjusting for all

covariates, having 3 or more diseases was no longer a significant predictor for death. To test whether this might be due to collinearity between the number of diseases and other health-related variables, such as BMI, self-rated health, and physical functioning, we utilised Variance Inflation Factor methods [97-99], the results of which are included in Supplementary Material 3. There was no indication of strong collinearity. However, having 1-2 diseases became a significant *protective* predictor for survival in Model 4. This may be due to the fact that women who survive longer have greater opportunity to develop disease.

One of the strongest predictors of survival was self-rated health. Although correlation between self-rated health and disease was only moderate ($r = -0.29$, $p < 0.0001$, Supplementary Material 3), it may be that self-rated health separates out those women whose health is suffering more from their morbidities (e.g., increased pain, disability, decreased quality of life), from those whose health remains fairly robust in spite of their diseases. Physical functioning was also an important predictor for survival, with improved physical function associated with lower hazard of death. Again, physical function scores may be identifying the women who suffer more from their diseases.

Our results also support the findings of other studies which have found that post-menopause, being overweight is protective against death [100, 101]. In fact our results suggest this protective effect may extend to obese subjects. This is in contrast to a recent meta-analysis [100], which found that obesity was linked to higher mortality; note however, that our study does not differentiate between the various levels of obesity. Conversely, we found that women who were underweight were at a greater risk of death compared to normal weight women. We stop short of recommending obesity in older ages; further evidence is needed of the impact of obesity in old age, ideally utilising models which treat BMI as an endogenous predictor.

Unlike previous analysis of this data [53], area of residence was a significant predictor of mortality, with women living in inner regional and remote or very remote areas at a greater hazard of death.

In studies where both a longitudinal outcome and a survival outcome exist, traditionally the effects are analysed separately, with one model for the longitudinal outcome, and another for the survival outcome [59]. The longitudinal outcome is often also included as a time-varying predictor for the survival outcome, usually within a Cox Proportional Hazards model. Such models assume that the longitudinal outcome is constant between observation times. Additionally, the longitudinal marker or predictor is assumed to be exogeneous – i.e., that the longitudinal trajectory of the covariate is independent of, and therefore contains no information about, the survival outcome. However, it is more likely that the longitudinal outcome may be influenced by the failure mechanism [59]. Furthermore, when assessing the trajectory of the longitudinal outcome, standard mixed effects models treat dropout as non-random. Estimation essentially handles the missing data, due to dropout, by interpolating their trajectories from the remaining sample. This treats the trajectories of the dropouts and survivors as the same, and can lead to bias [59, 82]. One of the strengths of this study is the use of JM. Jointly modelling the two outcomes avoids the aforementioned problems and allows for valid inferences on the longitudinal and survival processes [82].

We utilised ‘JM’ to simultaneously analyse longitudinal and survival data, however other options exist. Other software platforms that perform JM include ‘joineR’ which utilises shared latent random effects [102], STJM in stata [103], MATLAB [104, 105], SAS [106], and HETMIXSURV on Fortran 90 [107]. Selection models [80, 108], pattern mixture models [109], stochastic process modelling [110], or even joint latent class mixture models [111, 112] could be also have been utilised instead of shared-parameter JMs. Joint latent class mixture models (JLCMM) were initially trialled for this analysis; however the model complexity, and data size and complexity, proved to be too computationally demanding to be handled by standard computing power. Joint modelling using shared-parameter models was chosen as an alternative for this work, due to its widespread use in the field of biostatistics.

'JM' utilises maximum-likelihood estimation, however it should be noted that Bayesian estimation techniques are also available [80]. The expectation-maximisation algorithm of Dempster et al. [113] and Wulfsohn and Tsiatis [81] can lead to intractable integrals, since there is no closed-form solution for either the integral involving the random effects, or the integral involving the survival function [80]. Numerical approximations exist, such as adaptations of the Gauss-Hermite quadrature rule [81, 114-116]. More recently, Laplace approximations have been trialled to reduce the computational complexity [80]. However convergence rates remain slow, and prove to be a major disadvantage of using ML techniques for joint modelling. Bayesian methods have been developed which take advantage of the computational advantages associated with Bayesian approaches. The most common Bayesian method is that of Faucett and Thomas [117] which uses MCMC algorithm and Gibbs sampling [104, 118-120]. However, Bayesian methods require careful choice of prior distributions and sensitivity analysis for model validation [80].

One limitation of the current analysis is that only the longitudinal sleep and survival outcomes were modelled jointly. Other variables of interest were treated as exogeneous and included as time-varying covariates in the survival model as per a Cox Proportional Hazards model. Literature exists for modelling multiple longitudinal outcomes jointly with a survival outcome [115, 121-127]. However, Rizopoulos states that while such models are mathematically straightforward, in practice they are computationally demanding [82]. At this point in time, there is no readily available software [128] that allows for multiple longitudinal outcomes to be modelled jointly with a survival outcome. The high computational requirements of joint models limit the functionality of current freely available and accessible software packages to simpler models [80]. Sleep was chosen to be modelled as the longitudinal outcome, since it was the main predictor of interest in the current study. The current functionality of JM is also limited to a normal

longitudinal outcome [59]. However ideally, other variables such as BMI, disease and self-rated health would also be modelled jointly. Such a model would help determine, for instance, whether the significant effect of having 1-2 diseases on mortality in Model 4 was a true protective effect, or a result of survivors having greater opportunity to develop and report disease. Advancements in the field of JM may in the future allow for more complicated models to be computationally viable, for example, the work of Barrett et al. [128], which utilises a discretisation of the time-scale for the time-to-event outcome (as opposed to continuous time).

Another limitation of the current study is that the NHP sleep subscale is not directly comparable to the latent sleep classes of earlier work [53], which captured more complex longitudinal sleep difficulty patterns than simply ‘good sleep’ and ‘poor sleep’. In future work, we plan to model alternative continuous covariates, such as physical functioning, mental health and vitality, as the endogeneous longitudinal outcome, and to utilise the latent sleep classes [53] as an exogeneous predictor of survival.

Residual diagnostics were also performed by assessing residual plots from Models 1-4 across the multiple imputed analyses (results not included). There was some residual pattern evident in these plots, however this did not improve when various model parameters were changed (for instance, adding a quadratic time effect in the longitudinal sub-model). It is possible that the use of a different baseline hazard function for the survival model, such as a Weibull, may improve the residual plots. However, currently the possibility to use different hazard functions is not available in JM or any other available software when time-varying covariates are included in the survival sub-model.

An avenue identified for future extension of the analysis is to use exact chronological age, rather than follow-up years from baseline, as the time metameter in both the survival and longitudinal sub-models. However, this requires software that allows for left-censoring as well as right-censoring of the survival data. The ‘JMbates’ package is currently being developed to include left-censoring in its functionality [89, 129].

Another limitation of this study is that the possible collinearity between sleep, disease, self-rated health, and quality of life, may be masking the effect of sleep on mortality (Model 4). Future analysis will use Bayesian profile regression [130, 131] to jointly model the various sleep and health covariates and survival. Such an approach forms classes based on the latent relationships between the multivariate covariates and the sleep outcome, and can handle collinearity. Alternatively, in the future variable selection approaches may become available for use with JM [132]. Another possible alternative is Bayesian model averaging, which allows predictions to be based on a collection of models simultaneously, rather than a single model, and explicitly accounts for model uncertainty and ‘acknowledges that a single prognostic model may not be adequate for quantifying the risk of all future patients’ [133].

6. Conclusion

The association between sleep and mortality was investigated for a cohort of very old women, using shared random effects joint modelling to account for informative dropout. We have shown that sleep difficulty is correlated with other factors, such as disease, which have a proven causal effect on poorer survival. Sleep difficulty was significantly associated with increased disease count, decreased quality of life, decreased self-rated health, and education. Greater sleep difficulty was also associated with greater hazard of death in unadjusted and disease adjusted models. However, after adjusting for disease, other health-related covariates and demographics, sleep was not significantly associated with mortality. We expected that having 3 or more diseases would be a significant predictor of survival; however after adjusting for other health-related predictors the effect was non-significant. Women reporting 1-2 diseases had lower hazard of death in the fully-adjusted model, however future modelling with multiple longitudinal outcomes is required to investigate whether, once disease is also treated as an endogeneous rather than an exogeneous predictor, this protective effect remains significant. The main health

related predictors of mortality in this analysis were BMI, self-rated health, physical functioning and mental health. Therefore, it may be most useful for health practitioners to focus on these, rather than on sleep difficulty or disease, when assessing mortality risk in very old women.

Acknowledgements

This research was conducted as part of the Australian Longitudinal Study on Women's Health (ALSWH), the University of Newcastle and the University of Queensland. We are grateful to the Australian Government Department of Health and Ageing for funding, to the staff of ALSWH, and to the women who provided the survey data. Researchers in the Research Centre for Generational, Health and Ageing at the University of Newcastle are members of the Hunter Medical Research Institute (HMRI), Newcastle, Australia.

References

- [1] S. K. Roepke and S. Ancoli-Israel, Sleep disorders in the elderly, *Indian Journal of Medical Research* 131 (2010), 302-310.
- [2] B. Phillips and S. Ancoli-Israel, Sleep disorders in the elderly, *Sleep Medicine* 2 (2001), 99-114.
- [3] D. L. Bliwise, Sleep in normal aging and dementia, *Sleep* 16(1) (1993), 40-81.
- [4] E. J. W. Van Someren, Circadian and sleep disturbances in the elderly, *Experimental Gerontology* 35 (2000), 1229-1237.
- [5] X. Liu and L. Liu, Sleep habits and insomnia in a sample of elderly persons in China, *Sleep* 28(12) (2005), 1579-1587.
- [6] D. J. Foley et al., Sleep complaints among elderly persons: An epidemiologic study of three communities, *Sleep* 18 (1995), 425-32.
- [7] F. Mousavi and B. Golestan, Insomnia in the elderly population: A study in hospital clinics of Tehran, *Iran Journal of Sleep Research* 18 (2009), 481-482.
- [8] P. Salo et al., Trajectories of sleep complaints from early midlife to old age: Longitudinal modeling study, *Sleep* 35(11) (2012), 1559-1568.

- [9] C. C. Hoch et al., Longitudinal changes in diary- and laboratory-based sleep measures in health Old Old and Young Old subjects: A three-year follow-up, *Sleep* 20(3) (1997), 192-202.
- [10] M. Hägg et al., Sleep quality, use of hypnotics and sleeping habits in different age-groups among older people, *Scandinavian Journal of Caring Sciences* 28 (2014), 842-851.
- [11] J. M. Jacobs et al., Gender differences in survival in old age, *Rejuvenation Research* 17(6) (2014), 499-506.
- [12] R. Sterniczuk, B. Rusak and K. Rockwood, Sleep disturbances in older ICU patients, *Clinical Interventions in Aging* 9 (2014), 969-977.
- [13] P. Lavie and L. Lavie, Unexpected survival advantage in elderly people with moderate sleep apnoea, *Journal of Sleep Research* 18 (2009), 397-403.
- [14] L. Gallicchio and B. Kalesan, Sleep duration and mortality: A systematic review and meta-analysis, *Journal of Sleep Research* 18 (2009), 148-158.
- [15] F. P. Cappuccio et al., Sleep duration and all-cause mortality: A systematic review and metaanalysis of prospective studies, *Sleep* 33(5) (2010), 585-592.
- [16] L. Kurina et al., Sleep duration and all-cause mortality: A critical review of measurement and associates, *Annals of Epidemiology* 23 (2013), 361-370.
- [17] J. E. Gangwisch et al., Sleep duration associated with mortality in elderly, but not middle-aged, adults in a large US sample, *Sleep* 31(8) (2008), 1087-1096.
- [18] D. A. Katz and C. A. McHorney, Clinical correlates of insomnia on patients with chronic illness, *Archives of Internal Medicine* 158 (1998), 1099-1107.
- [19] D. Foley et al., Sleep disturbances and chronic disease in older adults, Results of the 2003 National Sleep Foundation Sleep in America survey, *Journal of Psychosomatic Research* 56 (2004), 497-502.
- [20] C. A. Magee et al., Short and long sleep duration are associated with prevalent cardiovascular disease in Australian adults, *Journal of Sleep Research* 21 (2011), 441-447.
- [21] S. Ancoli-Israel, Sleep and its disorders in aging populations, *Sleep Medicine* 10 (2009), S7-S11.
- [22] A. Koyanagi et al., Chronic conditions and sleep problems among adults aged 50 years or over in nine countries: A multi-country study, *PLoS ONE* 9(12) (2014), e114742.
- [23] J. E. Gangwisch et al., Sleep duration as a risk factor for diabetes incidence in a large US sample, *Sleep Medicine* 30(12) (2007), 1667-1673.
- [24] L. Mallon, J. E. Broman and J. Hetta, High incidence of diabetes in men with sleep complaints or short sleep durations: A 12-year follow-up study of a middle-aged population, *Diabetes Care* 28(11) (2005), 2762-2767.

- [25] J. E. Gangwisch et al., Short sleep duration as a risk factor for hypertension: Analyses of the First National Health and Nutrition Examination Survey, *Hypertension* 47(5) (2006), 833-839.
- [26] D. J. Gottlieb et al., Association of usual sleep duration with hypertension: The Sleep Heart Health Study, *Sleep* 29(8) (2006), 1009-1014.
- [27] M. A. Arias, S. Bartolomé and L. Rodriguez-Padial, Obesity, cardiovascular disease and obstructive sleep apnea [letter], *American Journal of Medicine* 121(5) (2008), e13.
- [28] E. Suzuki et al., Sleep duration, sleep quality and cardiovascular disease mortality among the elderly: A population-based cohort study, *Preventive Medicine* 49 (2009), 135-141.
- [29] H. Palomäki et al., Snoring, sleep apnea syndrome, and stroke, *Neurology* 42(7 suppl 6) (1992), 75-81.
- [30] K. N. Anderson et al., Assessment of sleep and circadian rhythm disorders in the very old: The Newcastle 85+ Cohort Study, *Age and Ageing* 43 (2014), 57-63.
- [31] J. E. Maglione et al., Subjective and objective sleep disturbance and longitudinal risk of depression in a cohort of older women, *Sleep* 37(7) (2014), 1179-1187.
- [32] R. Faubel et al., Usual sleep duration and cognitive function in older adults in Spain, *Journal of Sleep Research* 18 (2009), 427-435.
- [33] T. E. Gildner et al., Associations between sleep duration, sleep quality, and cognitive test performance among older adults from six middle income countries: Results from the Study on Global Ageing and Adult Health (SAGE), *Journal of Clinical Sleep Medicine* 10(6) (2014), 613-621.
- [34] T. Blackwell et al., Associations of objectively and subjectively measured sleep quality with subsequent cognitive decline in older community-dwelling men: The MrOS study, *Sleep* 37(4) (2014), 655-663.
- [35] J. P. Chaput et al., The association between sleep duration and weight gain in adults: A 6-year prospective study from the Quebec Family Study, *Sleep* 31(4) (2008), 517-523.
- [36] H. G. Bloom et al., Evidence-based recommendations for the assessment and management of sleep disorders in older persons, *Journal of the American Geriatrics Society* 57 (2009), 761-789.
- [37] K. L. Knutson et al., The metabolic consequences of sleep deprivation, *Sleep Medicine Reviews* 11(3) (2007), 163-178.
- [38] M. A. Miller and F. P. Cappuccio, Inflammation, sleep, obesity, and cardiovascular disease, *Current Vascular Pharmacology* 5(2) (2007), 93-102.
- [39] M. Kumari et al., Self-reported sleep duration and sleep disturbance are independently associated with cortisol secretion in the Whitehall II study, *Journal of Clinical Endocrinology and Metabolism* 94 (2009), 4801-9.
- [40] G. Copinschi, Metabolic and endocrine effects of sleep deprivation, *Essential Psychopharmacology* 6(6) (2005), 341-7.

- [41] M. A. Grander and S. P. A. Drummond, Who are the long sleepers? Towards an understanding of the mortality relationship, *Sleep Medicine Reviews* 11(5) (2007), 341-360.
- [42] W. V. McCall, Sleep in the elderly: Burden, diagnosis, and treatment, *Journal of Clinical Psychiatry* 6(1) (2004), 9-20.
- [43] K. L. Stone et al., Sleep disturbances and risk of falls in older community-dwelling men: The Outcomes of Sleep Disorders in Older Men (MrOS Sleep) Study, *Journal of the American Geriatrics Society* 62 (2014), 299-305.
- [44] K. L. Stone, K. E. Ensrud and S. Ancoli-Israel, Sleep, insomnia and falls in elderly patients, *Sleep Medicine Suppl.* 1 (2008), S18-S22.
- [45] K. E. Ensrud et al., Sleep disturbances and frailty status in older community-dwelling men, *Journal of the American Geriatrics Society* 57 (2009), 2085-2093.
- [46] S. Ancoli-Israel and J. R. Cooke, Prevalence and comorbidity of insomnia and effect on functioning in elderly populations, *Journal of the American Statistical Association* 83(Suppl.) (2005), S262-S271.
- [47] Institute of Medicine (US) Committee on Sleep Medicine and Research, *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*, H. R. Colton and B. M. Altevogt (editors), National Academies Press (US): Washington DC, 2006.
Available from: <http://www.ncbi.nlm.nih.gov/books/NBK19961/>
- [48] D. R. Hillman and L. C. Lack, Public health implications of sleep loss: The community burden, *Medical Journal of Australia* 199(8) (2013), S7-S10.
- [49] C. A. Smyth, Evaluating sleep quality in older adults, *American Journal of Nursing* 108(5) (2008), 42-50.
- [50] S. Hasan et al., Use of sleeping medication and quality of life among older women who report sleeping difficulty, *Australasian Journal on Ageing* 20 (2001), 29-35.
- [51] J. E. Byles, G. D. Mishra and M. A. Harris, The experience of insomnia among older women, *Sleep* 28(8) (2005), 972-979.
- [52] S. M. Hunt et al., The Nottingham Health Profile: Subjective health status and medical consultations, *Social Science and Medicine* 15A (1981), 221-229.
- [53] L. Leigh, I. L. Hudson and J. Byles, Sleeping difficulty, disease and mortality in older women: A latent class analysis and distal survival analysis, *Journal of Sleep Research* 24 (2015), 648-657.
- [54] L. M. Collins and S. T. Lanza, *Latent Class and Latent Transition Analysis with Applications in the Social, Behavioural and Health Sciences* Wiley Series in Probability and Statistics, D. J. Balding et al. (editor), Hoboken, New Jersey: John Wiley & Sons, 2010.
- [55] T. Asparouhov and B. O. Muthén, *Auxiliary Variables in Mixture Modeling: 3-Step Approaches Using Mplus*, 2013. Accessed 3/5/2014.

<http://www.statmodel.com/download/3stepOct28.pdf>

- [56] D. W. Gudicha and J. K. Vermunt, Mixture model clustering with explanatory variable: One step and three step approaches, in *Algorithms from and for Nature and Life; Studies in Classification, Data Analysis, and Knowledge Organisation*, B. Lausen, D. van den Poel and A. Ultsch (editors), Springer-Verlag GmbH: Heidelberg, p. 87-93 (2011).
- [57] S. T. Lanza, X. Tan and B. C. Bray, Latent Class Analysis with distal outcomes: A flexible modelbased approach, *Structural Equation Modeling* 20(1) (2013), 1-26.
- [58] C. K. Enders, *Applied Missing Data Analysis, Methodology in the Social Sciences*, T. D. Little (editor), New York: Guilford Press, 2010.
- [59] D. Rizopoulos, JM: An R package for the joint modeling of longitudinal and time-to-event data, *Journal of Statistical Software* 35(9) (2010).
- [60] J. Núñez et al., Red blood cell distribution width is longitudinally associated with mortality and anemia in heart failure patients, *Circulation Journal* 78(2) (2014), 410-418.
- [61] E. N. Njagi et al., A joint survival-longitudinal modelling approach for the dynamic prediction of rehospitalization in telemonitored chronic heart failure patients, *Statistical Modelling* 13(3) (2013), 179-198.
- [62] A. Seid et al., Joint modeling of longitudinal CD4 cell counts and time-to-default from HAART treatment: A comparison of separate and joint models *Electronic Journal of Applied Statistical Analysis* 7(2) (2014), 292-314.
- [63] R. Gueorguieva, R. A. Rosenheck and H. Lin, Joint modeling of longitudinal outcome and interval-censored competing risk dropout in a schizophrenia clinical trial, *Journal of the Royal Statistical Society, Series A, Statistics in Society* 175(2) (2012), 417-433.
- [64] B. Yu and P. Ghosh, Joint modeling for cognitive trajectory and risk of dementia in the presence of death, *Biometrics* 66 (2010), 294-300.
- [65] I. Guler et al., Joint modeling for longitudinal and time-to-event data: Application to liver transplantation data, *Lecture Notes in Computer Science* (2014), 8581 LNCS(PART 3) 580-593.
- [66] L. C. Diaz, *Joint Modelling for Longitudinal and Time-to-Event Data, Application to Liver Transplantation Data*, University of Santiago de Compostela, 2014.
http://eio.usc.es/pub/mte/descargas/ProyectosFinMaster/Proyecto_1079.pdf
- [67] N. Terrin, A. M. Rodday and S. K. Parsons, Joint models for predicting transplant-related mortality from quality of life data, *Quality of Life Research* 24 (2015), 31-39.
- [68] M. D. Schluchter, M. W. Konstan and P. B. Davis, Joint modelling the relationship between survival and pulmonary function in cystic fibrosis patients, *Statistics in Medicine* 21 (2002), 1271-1287.
- [69] E. Deslandes and S. Chevret, Joint modeling of multivariate longitudinal data and the dropout process in a competing risk setting: application to ICU data, *BMC Medical Research Methodology* 10(69) (2010).

- [70] L. A. Hatfield, M. E. Boyle and J. B. Carlin, Joint modeling of multiple longitudinal patient-reported outcomes and survival, *Journal of Biopharmaceutical Statistics* 21(5) (2011), 971-991.
- [71] Medicare Australia.
Available from: <http://www.humanservices.gov.au/customer/dhs/medicare>
- [72] W. Brown et al., Women's health Australia: Establishment of the Australian Longitudinal Study on Women's Health, *Journal of Women's Health* 5(5) (1996), 467-472.
- [73] C. Lee et al., Cohort profile: The Australian Longitudinal Study on Women's Health, *International Journal of Epidemiology* 34 (2005), 987-997.
- [74] S. M. Hunt et al., Nottingham Health Profile Scoring Instructions; Accessed 16/03/2015.
<http://reseauconceptuel.umontreal.ca/rid=1J1WXGYDK-CPP19Z-1ZQY/Nottingham.pdf>
- [75] World Health Organization, Obesity and Overweight, Fact Sheet 311, 2011 [14/6/2011].
Available from: <http://www.who.int/mediacentre/factsheets/fs311/en/index.html>
- [76] J. E. Ware, M. Kosinski and S. D. Keller, SF-36 Physical and Mental Health Summary Scales: A User's Manual, The Health Institute: Boston, MA, 1994.
- [77] Department of Health and Aged Care (GISCA), Measuring remoteness: Accessibility/remoteness index of Australia (ARIA), Revised Ed. Occasional Paper New Series No. 14. 2001. Department of Health and Aged Care, Canberra.
- [78] J. Powers et al., Effectiveness of the National Death Index for establishing the vita status of older women in the Australian Longitudinal Study on Women's Health, *Australian and New Zealand Journal of Public Health* 24(5) (2000), 526-528.
- [79] D. Rizopoulos, Joint Modeling of Longitudinal and Survival Data, Package 'JM'; Version 1.3.0; 2014, 2015 (access date 10/02/2015).
<http://cran.r-project.org/web/packages/JM/JM.pdf>
- [80] L. M. McCrink, A. H. Marshall and K. J. Cairns, Advances in joint modelling: A review of recent developments with application to the survival of end stage renal disease patients, *International Statistical Review* 81(2) (2013), 246-269.
- [81] M. S. Wulfsohn and A. A. Tsiatis, A joint model for survival and longitudinal data measured with error, *Biometrics* 53 (1997), 330-339.
- [82] D. Rizopoulos, Joint Models for Longitudinal and Time-to-Event Data: With Applications in R, Chapman & Hall/CRC Biostatistics Series, B. Jones et al. (editors), Boca Raton: CRC Press, 2012.

- [83] J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data*. Wiley Series in Probability and Mathematical Statistics, R. A. Bradley et al. (editors), New York: John Wiley and Sons, 1980.
- [84] A. Tsiatis and M. Davidian, Joint longitudinal and time-to-event data: An overview, *Statistica Sinica* 14 (2004), 809-834.
- [85] F. Hsieh, Y.-K. Tseng and J.-L. Wang, Joint modeling of survival and longitudinal data: Likelihood approach revisited, *Biometrics* 62 (2006), 1037-1043.
- [86] J. P. Klein and M. L. Moeschberger, *Survival Analysis, Techniques for Censored and Truncated Data*, Second Edition, *Statistics for Biology and Health*, NY: Springer Science and Business Media Inc., 2003.
- [87] D. Rizopoulos, Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive Gaussian quadrature rule, *Computational Statistics and Data Analysis* 56 (2012), 491-501.
- [88] C. Proust-Lima and J. Taylor, Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of post treatment PSA: A joint modeling approach, *Biostatistics* 10 (2009), 535-549.
- [89] D. Rizopoulos, The R package JM bayes for fitting joint models for longitudinal and time-to-event data using MCMC, 2014. Access date 16/03/2015.
 Accessed online: <http://arxiv.org/pdf/1404.7625v1.pdf>
- [90] D. Rizopoulos, Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data, *Biometrics* 67(3) (2011), 819-829.
- [91] S. van Buuren and K. Groothuis-Oudshoorn, MICE: Multivariate imputation by chained equations in R, *Journal of Statistical Software* 45(3) (2011), 1-67.
- [92] J. W. Graham, A. E. Olchowski and T. D. Gilreath, How many imputations are really needed? Some practical clarifications of multiple imputation theory, *Prevention Science* 8 (2007), 206-213.
- [93] T. E. Bodner, What improves with increased missing data imputations? *Structural Equation Modeling* 15 (2008), 651-675.
- [94] I. R. White, P. Royston and A. M. Wood, Multiple imputation using chained equations: Issues and guidance for practice, *Statistics in Medicine* 30 (2011), 377-399.
- [95] J. Starkweather, *Linear Mixed Effects Modeling Using R* [3/3/2015].
 Available from:
http://www.unt.edu/rss/class/Jon/Benchmarks/LinearMixedModels_JDS_Dec2010.pdf
- [96] J. Pinheiro et al., *Linear and Nonlinear Mixed Effects Models; Package nlme; Version 3.1.119*, 2015 [cited 2015 13/02/2015].
 Available from: <http://cran.rproject.org/web/packages/nlme/nlme.pdf>

- [97] J. Fox et al., Companion to Applied Regression. R package car; Version 2.0.24; 2015.
<http://cran.rproject.org/web/packages/car/car.pdf>
- [98] P. D. Allison, Survival Analysis, in *The Reviewer's Guide to Quantitative Methods in the Social Sciences*, G. Hancock and R. O. Mueller (editors), Routledge: New York, 2010.
- [99] M. H. Kutner et al., Applied Linear Statistical Models, Fifth Edition, in *Operations and Decision Sciences*, McGraw-Hill Irwin: New York, 2005.
- [100] K. M. Flegal et al., Association of all-cause mortality with overweight and obesity using standard Body Mass Index categories, *Journal of the American Medical Association* 309(1) (2013), 71-82.
- [101] A. M. Beleigoli et al., Overweight and Class 1 Obesity are associated with lower 10-year risk of mortality in Brazilian older adults: The Bambui Cohort Study of Ageing, *PLoS ONE* 7(12) (2012), e52111.
- [102] P. Philipson, I. Sousa and P. J. Diggle, Joint modelling of repeated measurements and time-to-event data, in R package joineR; Version 1.0.3; 2015.
<http://cran.rproject.org/web/packages/joineR/joineR.pdf>
- [103] M. J. Crowther, STJM: Stata module to fit shared parameter joint models of longitudinal and survival data, Technical Report, Statistical Software Components, Boston College Department of Economics, 2011.
- [104] X. Guo and J. B. Carlin, Separate and joint modeling of longitudinal and event time data using standard computer packages, *American Statistician* 58(1) (2004), 16-24.
- [105] S. J. Ratcliffe, W. Gue and H. T. R. T., Joint modeling of longitudinal and survival data via a common frailty, *Biometrics* 60(4) (2004), 892-899.
- [106] SAS Institute Inc., SAS/STAT Software, Version 9.3.
URL <http://www.sas.com/> Cary, NC.
- [107] C. Proust-Lima, HETMIXSURV program V2.0. n.d. 2/03/2015]. Available from:
<http://etudes.isped.u-bordeaux2.fr/BIOSTATISTIQUE/HETMIXSURV/US-Biostats-HETMIXSURV.html>
- [108] A. M. Gad, A selection model for longitudinal data with non-ignorable non-monotone missing values, *Journal of Data Science* 9 (2011), 171-180.
- [109] R. J. A. Little, Pattern-mixture models for multivariate incomplete data, *Journal of the American Statistical Association* 88(421) (1993), 125-134.
- [110] K. G. Arbeeve et al., Joint analyses of longitudinal and time-to-event data in research on ageing: Implications for predicting health and survival, *Frontiers in Public Health* 2(1-12) (2014).
- [111] C. Proust-Lima et al., Joint latent class models for longitudinal and time-to-event data: A review, *Statistical Methods in Medical Research* 23(1) (2014), 74-90.

- [112] C. Proust-Lima, V. Phillips and B. Liqueur, Estimation of extended mixed models using latent classes and latent processes: The R package lcmm, Arxiv, 2015.
- [113] A. P. Dempster, N. M. Laird and D. B. Rubin, Maximum likelihood for incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B Statistical Methodology* 39 (1977), 1-38.
- [114] R. Henderson, P. Diggle and A. Dobson, Joint modeling of longitudinal measurements and event time data, *Biostatistics* 1(4) (2000), 465-480.
- [115] X. Song, M. Davidian and A. Tsiatis, An estimator for the proportional hazards model with longitudinal covariates measured with error, *Biostatistics* 3(4) (2002), 511-528.
- [116] D. Rizopoulos, G. Verbeke and G. Molenberghs, Shared parameter models under random effects misspecifications, *Biometrika* 95(1) (2008), 63-74.
- [117] C. L. Faucett and D. C. Thomas, Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach, *Statistics in Medicine* 15 (1996), 1663-1685.
- [118] C. L. Faucett, N. Schenker and J. M. G. Taylor, Survival analysis using auxilliary variables via multiple imputation, with application to AIDS clinical trial data, *Biometrics* 58(1) (2002), 37-47.
- [119] J. Xu and S. L. Zeger, The evaluation of multiple surrogate endpoints, *Biometrics* 57 (2001), 81-87.
- [120] E. Brown, J. Ibrahim and V. DeGruttola, A flexible B-spline model for multiple longitudinal biomarkers and survival, *Biometrics* 61 (2005), 64-73.
- [121] E. R. Brown and J. G. Ibrahim, A Bayesian semiparametric joint hierarchical model for longitudinal and survival data, *Biometrics* 59 (2003), 221-228.
- [122] Y.-Y. Chi and J. G. Ibrahim, Joint models for multivariate longitudinal and multivariate survival data, *Biometrics* 62 (2006), 432-445.
- [123] D. Rizopoulos and P. Ghosh, A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event, *Statistics in Medicine* 30 (2011), 1366-1380.
- [124] T. Baghfalaki, M. Ganjali and D. Berridge, Joint modeling of multivariate longitudinal mixed measurements and time to event data using a Bayesian approach, *Journal of Applied Statistics* 41(9) (2014), 1934-1955.
- [125] C. Proust-Lima et al., Joint modelling of multivariate longitudinal outcomes and a time-toevent: A nonlinear latent class approach, *Computational Statistics and Data Analysis* 53 (2009), 1142-1154.
- [126] E. Dantan et al., Pattern mixture models and latent class models for the analysis of multivariate longitudinal data with informative dropouts, *International Journal of Biostatistics* 4(1) (2008), 1-26.

- [127] C. Proust-Lima, L. Letenneur and H. Jacqmin-Gadda, A nonlinear latent class model for joint analysis of multivariate longitudinal data and a binary outcome, *Statistics in Medicine* 26 (2007), 2229-2245.
- [128] J. Barrett et al., Joint modelling of repeated measurements and time-to-event outcomes: flexible model specification and exact likelihood inference, *Journal of the Royal Statistical Society, Series B, Statistical Methodology* 77(1) (2015), 131-148.
- [129] D. Rizopoulos, Personal Communication: Extension of JM bayes to include left-censoring, Personal Communication, 2015.
- [130] J. Molitor et al., Identifying vulnerable populations through an examination of the association between multipollutant profiles and poverty, *Environmental Science and Technology* 45 (2011), 7754-7760.
- [131] S. Liverani et al., PReMiuM: An R package for profile regression mixture models using Dirichlet processes, *Journal of Statistical Software*, Forthcoming in the *Journal for Statistical Software*.
- [132] Z. He et al., Simultaneous variable selection for joint models of longitudinal and survival outcomes, *Biometrics* 71(1) (2015), 178-187.
- [133] D. Rizopoulos et al., Combining dynamic predictions from joint models for longitudinal and time-to-event data using Bayesian model averaging, *Journal of the American Statistical Association* 109(508) (2014), 1385-1397.

Supplementary Material 1

A substudy of 1210 women from the ALSWH 1921-26 cohort was conducted in 2000. An alternative measure of sleep was measured alongside the NHP sleep subscale items, the Pittsburgh Sleep Quality Index (PSQI). The PSQI is a valid reliable measure of sleep quality which can be used to discriminate between good and poor sleepers and continues to be used regularly in sleep research [1-4], including studies of older women [5]. The PSQI produces seven component scores, and one global score, and has an overall reliability coefficient (α) of 0.83 and a test-retest reliability (Pearson) of 0.85 [6]. The PSQI global score ranges from a minimum of 0 (better), to a maximum of 21 (worse). A total score of ≤ 5 is associated with good sleep quality, while scores > 5 are associated with poor sleep quality [7]. The component scores cover the domains Sleep quality, Sleep latency, Sleep duration, Habitual sleep efficiency, Sleep disturbance, Sleeping medications, and Daytime dysfunction. Each is scored from 0 (better) to 3 (worse).

In order to investigate the reliability of the four NHP sleep items used to create the latent classes in this study, the PSQI was dichotomised, and the association between the sleep items and the PSQI dichotomised global score and component scores assessed using Kendall's Tau rank correlation tests [8, 9], which is an appropriate non-parametric test for the correlation between two ordinal variables. Kendall's Tau-B test statistics were used for the relationship between the NHP sleep items and the PSQI global score, and Kendall-Stuarts Tau-C [10] for the relationship between the NHP items and the PSQI component scores. The results of these tests are displayed in Table 1 below.

The NHP sleep item 'wake in the early hours of the morning' is strongly concordant with the PSQI components 'sleep quality' (0.32, 95%CI (0.25, 0.38)), 'sleep duration' (0.38, 95%CI (0.32, 0.44)), and 'habitual sleep efficiency' (0.36, 95%CI (0.29, 0.42)), as well as with the global PSQI (0.32, 95%CI (0.25, 0.38)). NHP item 'lie awake at night' was strongly

concordant with 'sleep quality' (0.25, 95% CI (0.20, 0.30)), 'sleep latency' (0.22, 95%CI (0.17, 0.28)), and 'sleep duration' (0.23, 95%CI (0.18, 0.29)). 'Taking a long time to fall asleep at night' had strong concordance with 'sleep latency' (0.62, 95%CI (0.57, 0.68)), as well as 'sleep quality' (0.31, 95%CI (0.25, 0.37)), 'sleep duration' (0.31, 95%CI (0.25, 0.36)), 'habitual sleep efficiency' (0.31, 95%CI (0.24, 0.37)), and the global PSQI score (0.32, 95%CI (0.27, 0.36)). Finally, the NHP sleep item 'sleep badly' was strongly concordant with 'sleep quality' (0.44, 95%CI (0.38, 0.49)), 'sleep duration' (0.30, 95%CI (0.25, 0.36)), 'sleep latency' (0.283, 95%CI (0.22, 0.35)) and 'habitual sleep efficiency' (0.27, 95%CI (0.22, 0.33)).

Correlation between the global PSQI score and the total NHP sleep subscale score was also investigated using Pearson's and non-parametric Spearman's tests. Both the Pearson and Spearman's correlation test between the global PSQI and NHP sleep subscale found strong correlation between the two measures. The correlation coefficient of the Pearson's was 0.69 (95%CI (0.65, 0.72)), and for the Spearman's was 0.71 (95%CI (0.67, 0.74)). Concordance between the global PSQI and NHP scores was also assessed using Kendall-Stuarts Tau-C. There was strong concordance between the two measures, with concordance coefficient of 0.54 (95%CI (0.50, 0.57)).

The results of the concordance tests show that there is strong concordance between both the NHP items and the PSQI component scores, as well as with the global PSQI score. The NHP items and subscale score therefore capture a similar range of sleep difficulty as the PSQI, and are therefore reliable measures upon which to base the latent class sleep analysis.

Table 1. Concordance test statistics and *p*-values

NHP Sleep Item	PSQI	Tau-B/C*	95%CI	<i>p</i> -value
Do you wake in the early hoursof the morning?	SQ	0.32	(0.25,0.38)	< 0.0001
	SL	0.16	(0.09,0.23)	< 0.0001
	SD	0.38	(0.32,0.44)	< 0.0001
	HSE	0.36	(0.29,0.42)	< 0.0001
	SDi	0.15	(0.10,0.20)	< 0.0001
	SM	0.10	(0.03,0.17)	0.0046
	DD	0.10	(0.04,0.16)	0.0016
Global PSQI^	0.32	(0.25,0.38)	< 0.0001	
Do you lie awake most of the night?	SQ	0.25	(0.20,0.30)	< 0.0001
	SL	0.22	(0.17,0.28)	< 0.0001
	SD	0.23	(0.18,0.29)	< 0.0001
	HSE	0.20	(0.15,0.25)	< 0.0001
	SDi	0.11	(0.07,0.16)	< 0.0001
	SM	0.02	(-0.03,0.08)	0.3674
	DD	0.12	(0.07,0.17)	< 0.0001
Global PSQI^	0.16	(0.12,0.20)	< 0.0001	
Do you take a long time to get to sleep?	SQ	0.31	(0.25,0.37)	< 0.0001
	SL	0.62	(0.57,0.68)	< 0.0001
	SD	0.31	(0.24,0.36)	< 0.0001
	HSE	0.31	(0.24,0.37)	< 0.0001
	SDi	0.16	(0.11,0.21)	< 0.0001
	SM	0.23	(0.16,0.29)	< 0.0001
	DD	0.07	(0.01,0.13)	0.0311
Global PSQI^	0.32	(0.27,0.36)	< 0.0001	

Table 1. (Continued)

	SQ	0.44	(0.38,0.49)	< 0.0001
	SL	0.28	(0.22,0.35)	< 0.0001
	SD	0.30	(0.25,0.36)	< 0.0001
Do you sleep badly at night?	HSE	0.27	(0.22,0.33)	< 0.0001
	SDi	0.18	(0.13,0.23)	< 0.0001
	SM	0.12	(0.05,0.18)	0.0004
	DD	0.13	(0.08,0.19)	< 0.0001
	Global PSQI [^]	0.24	(0.20,0.28)	< 0.0001
NHP Sleep Subscale	Global PSQI	0.54	(0.50, .57)	< 0.0001

Abbreviations: Sleep Quality (SQ), Sleep Latency (SL), Sleep Duration (SD), Habitual Sleep Efficiency (HSE), Sleep Disturbance (SDi), Sleeping Medications (SM), Daytime Dysfunction (DD), Global Score (Global PSQI); PSQI, Pittsburgh Sleep Quality Index; NHP, Nottingham Health Profile; [^]Global PSQI dichotomised, = 5/>5.*Kendall-Stuarts Tau-C for correlation with PSQI components, Kendall's Tau-B for correlation with dichotomised global PSQI.

References

- [1] J. E. Maglione, S. Ancoli-Israel, S. A. Peters, M. L. Paudel, K. Yaffe, K. E. Ensrud et al., Subjective and objective sleep disturbance and longitudinal risk of depression in a cohort of older women, *Sleep* 37(7) (2014), 1179-1187.
- [2] T. Blackwell, K. Yaffe, A. Laffan, S. Ancoli-Israel, S. Redline, K. E. Ensrud et al., Associations of objectively and subjectively measured sleep quality with subsequent cognitive decline in older community-dwelling men: The MrOS study, *Sleep* 37(4) (2014), 655-663.
- [3] M. L. Paudel, B. C. Taylor, S. Ancoli-Israel, T. Blackwell, J. E. Maglione, K. L. Stone et al., Sleep disturbances and risk of depression in older men, *Sleep* 36(7) (2013), 1033-1040.
- [4] S. R. Patel, T. Blackwell, S. Ancoli-Israel and K. L. Stone, Sleep characteristics of self-reported long sleepers, *Sleep* 35(5) (2012), 641-648.
- [5] S. A. Beaudreau, A. P. Spira, A. Stewart, E. J. Kezirian, L. Y. Lui, K. E. Ensrud et al., Validation of the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Scale in older black and white women, *Sleep Med.* 13(1) (2012), 36-42.
- [6] D. J. Buysse, C. F. Reynolds III, T. H. Monk, S. Berman and D. J. Kupfer, The Pittsburgh Sleep Quality Index: A new instrument for Psychiatric Practice and Research, *Psychiatr Res. Rep. Am. Psychiatr Assoc.* 28 (1989), 193-213.

- [7] D. J. Buysse, C. F. Reynolds III, T. H. Monk, S. Berman and D. J. Kupfer, Pittsburgh Sleep Quality Index (PSQI), University of Pittsburgh Sleep Medicine Institute; 1989 [cited 2014 9/9/2014].
<http://www.sleep.pitt.edu/content.asp?id=1484&subid=2316>
- [8] M. Kendall, A new measure of rank correlation, *Biometrika* 30(1-2) (1938), 81-89.
- [9] A. Agresti, *Analysis of Ordinal Categorical Data*, Second Edition, New York: John Wiley & Sons; 2010.
- [10] A. Göktaş and Ö. Işci, A comparison of the most commonly used measures of association for doubly ordered square contingency tables via simulation, *Metodološki zvezki* 8(1) (2011), 17-37.

Supplementary Material 2**Table 1.** Participation, attrition, and key variables distributions across all surveys

	Survey 1 (1996)	Survey 2 (1999)	Survey 3 (2002)	Survey 4 (2005)	Survey 5 (2008)	Survey 6 (2011)
Response (N)	12432 (100%)	10434 (83.9%)	8646 (69.6%)	7158 (57.6%)	5560 (44.7%)	4055 (32.6%)
Deaths		489 (3.9%)	1127 (9.1%)	2098 (16.9%)	3330 (26.8%)	4869 (39.2%)
Other*		1509 (12.1%)	3659 (29.4%)	3176 (25.5%)	3542 (28.5%)	3508 (28.2%)
Disease Count						
0	2276 (18.3%)	2857 (27.4%)	1375 (15.9%)	960 (13.4%)	858 (15.4%)	444 (11.0%)
1-2	7097 (57.1%)	5871 (56.3%)	5098 (59.0%)	4119 (57.5%)	3255 (58.5%)	2061 (50.8%)
3+	3034 (24.4%)	1480 (14.2%)	1994 (23.1%)	1993 (27.8%)	1388 (25.0%)	1506 (37.1%)
Missing	25 (0.2%)	226 (2.2%)	179 (2.1%)	86 (1.2%)	59 (1.1%)	44 (1.1%)
BMI						
Underweight	360 (2.9%)	319 (3.1%)	261 (3.0%)	272 (3.8%)	203 (3.7%)	228 (5.6%)
Normal weight	5575 (44.8%)	4453 (42.7%)	2975 (34.4%)	3078 (43.0%)	2276 (40.9%)	1665 (41.4%)
Overweight	3698 (29.8%)	3049 (29.2%)	2002 (23.2%)	2082 (29.1%)	1462 (26.3%)	990 (24.4%)
Obese	1486 (39.8%)	1223 (11.7%)	844 (9.8%)	873 (12.2%)	670 (12.1%)	419 (10.3%)
Missing	1313 (10.6%)	1390 (13.3%)	2564 (29.7%)	853 (11.9%)	949 (17.1%)	751 (18.5%)

Table 1. (Continued)

Marital Status						
Married/de facto	6934 (55.8%)	5314 (50.9%)	3849 (44.5%)	2624 (36.7%)	1639 (29.5%)	880 (21.7%)
Separated/divorced	699 (5.6%)	505 (4.8%)	423 (4.9%)	307 (4.3%)	217 (3.9%)	163 (4.0%)
Widowed	4224 (34.0%)	4284 (41.1%)	4099 (47.4%)	3978 (55.6%)	3526 (63.4%)	2892 (71.4%)
Never married	351 (2.8%)	290 (2.8%)	238 (2.8%)	204 (2.9%)	142 (2.5%)	102 (2.5%)
Missing	224 (1.8%)	41 (0.4%)	37 (0.4%)	45 (0.6%)	36 (0.7%)	16 (0.4%)
Area of Residence						
Major City	5173 (41.6%)	4034 (38.7%)	3769 (43.6%)	3205 (44.8%)	2421 (43.5%)	1863 (46.0%)
Inner regional	4803 (38.6%)	4285 (41.1%)	3246 (37.5%)	2629 (36.7%)	2023 (36.4%)	1453 (35.9%)
Outer regional	2185 (17.6%)	1858 (17.8%)	1461 (16.9%)	1188 (16.6%)	884 (15.9%)	655 (16.2%)
Remote/Very Remote	271 (2.2%)	237 (2.3%)	168 (1.9%)	134 (1.9%)	101 (1.8%)	79 (2.0%)
Missing	0 (0%)	20 (0.2%)	2 (0.02%)	2 (0.03%)	131 (2.4%)	3 (0.1%)
Self-rated Health						
Good/Very Good/Excellent	8666 (69.7%)	7524 (72.1%)	6306 (72.9%)	4794 (67.0%)	3557 (64.0%)	2565 (63.3%)
F/P	3376 (27.2%)	2871 (27.5%)	2280 (26.4%)	2324 (32.5%)	1974 (35.5%)	1467 (36.2%)
Missing	390 (3.1%)	39 (0.4%)	60 (0.7%)	40 (0.6%)	29 (0.5%)	21 (0.5%)
NHP Sleep Subscale Items						
NHP1 – Y		4313 (41.3%)	4018 (46.5%)	4090 (57.1%)	3043 (54.7%)	2152 (53.1%)
NHP1 - N		5895 (56.5%)	4474 (51.8%)	2840 (39.7%)	2403 (43.2%)	1842 (45.6%)
NHP2 – Y		1112 (10.7%)	778 (9.0%)	1103 (15.4%)	675 (12.1%)	491 (12.1%)
NHP2 - N		9096 (87.2%)	7714 (89.2%)	5827 (81.4%)	4771 (85.8%)	3503 (86.4%)

Table 1. (Continued)

NHP3 – Y	2670 (25.6%)	2335 (27.0%)	2584 (36.1%)	1833 (33.0%)	1285 (31.7%)
NHP3 - N	7538 (972.2%)	6157 (71.2%)	4346 (60.7%)	3613 (65.0%)	2709 (66.8%)
NHP4 – Y	2147 (20.6%)	1710 (19.8%)	1977 (27.6%)	1359 (24.4%)	923 (22.8%)
NHP4 - N	8061 (77.3%)	6782 (78.4%)	4953 (69.2%)	4087 (73.5%)	3071 (75.8%)
Missing	226 (2.2%)	154 (1.8%)	228 (3.2%)	114 (2.1%)	59 (1.5%)
Sleep medication use					
NHP5 – Y	2110 (16.7%)	1561 (15.0%)	1578 (18.3%)		936 (23.1%)
NHP5 - N	10103 (81.3%)	8602 (82.4%)	6915 (80.0%)		3058 (75.5%)
Missing	219 (1.8%)	271 (2.6%)	153 (1.8%)		59 (1.5%)
Education					
None	4023 (32.4%)				
School/Higher School Certificate	6003 (48.3%)				
Trade/Diploma	1311 (10.6%)				
Higher education	431 (3.5%)				
Missing	664 (5.3%)				

Percentages expressed as proportion of respondents at each survey (N); Abbreviations: NHP1, NHP Sleep Subscale item 'Do you wake in the early hours of the morning?'; NHP2, NHP Sleep Subscale item 'Do you lie awake most of the night?'; NHP3, NHP Sleep Subscale item 'Do you take a long time to get to sleep?'; NHP4, NHP Sleep Subscale item 'Do you sleep badly a night?'; NHP5, NHP Sleep Subscale item 'Do you take sleep medication?'; Other* encompasses missing due to withdrawal, ineligibility due to frailty, nonresponse or loss to follow-up – actual number of deaths is higher, however for attrition purposes those who die after withdrawing for other reasons are counted here under 'other'.

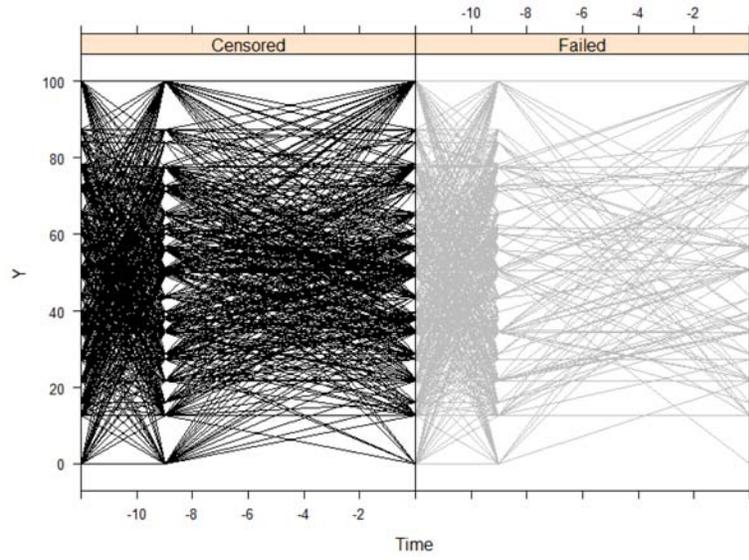
Table 2. Odds ratios (and 95%CI) multinomial logistic regressions of latent class on disease

Disease	Comorbidity adjusted			
	T	EW	TF	UT
Arthritis	2.27 (1.98,2.61)	1.64 (1.42,1.9)	1.74 (1.46,2.07)	Ref
CC	1.5 (1.41,1.59)	1.19 (1.11,1.27)	1.23 (1.13,1.32)	
Heart disease	1.8 (1.5,2.16)	1.32 (1.08,1.62)	1.5 (1.19,1.89)	Ref
CC	1.61 (1.52,1.71)	1.25 (1.18,1.34)	1.3 (1.21,1.4)	
Asthma	1.48 (1.21,1.82)	1.22 (0.98,1.52)	1.29 (1.0,1.67)	Ref
CC	1.65 (1.56,1.75)	1.26 (1.18,1.34)	1.32 (1.22,1.42)	
Bronchitis/emphysema.	1.73 (1.45,2.08)	1.33 (1.1,1.62)	1.1 (0.86,1.4)	Ref
CC	1.62 (1.53,1.72)	1.25 (1.17,1.33)	1.36 (1.26,1.46)	
Diabetes	1.45 (1.15,1.83)	0.96 (0.73,1.27)	1.23 (0.91,1.66)	Ref
CC	1.65 (1.56,1.74)	1.28 (1.2,1.36)	1.33 (1.24,1.43)	
Hypertension	1.35 (1.18,1.54)	1.15 (1.01,1.32)	1.14 (0.97,1.34)	Ref
CC	1.71 (1.6,1.82)	1.29 (1.2,1.38)	1.37 (1.26,1.48)	
Stroke	1.31 (0.96,1.79)	0.98 (0.68,1.40)	1.39 (0.95,2.04)	Ref
CC	1.64 (1.56,1.74)	1.26 (1.19,1.34)	1.3 (1.21,1.40)	

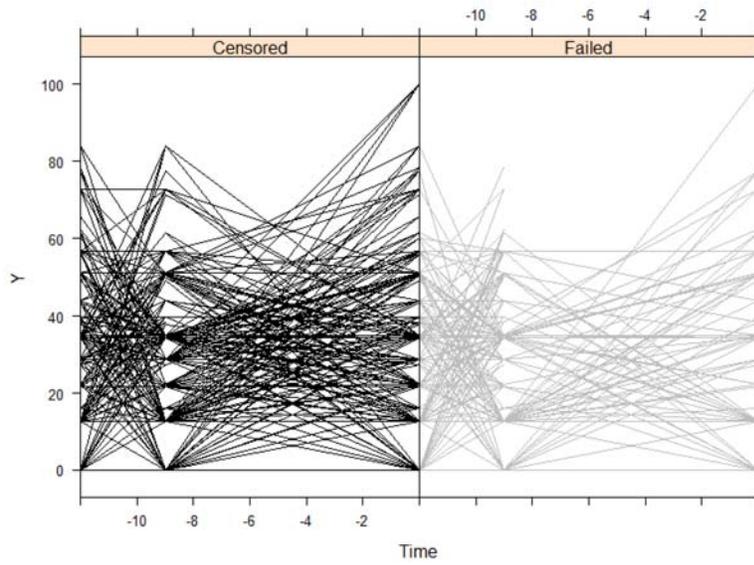
Table 2. (Continued)

Osteoporosis	1.67 (1.41,1.97)	1.25 (1.05,1.33)	1.14 (0.92,1.43)	Ref
CC	1.63 (1.54,1.73)	1.25 (1.17,1.33)	1.34 (1.24,1.44)	
Cancer	1.19 (0.93,1.52)	1.0 (0.77,1.3)	1.26 (0.94,1.69)	Ref
CC	1.65 (1.57,1.75)	1.28 (1.2,1.35)	1.32 (1.23,1.42)	

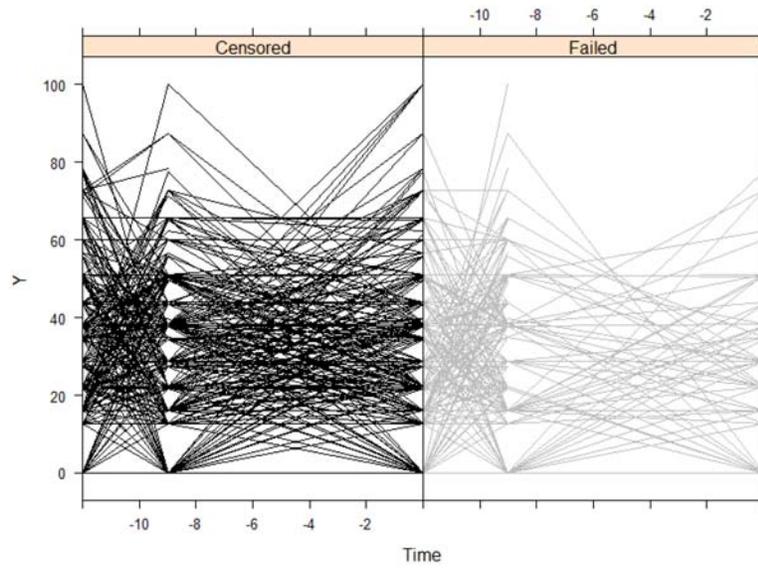
Bolded values are significant at 0.05 level of significance; Abbreviations: OR, Odds Ratio; T, troubled sleepers class; EW, early wakers class; TF, trouble falling asleep class; UT, untroubled sleepers class; CC, Comorbidity count (calculated as the sum of all other diseases).



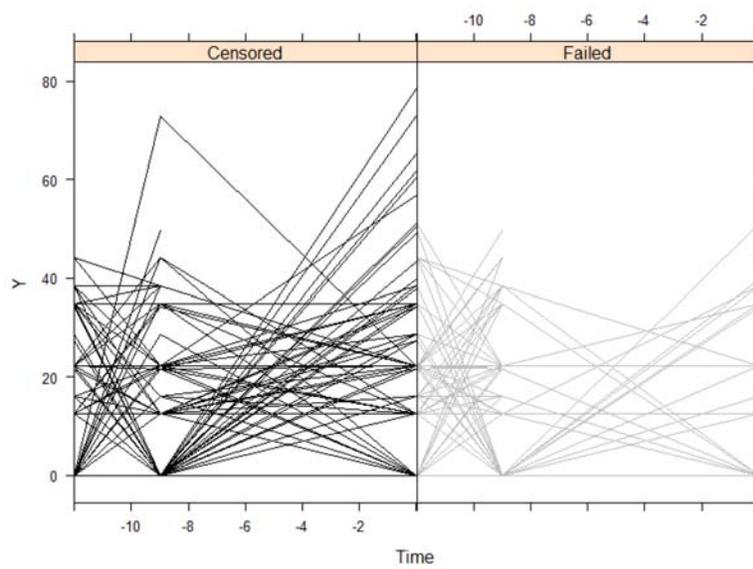
(a)



(b)



(c)



(d)

Figure 1. NHP Sleep Subscale trajectory plots for censored and uncensored women, within the (a) 'troubled sleepers', (b) 'early wakers', (c) 'trouble falling asleep', and (d) 'untroubled sleepers' classes.

In order to compare the current analysis, which incorporates NMAR dropout, with the previous analysis of this cohort (Leigh et al. [1]) which assumed MAR dropout, we compared how the sleep trajectories (as given by the NHP sleep subscale score) varied across the ‘troubled’ sleepers, ‘early wakers’, ‘trouble falling asleep’ and ‘untroubled’ sleepers classes obtained from the previous analysis (Leigh et al. [1]). Trajectory plots (Figure 1) of the longitudinal NHP sleep subscale were created using ‘joineR’ (Philipson et al. [102]), and compared censored and uncensored women, stratified across the four latent sleep classes from earlier work on this cohort (Leigh et al. [1]). It is clear from Figure 1 that the NHP sleep subscale score is generally greater and highly variable in the ‘troubled’ sleepers, slightly lower for ‘early wakers’ and in women who had ‘trouble falling asleep’, and lowest and less variable for the ‘untroubled’ sleepers class. This demonstrates that the latent classes are also reflected by the variation in scoring of the NHP sleep subscale itself.

Table 3. Dynamic survival predictions (mean, median, and 95% CIs) for subjects surviving until survey 6, model 1. Participants belonged to (i) ‘troubled sleepers’, (ii) ‘early wakers’, (iii) ‘trouble falling asleep’, and (iv) ‘untroubled sleepers’ class in previous analysis (Leigh et al. [1])

(i)

Times	Mean	Median	Lower CI	Upper CI
12.0	1.0	1.0	1.0	1.0
12.25	0.98	0.98	0.98	0.98
12.5	0.97	0.97	0.96	0.97
13.0	0.93	0.93	0.92	0.94
13.5	0.90	0.90	0.89	0.91
14.0	0.87	0.87	0.85	0.88
14.5	0.83	0.84	0.81	0.85
14.77	0.82	0.82	0.79	0.84

(ii)

Times	Mean	Median	Lower CI	Upper CI
12.0	1.0	1.0	1.0	1.0
12.25	0.99	0.99	0.98	0.99
12.5	0.97	0.97	0.97	0.97
13.0	0.94	0.94	0.94	0.95
13.5	0.91	0.91	0.91	0.92
14.0	0.88	0.88	0.87	0.90
14.5	0.86	0.86	0.84	0.87
14.77	0.84	0.84	0.82	0.86

(iii)

Times	Mean	Median	Lower CI	Upper CI
12.0	1.0	1.0	1.0	1.0
12.25	0.99	0.99	0.98	0.99
12.5	0.97	0.97	0.97	0.97
13.0	0.94	0.94	0.94	0.95
13.5	0.92	0.92	0.90	0.93
14.0	0.89	0.89	0.87	0.90
14.5	0.86	0.86	0.84	0.87
14.77	0.84	0.84	0.83	0.86

(iv)

Times	Mean	Median	Lower CI	Upper CI
12.0	1.0	1.0	1.0	1.0
12.25	0.99	0.99	0.99	0.99
12.5	0.97	0.97	0.97	0.98
13.0	0.95	0.95	0.94	0.95
13.5	0.92	0.92	0.92	0.93
14.0	0.90	0.90	0.89	0.91
14.5	0.87	0.87	0.86	0.89
14.77	0.86	0.86	0.85	0.88

Table 4. Dynamic survival predictions (mean, median, and 95% CIs) for subjects surviving until survey 3, model 1. Participants belonged to (i) ‘troubled sleepers’, (ii) ‘early wakers’, (iii) ‘trouble falling asleep’, and (iv) ‘untroubled sleepers’ class in previous analysis (Leigh et al. [1])

(i)

Times	Mean	Median	Lower CI	Upper CI
3.0	1.0	1.0	1.0	1.0
3.5	0.99	0.99	0.99	0.99
5.0	0.96	0.96	0.95	0.96
7	0.92	0.92	0.90	0.92
9	0.84	0.84	0.81	0.86
11	0.77	0.77	0.73	0.79
13	0.66	0.67	0.61	0.70
14.77	0.58	0.58	0.52	0.63

(ii)

Times	Mean	Median	Lower CI	Upper CI
3.0	1.0	1.0	1.0	1.0
3.5	0.99	0.99	0.99	0.99
5.0	0.97	0.97	0.96	0.97
7	0.93	0.93	0.93	0.94
9	0.87	0.87	0.86	0.89
11	0.81	0.81	0.79	0.83
13	0.72	0.72	0.69	0.75
14.77	0.64	0.64	0.60	0.68

(iii)

Times	Mean	Median	Lower CI	Upper CI
3.0	1.0	1.0	1.0	1.0
3.5	0.99	0.99	0.99	0.99
5.0	0.97	0.97	0.93	0.97
7	0.93	0.93	0.92	0.94
9	0.87	0.87	0.85	0.88
11	0.80	0.80	0.78	0.82
13	0.71	0.71	0.67	0.73
14.77	0.63	0.63	0.59	0.66

(iv)

Times	Mean	Median	Lower CI	Upper CI
3.0	1.0	1.0	1.0	1.0
3.5	0.99	0.99	0.99	0.99
5.0	0.97	0.97	0.97	0.97
7	0.94	0.94	0.93	0.95
9	0.88	0.88	0.87	0.90
11	0.82	0.82	0.80	0.84
13	0.74	0.74	0.71	0.77
14.77	0.67	0.67	0.63	0.71

Table 5. Dynamic survival predictions (mean, median, and 95% CIs) for subjects surviving until survey 2, model 1. Participants belonged to (i) ‘troubled sleepers’, (ii) ‘early wakers’, (iii) ‘trouble falling asleep’, and (iv) ‘untroubled sleepers’ class in previous analysis (Leigh et al. [1])

(i)

Times	Mean	Median	Lower CI	Upper CI
0.0	1.0	1.0	1.0	1.0
3.0	0.94	0.94	0.93	0.95
5.0	0.91	0.91	0.89	0.92
7.0	0.87	0.87	0.84	0.88
9.0	0.80	0.80	0.77	0.82
11.0	0.73	0.73	0.69	0.76
13.0	0.64	0.64	0.58	0.67
14.77	0.56	0.56	0.50	0.60

(ii)

Times	Mean	Median	Lower CI	Upper CI
0.0	1.0	1.0	1.0	1.0
3.0	0.95	0.95	0.94	0.96
5.0	0.92	0.92	0.91	0.93
7.0	0.89	0.89	0.87	0.90
9.0	0.83	0.83	0.81	0.85
11.0	0.77	0.77	0.74	0.79
13.0	0.68	0.68	0.64	0.72
14.77	0.61	0.61	0.56	0.65

(iii)

Times	Mean	Median	Lower CI	Upper CI
0.0	1.0	1.0	1.0	1.0
3.0	0.95	0.95	0.95	0.96
5.0	0.92	0.92	0.91	0.93
7.0	0.89	0.89	0.88	0.90
9.0	0.83	0.83	0.81	0.85
11.0	0.77	0.77	0.74	0.80
13.0	0.69	0.68	0.65	0.72
14.77	0.61	0.61	0.57	0.66

(iv)

Times	Mean	Median	Lower CI	Upper CI
0.0	1.0	1.0	1.0	1.0
3.0	0.95	0.95	0.95	0.96
5.0	0.92	0.92	0.92	0.93
7.0	0.89	0.89	0.88	0.90
9.0	0.83	0.83	0.82	0.85
11.0	0.77	0.77	0.75	0.80
13.0	0.69	0.69	0.66	0.72
14.77	0.62	0.62	0.58	0.66

Reference

- [1] L. Leigh, I. L. Hudson and J. Byles, Sleeping difficulty, disease and mortality in older women: A latent class analysis and distal survival analysis, *Journal of Sleep Research* 24 (2015), 648-657.

doi: 10.1111/jsr.12324.

Supplementary Material 3

As an initial investigation into the possibility of multi-collinearity between the covariates in the survival sub-model, Spearman's correlations (Zar [5]) were performed for all health related covariates, which were deemed as the covariates most likely to suffer from collinearity. The results are presented in Table 1 below. Strong correlation is not exhibited for any pair, however many display moderate correlations with more than one other covariate.

Table 1.

Variable 1	Variable 2	Spearman's Rho	<i>P</i> -value (rho ne 0)
Physical Function	Mental health	0.35	< 0.00001
Physical Function	Vitality	0.62	< 0.00001
Physical Function	BMI	- 0.24	< 0.00001
Physical Function	NHP Sleep	- 0.24	< 0.00001
Physical Function	Self-rated health	0.50	< 0.00001
Physical Function	Disease	- 0.37	< 0.00001
Mental Health	Vitality	0.59	< 0.00001
Mental Health	BMI	- 0.001	0.9226
Mental Health	NHP Sleep	- 0.34	< 0.00001
Mental Health	Self-rated health	0.36	< 0.00001
Mental Health	Disease	- 0.20	< 0.00001
Vitality	BMI	- 0.12	< 0.00001
Vitality	NHP Sleep	- 0.34	< 0.00001
Vitality	Self-rated health	0.52	< 0.00001
Vitality	Disease	- 0.31	< 0.00001
BMI	NHP Sleep	0.03	0.01978
BMI	Self-rated health	- 0.08	< 0.00001
BMI	Disease	0.14	< 0.00001
NHP sleep	Self-rated health	- 0.24	< 0.00001
NHP Sleep	Disease	0.18	< 0.00001
Self-rated health	Disease	- 0.29	< 0.00001

P-value: H_0 : spearman's rho = 0, H_A : spearman's rho \neq 0.

To further investigate the possibility of multi-collinearity between the predictors in the survival part of the joint model, VIF (variance inflation factor) tests were performed. Allison [1] states that for survival analyses, ‘one can simply do a preliminary check with a linear regression program, while specifying the event time as the dependent variable. Because multi-collinearity is all about linear relations among the covariates, it is not necessary to evaluate it within the context of a survival analysis.’ This was performed on a single imputed data set to investigate whether multi-collinearity was likely to be present. VIF tests were performed in the R package ‘car’ (Fox et al. [2]), using survival time as the outcome in a linear mixed effects model within the R package ‘JM’ (Rizopoulos [4]), and the NHP sleep subscale and all other covariates as predictors (all timevarying except for education). VIF is thought to indicate a problem for any value over ten (Kutner et al. [3]). All VIF values produced were less than 10 (see Table 1 below). Therefore, multi-collinearity may not be a cause for concern in the JM analysis.

Table 2.

Variable	VIF
Survey Time	1.185096
NHP sleep subscale score	1.173700
Physical Functioning	1.981889
Mental Health	1.591845
Vitality	2.284028
Underweight	1.053690
Overweight	1.151908
Obese	1.203247
Disease Count – 1-2	1.712128
Disease Count – 3+	2.000932
Self-rated health	1.541110
Separated/divorced	1.062193
Widowed	1.165306
Never married	1.054234
Inner regional	1.175291
Outer regional	1.167005
Remote/very remote	1.030432
SC/HSC	1.305320
Trade/Diploma	1.249721
Higher Education	1.125777
Baseline age	1.030703

References

- [1] P. D. Allison, Survival Analysis. In: G. Hancock and R. O. Mueller (editors), The Reviewer's Guide to Quantitative Methods in the Social Sciences, Routledge, New York, 2010.
- [2] J. Fox, S. Weisberg, D. Adler et al., Companion to Applied Regression, R package 'car'; Version 2.0.24, 2015.

<http://cran.r-project.org/web/packages/car/car.pdf>

- [3] M. H. Kutner, C. J. Nachtsheim, J. Neter and W. Li, Applied Linear Statistical Models, Fifth Edition, (In), Operations and Decision Sciences, McGraw-Hill Irwin, New York, 2005.
- [4] D. Rizopoulos, Joint Modeling of Longitudinal and Survival Data, Package 'JM'; Version 1.3.0; 2014, 2015.
<http://cran.r-project.org/web/packages/JM/JM.pdf>
- [5] J. H. Zar, Spearman Rank Correlation, (In), Encyclopedia of Biostatistics 7, 2005.

