A joint survival-longitudinal modelling approach for the dynamic prediction of rehospitalization in telemonitored chronic heart failure patients.

DOI: 10.1177/1471082X13478880
Handle: http://hdl.handle.net/1942/16470

A Joint Survival-Longitudinal Modelling Approach for the Dynamic Prediction of Rehospitalization in Telemonitored Chronic Heart Failure Patients

Edmund Njeru Njagi*
Research Assistant, I-BioStat, Universiteit Hasselt, Diepenbeek, Belgium

Dimitris Rizopoulos
Professor, Department of Biostatistics, Erasmus University Medical Center, Rotterdam, the Netherlands

Geert Molenberghs
Professor, I-BioStat, Universiteit Hasselt & KU Leuven, Diepenbeek & Leuven, Belgium

Paul Dendale
Professor, Jessa Hospital, Heart Centre Hasselt, Hasselt, Belgium
& Universiteit Hasselt, Faculty of Medicine and Life Sciences, Diepenbeek, Belgium

Koen Willekens
Medical Student, Katholieke Universiteit Leuven, Faculty of Medicine, B-3000 Leuven, Belgium

Abstract

Telemonitoring in chronic heart failure involves remote monitoring, by clinicians, daily patient measurements of biomarkers such as blood pressure and heart rate. As a strategy in heart failure management, the aim is for clinicians to use these measurements to predict rehospitalization, so that intervention decisions can be made. This is important for clinical practice since heart failure patients have a very high rehospitalization rate. We present a dynamic prediction approach, based on calculating dynamically-updated patient-specific conditional survival probabilities, and their confidence intervals, from a joint model for the time-to-rehospitalization and the time-varying and possibly error-contaminated biomarker. We quantify the ability of the biomarker to discriminate between patients who are and those who are not going to get rehospitalized within a given time window of interest. This approach does not only provide a sound statistical modelling approach to the substantive problem, a problem which to the best of our knowledge has not previously been addressed using a statistical modelling approach, it also provides clinicians with a valuable additional tool on which to base their intervention decisions, and thus provides immense contribution to heart failure management.

Some Keywords: Area under the receiver operating characteristic curve (AUC); Dynamic discriminative index; Dynamic prediction; Joint modelling.

*Address for correspondence: I-BioStat, Universiteit Hasselt, Agoralaan 1, B-3590 Diepenbeek, Belgium. Email: edmund.njagi@uhasselt.be

1 Introduction

Joint modelling of longitudinal and time-to-event data, in response to the fact that the longitudinal outcome is usually error-contaminated and only intermittently observed, has evolved over the past few decades. The field has developed, starting from the simple so-called Last Value Carried Forward approach, to the two-stage procedures, culminating in the current shared-parameter joint modelling approaches (Verbeke *et al.*, 2010; Tsiatis and Davidian, 2004). Indeed, a lot of work has been done in related aspects, and some key references include Pawitan and Self (1993), DeGruttola and Tu (1994), Taylor, Cumberland, and Sy (1994), Faucett and Thomas (1996), Lavalley and DeGruttola (1996), Hogan and Laird (1997), Hogan and Laird (1998), Tsiatis, DeGruttola, and Wulfsohn (1995), Henderson, Diggle, and Dobson (2000), and Xu and Zeger (2001), with excellent overviews being given by Tsiatis and Davidian (2004) and Yu *et al.* (2004).

Within classical survival analysis, there has been a lot of interest on the ability of models in discriminating between patients who will and patients who will not experience the event of interest. Harrell *et al.* (1996) discussed an index, analogous to the receiver operating curve (ROC), based on comparing survival probabilities for the so-called comparable subjects. Antolini *et al.* (2005) extended the discrimination index to include time-dependent covariates. This included a time-dependent area under the curve (AUC) approach, allowing for the evaluation of the discriminative ability at any time point, as well as overall. Among other key references in this rapidly evolving field of prospective accuracy include Zheng and Heagerty (2007) and Heagerty and Zheng (2005). Within the joint modelling framework, in contrast, there has not been as much work related to discrimination. Yu *et al.* (2008) considered a Bayesian approach to individual prediction of recurrent probabilities using a joint longitudinal survival-cure model. Rizopoulos (2011) recently focused on dynamic prediction of conditional survival probabilities. By first taking into account the endogenous nature of the time-dependent biomarkers, this author discussed a Monte Carlo based procedure for computing dynamic conditional survival predictions, as well as their confidence intervals. He also presented a general definition of prediction rules, based on a more elaborate function of the longitudinal history, as well as the consideration of different threshold values. Similar in spirit to the approach of Antolini *et al.* (2005), the paper discussed the so-called dynamic discrimination index, following from the time-dependent AUCs, which offers an overall measure of the discriminative ability.

Existing methodology on dynamic prediction and discrimination is applicable in many areas of medical

research. These include HIV research, where CD4 counts as a marker for survival may be analyzed to provide conditional survival probabilities for individual patients based on their available information (Rizopoulos, 2012b, 2011). In liver cirrhosis research, prothrombin measurements may be analyzed for similar purposes (Rizopoulos, 2012b). The vast application areas extend to research in breast cancer (Antolini *et al.*, 2005). Therefore, the area of dynamic prediction and discrimination is an important one.

In this manuscript, we explore the practical question of dynamically predicting rehospitalization in telemonitored chronic heart failure (CHF) patients. Heart failure is a condition in which the heart fails to pump enough blood for the needs of the body. The body then initiates mechanisms to compensate for the heart's failure. Over time, these mechanisms may overshoot and by themselves cause problems, and this is referred to as cardiac decompensation. Because heart failure patients are known to have a very high rate of rehospitalization, which reaches 50% per year in the most severe cases, the prediction of threatening decompensation is very important in clinical practice. In the heart failure literature, Chin and Goldman (1997), Lewin *et al.* (2005), Chaudhry *et al.* (2007), Zhang *et al.* (2009), and Dendale *et al.* (2011), have addressed various aspects, including the challenges in determining a patient's rehospitalization risk.

To the best of our knowledge, a statistical modelling approach has not previously been used to predict rehospitalizations in telemonitored chronic heart failure patients. Given that these data comprise time-to-event (time-to-rehospitalization) and longitudinal outcomes (the various longitudinal biomarkers), and given that interest is on predicting rehospitalization within a given time window of interest so that intervention can be done as appropriate, the practice of heart failure management through telemonitoring would immensely benefit from dynamic prediction as developed within the joint modelling framework. In this paper, we aim to recast the framework by Rizopoulos (2012b, 2011) to propose a solution in this area. The first step is to fit a shared random effects joint model for the time-to-rehospitalization and the biomarker, with the latter as a time-varying and error-contaminated covariate (Rizopoulos, 2011; Verbeke *et al.*, 2010; Tsiatis and Davidian, 2004). Note that in this step, we take into account the fact that the biomarker may be measured with error, an aspect that is clearly ignored in the existing (non-statistical) approaches. The second step involves calculating (patient-specific) conditional survival probabilities, and their respective confidence intervals, based on the fitted model (Rizopoulos, 2011). These probabilities, which take into account the patient's available biomarker measurements, and get (dynamically) updated as more measurements

become available, would clearly provide physicians with an additional tool on which to base their intervention decisions. The third step is to quantify the discriminative ability of the biomarker (Rizopoulos, 2012b, 2011). This would aid physicians in assessing the performance of the predictions as provided by the statistical model.

1.1 The Chronic Heart Failure Data

These data originated from a study conducted in Belgium between 2008 and 2010 (Dendale *et al.*, 2011), and whose objective was to study whether follow-up of chronic heart failure (CHF) patients, by means of a telemonitoring programme, reduced mortalilty and rehospitalization rates. Systolic and diastolic blood pressure, heart rate and weight measurements were daily remotely collected from 80 patients. This was through a set of apparatuses availed to the patients at hospital discharge, and through which they not only made the measurements, but through which the measurements were remotely availed to the medical personnel. The measurements were recorded each day for a period of about 6 months. 16 patients got hospitalized during the study period. 13 of these got hospitalized once, two got hospitalized twice, and one got hospitalized thrice, hence a total of 20 hospitalizations. All four longitudinal biomarkers contained some missing values, due to a variety of reasons. In addition, the following (baseline) patient characteristics were also available: patient's sex, age, heart rhythm, cardiac muscle fiber stretch as measured through NTproBNP, patient fitness indicator (given using the New York Heart Association (NYHA) class indication), and the Left Ventricle Ejection Fraction (LVEF), which is a measure of heart performance. Our analysis of these data will have two objectives. One, we will explore prediction of patient-specific survival probabilities, given the available biomarker measurements, and assess how well the biomarkers discriminate between patients who are and patients who are not going to get hospitalized. Secondly, we will evaluate the added value of correcting for each of the baseline covariates on the discriminative ability of the biomarkers. Four patients with missing baseline cardiac muscle fiber stretch (NTproBNP) are not included in the analysis. In Figure 1, we present the Kaplan-Meier survival estimate, with the 95% confidence interval, for the time to first hospitalization.

Figure 1: *Kaplan-Meier Survival Estimate, for Time to First Hospitalization.*

2 The Joint Model

2.1 Specification, Assumptions, and Estimation

Models in which a sub-model for the time-to-event outcome is linked to that for the longitudinal outcome through a shared latent structure have received considerable focus. Distributional assumptions may be placed on the latent structure, or the same may be relaxed. Tsiatis and Davidian (2004), for instance, review both the parametric and conditional score approach. Verbeke *et al.* (2010) and Rizopoulos (2011) fit shared random effects joint models, with normality assumptions on the random effects.

For the i^{th} subject, we let T_i be the observed event time, $T_i = \min(T_i^*, C_i)$, with T_i^* the true event time, and C_i the censoring time. We assume the following hazard model:

$$h_i(t|\mathcal{M}_i(t), \boldsymbol{w}_i) = h_0(t)\exp\{\boldsymbol{\gamma}'\boldsymbol{w}_i + \alpha m_i(t)\}, \tag{2.1}$$

where $\mathcal{M}_i(t) = \{m_i(s), 0 \leq s < t\}$ is the history of the true, unobserved longitudinal process up to time point t, α the parameter representing the effect of the longitudinal process on the hazard, and $h_0(\cdot)$ the baseline risk function. In addition, $\boldsymbol{w}_i$ is a baseline covariates' vector with associated parameter vector $\boldsymbol{\gamma}$. The hazard for an event at time t therefore depends on the baseline hazard, baseline covariates, and the true value of the longitudinal covariate at that time. The risk ratio associated with unit changes in the baseline covariates is given by $\exp(\boldsymbol{\gamma})$, while the relative change in the risk for a unit change in the true value of the longitudinal covariate is $\exp(\alpha)$.

For the longitudinal outcome, we assume that the unobserved true value for the i^{th} subject at any time point t is related with the observed value $y_i(t)$ via the model:

$$y_i(t) = m_i(t) + \varepsilon_i(t) = \boldsymbol{x}_i'(t)\boldsymbol{\beta} + \boldsymbol{z}_i'(t)\boldsymbol{b}_i + \varepsilon_i(t), \tag{2.2}$$

with $\varepsilon_i(t) \sim N(0, \sigma^2)$, $\boldsymbol{\beta}$ the parameter vector, $\boldsymbol{x}_i'$ and $\boldsymbol{z}_i'$ the vectors of the fixed effects and random effects design matrices, respectively. The measurement error $\varepsilon_i(t)$ is assumed independent of the random effects $\boldsymbol{b}_i$, with $\boldsymbol{b}_i \sim N(0, D)$. The longitudinal outcome at time t therefore comprises of the true value, $m_i(t)$, contaminated by a random error term, $\varepsilon_i(t)$. The true value, as shown, is represented by a mixed model.

Note that the formulation so far assumes that the longitudinal outcome is observed at any time t. This is normally not the case, since measurements are only observed intermittently, at the time

points t_{ij}. Therefore, as discussed in Verbeke *et al.* (2010), the aim is to estimate $m_i(t)$ using the available measurements, $y_i(t_{ij}), j = 1, \ldots, n_i$, combined with model (2.2).

As discussed by the above-mentioned authors, the likelihood contribution for the i^{th} patient is:

$$f(T_i, \delta_i, \boldsymbol{y}_i; \boldsymbol{\theta}) = \int f(T_i, \delta_i | \boldsymbol{b}_i; \boldsymbol{\theta}) \left[\prod_j f\{\boldsymbol{y}_i(t_{ij}) | \boldsymbol{b}_i; \boldsymbol{\theta}\} \right] f(\boldsymbol{b}_i; \boldsymbol{\theta})\, d\boldsymbol{b}_i,$$

with $\boldsymbol{\theta}$ the parameter vector, $\boldsymbol{y}_i$ the longitudinal information for the i^{th} subject, δ_i the event indicator, and

$$f(T_i, \delta_i | \boldsymbol{b}_i; \boldsymbol{\theta}) = [h_0(T_i) \text{exp}\{\boldsymbol{\gamma}' \boldsymbol{w}_i + \alpha m_i(T_i)\}]^{\delta_i} \exp\left(- \int_0^{T_i} [h_0(t) \text{exp}\{\boldsymbol{\gamma}' \boldsymbol{w}_i + \alpha m_i(t)\}]\, dt \right).$$

Note that the definition evokes the assumption of non-informativeness of the censoring mechanism and the so-called visiting process. Tsiatis and Davidian (2004) offer an excellent overview on the usually made assumptions.

We note that different parameterizations can be considered for the true marker, $m_i(t)$, in (2.1). Instead of the true value, the true trajectory, as well as both the true value and the true trajectory, can easily be considered. In addition, in case of highly non-linear longitudinal profiles, splines, or higher order polynomials, could be considered (Rizopoulos, 2012a). As such, the model allows for a lot of flexibility.

2.2 Predicted Conditional Survival Probabilities

Based on a fitted joint model, interest here lies in calculating survival probabilities for a new subject who has provided a set of longitudinal measurements $\mathcal{Y}_i(t) = \{y_i(s); 0 \leq s \leq t\}$. Rizopoulos (2011) takes into consideration the fact that a subject providing longitudinal measurements up to time t implies they have survived up to that point; in other words, the longitudinal process provides a time-dependent endogenous covariate. Therefore, it is of relevance to consider the conditional probability that the subject survives time $u > t$, given survival up to t:

$$\pi_i(u|t) = \text{Pr}(T_i^* \geq u | T_i^* > t, \mathcal{Y}_i(t), \mathcal{D}_n; \boldsymbol{\theta}), \tag{2.3}$$

with $\mathcal{D}_n = \{T_i, \delta_i, y_i; i = 1, \ldots, n\}$ being the sample on which the model was fitted, and on which we wish to base our predictions (Rizopoulos, 2011). The problem is written in a Bayesian formulation to facilitate the calculation of standard errors. In particular, the subject-specific survival probabilities

(2.3) are decomposed into three factors. The first is the posterior of the parameters $\boldsymbol{\theta}$ given the data, which is approximated by a normal distribution with the MLEs and variance-covariance the asymptotic covariance matrix of the MLEs. The second factor is the posterior of the random effects for subject i given his oberved data $\{\mathcal{Y}_i(t), T_i^* > t\}$. We sample from this distribution using a Metropolis-Hastings algorithm with multivariate t proposals. The last factor is the ratio of the conditional survival probabilties $S_i(u|\mathcal{M}_i(u))/S_i(t|\mathcal{M}_i(t))$, which is calculated using realization of the parameters $\boldsymbol{\theta}$ from the asymptotic normal distribution and of the random effects from factor 2.

Repeating these steps a desired number of times, we obtain a Monte Carlo sample of $\pi_i(u|t)$ based on which can derive standard errors and calculate confidence intervals. For the interested reader, technical details are provided by Rizopoulos (2011).

2.3 Prospective Accuracy: Time-dependent AUCs and the Dynamic Discrimination Index

It is usually of interest to assess the predictive performance of a joint model. This is commonly done in the context of calibration, i.e., how well the model predicts what is observed, as well as in the setting of discrimination, i.e., how well the model creates a distinction between patients who experience the event of interest and those who do not. We focus on the latter approach, and use ROC methodology.

Given that conditional survival probabilities are progressively updated as more measurements become available, of medical interest is then to distinguish between patients who are going to experience the event within a given time window from those who are not. This can then aid making appropriate intervention decisions. Formally, given longitudinal measurements $\mathcal{Y}_i(t)$, interest is on the time window $(t, t + \Delta t]$. A prediction rule is then defined using $\pi_i(t + \Delta t|t)$, where, for c in $[0, 1]$, $\pi_i(t + \Delta t|t) \leq c$ is termed a success (occurrence of the event), and $\pi_i(t + \Delta t|t) > c$ a failure. Sensitivity and specificity are thus defined as

$$\begin{aligned} Se(c,t) &= \Pr\{\pi_i(t+\Delta t|t) \leq c | T_i^* \in (t+\Delta t]\}, \\ Sp(c,t) &= \Pr\{\pi_i(t+\Delta t|t) > c | T_i^* > t+\Delta t\}. \end{aligned}$$

The AUC at time t, $AUC(t)$, is obtained by varying c, as

$$AUC(t, \Delta t) = \Pr[\pi_i(t + \Delta t|t) < \pi_j(t + \Delta t|t) | \{T_i^* \in (t, t + \Delta t]\} \cap \{T_j^* > t + \Delta t\}],$$

(Rizopoulos, 2012b), where i and j represent a pair of comparable subjects (comparable pairs) (Antolini *et al.*, 2005; Harrell *et al.*, 1996). Here, the logic is that at each time point, and for a

given time period of interest, if we consider two subjects, one of whom experiences the event within the time window, and the other who survives the time window, the calculated conditional survival probability for the first patient should be lower. Predictive accuracy can be assessed at certain time points and for given time windows, using the time-dependent AUC, and the overall performance can be assessed using a summary of the AUCs, in the form of the dynamic discrimination index, $C_{dyn}^{\Delta t}$. Rizopoulos (2012b) uses the following:

$$C_{dyn}^{\Delta t} = \frac{\int AUC(t, \Delta t) \mathsf{Pr}\{\varepsilon(t, \Delta t)\}\, dt}{\int \mathsf{Pr}\{\varepsilon(t, \Delta t)\}\, dt}, \tag{2.4}$$

where

$$\varepsilon(t, \Delta t) = [\{T_i^* \in (t, t + \Delta t]\} \cap \{T_j^* > t + \Delta t\}].$$

$\mathsf{Pr}\{\varepsilon(t, \Delta t)\}$ is the probability that a random pair of subjects is comparable at time t. Technical details regarding the estimation of these quantities are provided in Rizopoulos (2011). It should be mentioned that as with Harrell's C-index, the dynamic discrimination index (DDI) does not fully account for censoring; nonetheless, just as the Harrell's C-index has found routine use, we believe that this DDI has a place in practice.

3 Analysis of Chronic Heart Failure Data

3.1 Model Formulation

For the time to first hospitalization and for each of the longitudinal markers separately, the joint model consisting of (2.1) and (2.2) was considered. At a first step, baseline covariates were not included into the joint model. In this case, the linear predictor of the survival sub-model only contained the effect of the biomarker, while the fixed-effects structure of the longitudinal sub-model only included the linear time evolution. We will forthwith refer to this as the first step model. At a second step, each of the baseline covariates were considered in turn. In this case, the concerned baseline covariate was not only included in the survival but also in the longitudinal sub-model. This will be referred to as the second step model. The random-effects structure in longitudinal sub-model, and hence the shared random-effects structure, comprised of a random intercept. The Weibull baseline risk function was assumed for $h_0(t)$. For convenience, (2.1) was re-parameterized as:

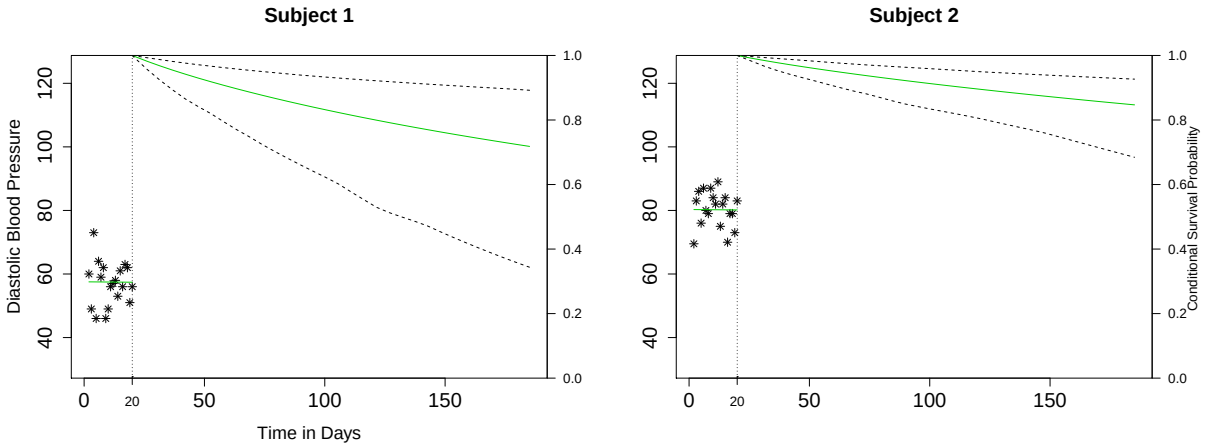
$$h_i(t|\mathcal{M}_i(t), w_i) = \rho\, t^{\rho-1} \exp\{\gamma_0 + \boldsymbol{\gamma}' \boldsymbol{w}_i + \alpha m_i(t)\}, \tag{3.1}$$

with ρ the shape parameter, and scale parameter $\exp(\gamma_0)$.

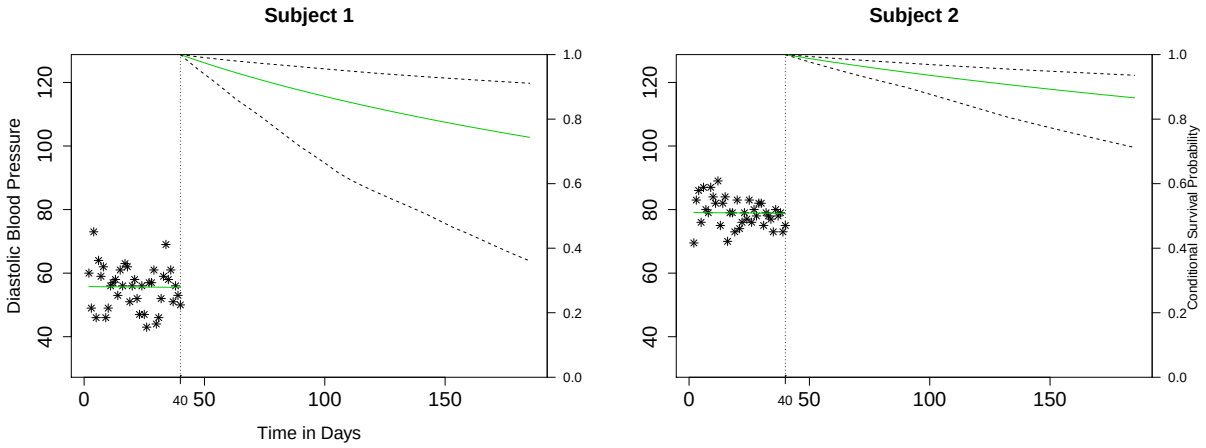
3.2 Diastolic Blood Pressure

Given the model fit, we now consider dynamic prediction of conditional survival probabilities. For illustrative purposes, we will consider two patients, who provided diastolic blood pressure measurements in a similar pattern for the first 100 days, where 'pattern' refers to missingness. For each of these patients, we consider their measurements during the first 20, 40, 80, and 100 days and, given that they had survived up to each of these time points, we compute the predicted conditional survival probabilities at each of the remaining time points until the study end. Two hundred Monte Carlo samples are generated and their median considered. We plot these probabilities, for both patients, in Figure 2 (considering measurements during the first 20 and 40 days), and Figure 3 (considering measurements during the first 60 and 80 days). The scatter points appearing before the vertical dashed line represent a plot of the longitudinal measurements up to that particular time point. To the right of this line, the conditional survival probability curve is shown, with the solid line representing the median over the Monte Carlo replications, and the dashed curves representing the confidence intervals. We notice how the diastolic blood pressure profile of the patient reflects in the conditional survival probability profile. For the first patient, the diastolic blood pressure measurements show a general clustering below 60, in comparison to the second patient, where the same seems to generally remain between 60 and 90. We notice that the conditional survival probability profile for the first patient declines rapidly, in comparison to the one for the second patient. This seems logical given that diastolic blood pressure measurements below 60 may be indicative of hypotension, hence it may be expected that the first patient is less likely to survive later times, in comparison to the second.

To show more clearly how the conditional survival probabilities are updated as more measurements become available, Figure 4 shows, for the same couple of patients as above, how the conditional probabilities of surviving an extra 20, 40, 60, and 80 days are updated, with each additional 20 days of measurements, starting from day 20 to the 100^{th} day. The same number of Monte Carlo samples as above is considered. The large dots represent the estimates, as given by the median, and the lines indicate the confidence intervals. We now consider how well the model performs in terms of discriminating between subjects who are going to experience hospitalization, and those who are not. We pre-specify that AUCs will be calculated every 2 weeks (14 days), and the time windows of interest will be 2, 4, 8 and, 16 days. Hence, DDIs will be calculated for the windows of interest Δt=2, 4, 8 and 16 days. In Table 1, we present the results. From the AUCs, we notice varying degrees of discriminative ability for different time windows at different time points, from a high of 0.9552 for



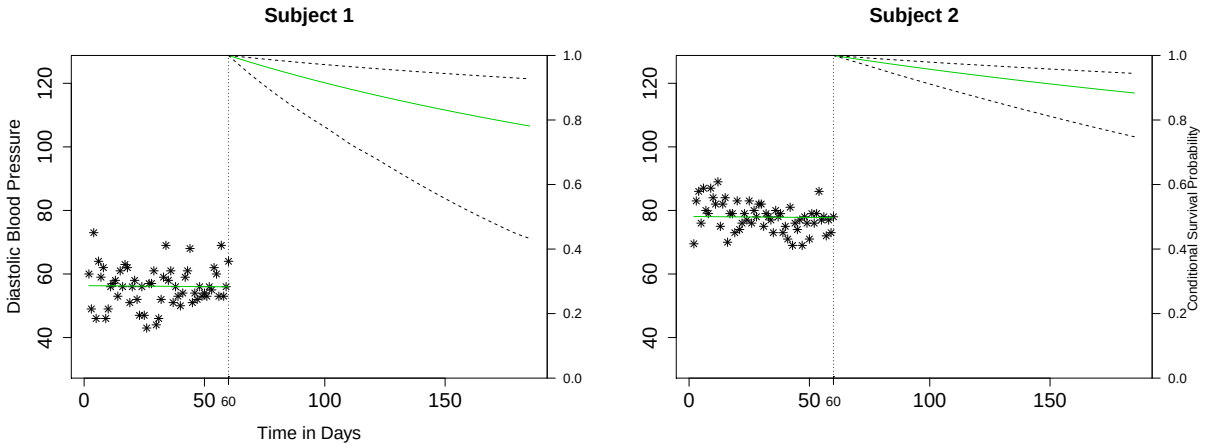
(a) Considering measurements during the first 20 days



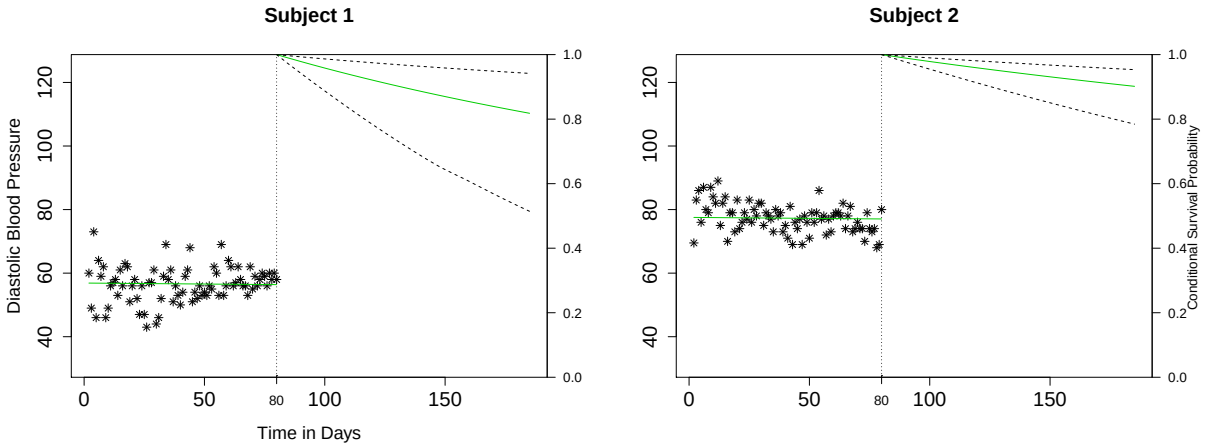
(b) Considering measurements during the first 40 days

Figure 2: *Conditional survival probabilities at each of the remaining time points until study end.*

Δt=2 at 6 weeks (42 days), to a low of 0.0517 for Δt=16 at day 154. This means, for instance, that if interest was on predicting rehospitalization within a 2-day window, and such a prediction was done at 6 weeks, then the probabilty that the model would allocate a lower conditional survival probability to a patient who was going to be rehospitalized within the next 2 days as compared to one that was not, would be 0.9552. To get a summarized measure of the discriminative ability over the follow-up period, we look at the DDIs, which provide a weighted average of the AUCs, with weights accounting for how many patients are still at risk. The indices range from 0.4875 for a time window of 2 days to 0.5814 for a time window of 8 days.



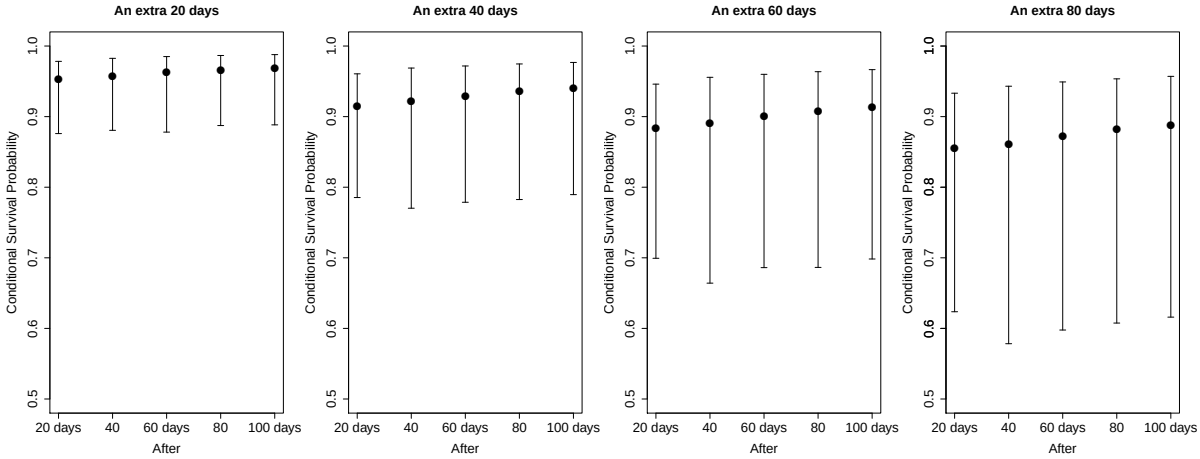
(a) Considering measurements during the first 60 days



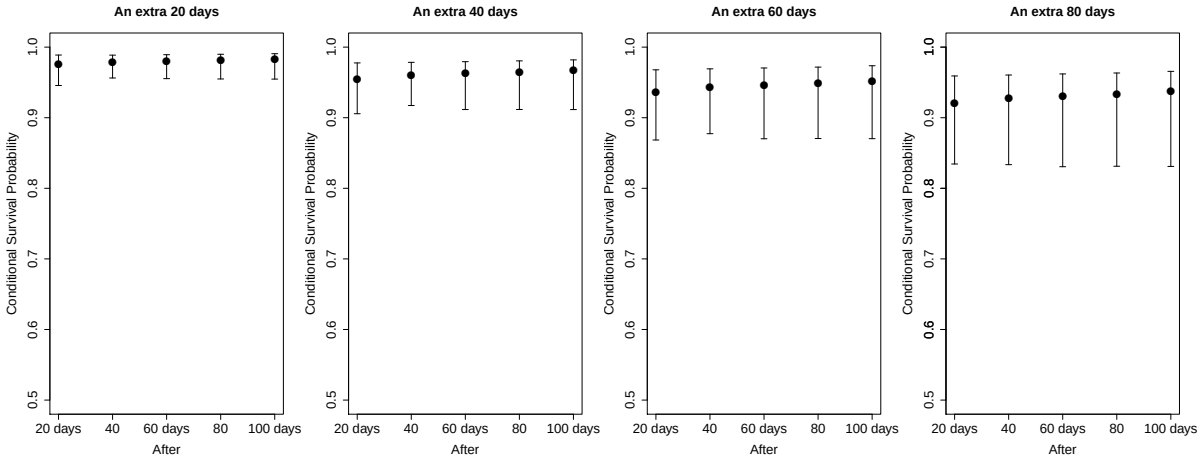
(b) Considering measurements during the first 80 days

Figure 3: *Conditional survival probabilities at each of the remaining time points until study end.*

Second step Here, interest will only be on assessing the added overall discriminative value of correcting for each of the baseline covariates. In Table 2, we provide DDIs, calculated under the same time-point and time-window specifications as above. We also include DDI results from the first step model as above. We notice that correction for NTproBNP has the highest positive effect on the overall discriminative ability under all the time windows, with the index improving to above 0.7 for an 8-day window. Patient age has the second-highest positive effect, with the index reaching above 0.6 for 8-day and 16-day windows.



(a) Patient 1



(b) Patient 2

Figure 4: *Conditional survival probabilities of surviving an extra 20, 40, 60 and 80 days, with each additional 20 days of measurement.*

3.3 Systolic Blood Pressure, Heart Rate, and Weight

In the foregoing elaborate analysis and discussion on diastolic blood pressure, we have looked at the dynamic prediction of conditional survival probabilities, how these probabilities get updated as more measurements become available, seen how the longitudinal profiles reflect in these probabilities, and assessed the discriminative ability under the first step and the second step model. Clearly, such a detailed procedure can be repeated for each of the three remaining markers (systolic blood pressure, heart rate and weight). For each of these three, however, we will focus on assessing the discriminative ability. We will consider the DDIs computed following the same time-point and

Table 1: *Diastolic Blood Pressure, First Step Model. AUCs and DDIs.*

Time window Δt	Time point t	AUC(t)	DDI
2	14	0.6944	0.4875
	28	0.7429	
	42	0.9552	
	84	0.3770	
	168	0.0862	
4	14	0.6944	0.4949
	28	0.7714	
	42	0.8955	
	84	0.3770	
	168	0.0862	
8	14	0.7500	0.5814
	28	0.8696	
	42	0.4615	
	56	0.7619	
	84	0.3934	
	168	0.0862	
16	14	0.7810	0.5745
	28	0.6493	
	42	0.4688	
	56	0.7619	
	70	0.4590	
	84	0.7119	
	154	0.0517	
	168	0.2692	

time-window specifications as above.

Systolic Blood Pressure For numerical stability, various variables were transformed or rescaled. In the first step model, systolic blood pressure measurements were rescaled to within unit magnitude, by

Table 2: *Diastolic Blood Pressure, DDIs for different time windows.*

		Second Step. Covariate controlled for:					
Δt	First Step	NTproBNP	Heart Rhythm	NYHA	Sex	LVEF	Age
2	0.4875	0.6054	0.4842	0.5087	0.5156	0.4725	0.5425
4	0.4949	0.6167	0.4956	0.5274	0.5267	0.4838	0.5535
8	0.5814	0.7206	0.5819	0.6126	0.5909	0.5702	0.6140
16	0.5745	0.6260	0.5874	0.5933	0.5967	0.5864	0.6254

dividing with the largest value. This was also done in the second step model that controlled for each of the following: NTproBNP, NYHA, sex, LVEF and age, with, in addition, NTproBNP values being transformed through the square root, and age values being rescaled by dividing with the minimum, in the respective models. In the second step model controlling for heart rhythm, however, systolic blood pressure was transformed through the cube root. In Table A1 in the Web Appendix, where we present the results, both for the first and the second step, we notice discriminative power of at least 0.6 for all of the time windows in the first step model. Patient age is seen to provide the highest positive impact on discriminative power.

Heart Rate Heart rate values were rescaled to within unit magnitude, by dividing with the largest value, for numerical stability. In Table 3, DDIs for the various time windows are given. The discriminative ability in the first step model was above 0.65. Controlling for NTproBNP provided the highest positive enhancement on the discriminative ability for the time windows of 2, 4 and 16, while for the 8-day window it was NYHA.

Weight Weight values were rescaled in a similar manner as above. For the second step model containing age, age values were rescaled as in the systolic blood pressure case. As shown in Table A2 in the Web Appendix, the first step model showed discriminative indices of between 0.3877 and 0.5020, while patient age had the highest positive effect on the discriminative ability.

3.4 Overall Findings

Diastolic blood pressure measurements on their own showed generally poor discriminative ability, with patient's cardiac muscle fiber stretch and patient age being seen to provide, in that order, the best

enhancement to this ability. Systolic blood pressure measurements had fairly moderate discriminative ability, and controlling for patient age provided the best impact. Heart rate measurements showed moderate to good discriminative power, with the DDI in the first step model for an 8-day time window being above 0.7. In this case, controlling for patient's cardiac muscle fiber stretch had the best impact. Patient's age provided the best enhancement to the discriminative ability of the longitudinal weight measurements. Finally, in our particular case of sensitivity analysis, consideration of a more elaborate random effects structure comprising of both random intercepts and slopes, as opposed to intercepts alone, was seen to produce noticeable improvement in discriminative ability.

4 Sensitivity Analysis

Considerable flexibility exists when specifying the joint model. The hazard for an event can be taken to depend on the true value of the biomarker, its true trajectory, or both, or even on its previous values. The true biomarker can be estimated using elaborate fixed- and random-effects structures, through higher order polynomials or splines. Finally, piecewise-constant and spline representations can be used for the survival baseline, in addition to the common parametric choices.

We briefly explore a number of these options, in terms of the discriminative ability of the model, and its fit. For our illustration, we consider the first step model for heart rate, and focus on the baseline hazard function, the lag of the true value of the longitudinal covariate, and the random-effects structure. The linear predictor in the survival sub-model, (2.1), was expressed as

$$\gamma' w_i + \alpha m_i\{\max(t-k,0)\},$$

where k represents the lag. Lag values of zero and one, respectively denoting the assumption that the hazard depended on the true current value or the true previous value of the longitudinal covariate, were considered. Two forms for the baseline hazard function were considered: the Weibull and the piecewise constant baseline risk function, with the knots in the latter being equally-spaced in the percentiles of the observed event times. Both random intercepts only as well as random intercepts and slopes were considered. In Table 4, DDIs under the various combinations of assumptions are provided. In this particular case, while for a given random-effects structure we do not notice a substantial impact as assumptions regarding the lag of the true value of the longitudinal covariate and the baseline hazard function are varied, there is noticeable improvement in discriminative ability when the random effects structure is extended to include random slopes. Therefore, in this particular

case, consideration of a more elaborate random effects structure comprising of both random intercepts and slopes, as opposed to intercepts alone, produces noticeable improvement in discriminative ability.

The Akaike Information Criterion (AIC) values for the various models are also given, for assessment of model fit. In line with model selection practices, this can be used, alongside the DDI, to settle on a model. For the set of models considered, the AIC is the smallest for the Weibull baseline hazard model incorporating both random intercepts and slopes, and with dependence of the hazard on the current value of the true longitudinal marker.

5 Discussion

We have explored how dynamic prediction can assist physicians make intervention decisions in telemonitored CHF patients. The dynamically updated conditional survival probabilities, and their confidence intervals, can provide physicians with additional information on which to base such decisions. We have also explored how well each of the available biomarkers discriminate between patients who are and those who are not going to get rehospitalized. This approach does not only provide a sound statistical modelling approach to predicting rehospitalizations in telemonitored chronic heart failure patients, it also provides a practical solution in heart failure management.

In this paper, we have only addressed dynamic prediction in relation to the time to first hospitalization. There is therefore need for methodological extension to cope with dynamic prediction in the context of recurrent events . We have also analyzed each biomarker separately. It would be important to consider them jointly, taking their association structure into account. More software implementation work is needed to allow dynamic prediction and calculation of accuracy measures when multiple longitudinal biomarkers are available. The main challenge in this setting is that when marker-specific random effects are incorporated, computation will become prohibitive as the number of random effects increases.

Acknowledgment

The authors gratefully acknowledge support from IAP research Network P6/03 of the Belgian Government (Belgian Science Policy).

References

Antolini, F., Boracchi, P., Biganzoli, E. (2005). A time-dependent discrimination index for survival data. *Statistics in Medicine*, **24**, 3927–3944.

Chaudhry, S.I., Wang, Y., Concato, J., Gill, T.M. *et al.* (2007). Patterns of weight change preceding hospitalization for heart failure. *Circulation*, **116**, 1549–1554.

Chin, M.H. and Goldman, L. (1997). Correlates of Early Hospital Readmission or Death in Patients With Congestive Heart Failure. *American Journal of Cardiology*, **79**,1640–1644.

DeGruttola, V. and Tu, X.M. (1994). Modeling progression of CD4 lymphocyte count and its relationship to survival time. *Biometrics*, **50**, 1003–1014.

Dendale, P., De Keulenaer, G., Troisfontaines, P., Weytjens, C. *et al.* (2011). Effect of a telemonitoring-facilitated collaboration between general practitioner and heart failure clinic on mortality and rehospitalization rates in severe heart failure: the TEMA-HF 1 (TElemonitoring in the MAnagement of Heart Failure) study. *European Journal of Heart Failure*, **14**, 333–340.

Dickstein, K., Cohen-Solal, A., Filippatos, G., McMurray, J.J. *et al.* (2008). The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM). *European Heart Journal*, **29**, 2388–2442

Faucett, C.L. and Thomas, D.C. (1996). Simultaneously modeling censored survival data and repeatedly measured covariates: a Gibbs sampling approach. *Statistics in Medicine*, **15**, 1663–1685.

Harrell, F., Kerry, L., Mark, D. (1996). Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, **15**, 361–387.

Heagerty, P., and Zheng, Y. (2005). Survival model predictive accuracy and ROC curves. *Biometrics*, **61**, 92–105.

Henderson, R., Diggle, P., and Dobson, A. (2000). Joint modeling of longitudinal measurements and event time data. *Biostatistics*, **1**, 465–480.

Hogan, J.W. and Laird, N.M. (1998) Increasing efficiency from censored survival data by using random effects to model longitudinal covariates. *Statistical Methods in Medical Research*, **7**, 28–48.

Hogan, J.W. and Laird, N.M. (1997). Mixture models for the joint distribution of repeated measures and event times. *Statistics in Medicine*, **16**, 239–258.

Kurland, B.F. and Heagerty, P.J. (2005) Directly parameterized regression conditioning on being alive: analysis of longitudinal data truncated by death. *Biostatistics*, **6**, 241–258.

Kurland, B.F., Johnson, L.L., Egleston, B.L., and Diehr, P.H. (2009) Longitudinal data with follow-up truncated by death: Match the analysis method to research aims. *Statistical Science*, **24**, 211–222.

Lavalley, M.P. and DeGruttola, V. (1996). Models for empirical Bayes estimators of longitudinal CD4 counts. *Statistics in Medicine*, **15**, 2289–2305.

Lewin, J., Ledwidge, M., OLoughlin, C., McNally, C. *et al.* (2005). Clinical deterioration in established heart failure: What is the value of BNP and weight gain in aiding diagnosis? *European Journal of Heart Failure* **7**, 953–957.

Pawitan, Y. and Self, S. (1993). Modeling disease marker processes in AIDS. *Journal of the American Statistical Association*, **88**, 719–726.

Rizopoulos, D. (2010). JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software*, **35**(9), 1–33.

Rizopoulos, D. (2011). Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics*, **67**, 819–829.

Rizopoulos, D. (2012a). *Joint Models for Longitudinal and Time-to-Event Data*. Boca Raton: Chapman and Hall/CRC.

Rizopoulos, D. (2012b). Package "JM": R package version 1.0-0, URL http://cran.r-project.org/web/packages/JM/JM.pdf.

Taylor, J.M.G., Cumberland, W.G., and Sy, J.P. (1994). A stochastic model for analysis of longitudinal AIDS data. *Journal of the American Statistical Association*, **89**, 727–736.

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

Tsiatis, A.A., DeGruttola, V., and Wulfsohn, M.S. (1995). Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *Journal of the American Statistical Association*, **90**, 27–37.

Verbeke, G., Molenberghs, G., and Rizopoulos, D. (2010). Random effects models for longitudinal data. *Longitudinal Research with Latent Variables*. K. van Montfort, H. Oud, and Al Satorra (Eds.). New York: Springer, pp. 37–96.

Xu, J. and Zeger, S.L. (2001). Joint analysis of longitudinal data comprising repeated measures and times to events. *Applied Statistics*, **50**, 375–387.

Yu, M., Law, N., Taylor, J., and Sandler, H. (2004). Joint longitudinal-survival-cure models and their application to prostate cancer. *Statistica Sinica*, **14**, 835–832.

Yu, M., Taylor J.M.G., and Sandler H.M. (2008). Individual Prediction in Prostate Cancer Studies Using a Joint Longitudinal Survival-Cure Model. *Journal of the American Statistical Association*, **103**(481), 178–187.

Zhang, J., Goode, K.M., Cuddihy, P.E., and Cleveland, J.G. (2009). Predicting hospitalization due to worsening heart failure using daily weight measurement: analysis of the Trans- European Network-Home-Care Management System (TEN-HMS) study. *European Journal of Heart Failure*, **11**, 420–427.

Zheng, Y., Heagerty, P. (2007). Prospective Accuracy for Longitudinal Markers. *Biometrics*, **63**, 332–341.

Table 3: *Heart Rate. DDIs.*

	Time window Δt	DDI
First Step Model	2	0.6698
	4	0.6698
	8	0.7433
	16	0.6576
Second Step Model		
NTproBNP	2	0.6983
	4	0.7061
	8	0.7524
	16	0.6706
Heart Rhythm	2	0.6700
	4	0.6700
	8	0.7434
	16	0.6532
NYHA	2	0.6801
	4	0.6842
	8	0.7554
	16	0.6670
Sex	2	0.6837
	4	0.6917
	8	0.7459
	16	0.6671
LVEF	2	0.6622
	4	0.6698
	8	0.7452
	16	0.6585
Age	2	0.6644
	4	0.6685
	8	0.7341
	16	0.6459

Table 4: *Heart Rate. DDI under various assumptions.*

Random effects	Lag	Baseline hazard	Δt	DDI	AIC
Intercept	0	Weibull	2	0.6698	-26413.15
			4	0.6698	
			8	0.7433	
			16	0.6576	
		Piecewise Constant	2	0.6698	-26408.43
			4	0.6698	
			8	0.7433	
			16	0.6595	
	1	Weibull	2	0.6698	-26413.14
			4	0.6698	
			8	0.7433	
			16	0.6576	
		Piecewise Constant	2	0.6698	-26408.42
			4	0.6698	
			8	0.7433	
			16	0.6595	
Intercept and slope	0	Weibull	2	0.7427	-27995.78
			4	0.7508	
			8	0.7902	
			16	0.6878	
		Piecewise Constant	2	0.7427	-27991.15
			4	0.7508	
			8	0.7902	
			16	0.6865	
	1	Weibull	2	0.7427	-27995.72
			4	0.7508	
			8	0.7941	
			16	0.6884	
		Piecewise Constant	2	0.7427	-27991.09
			4	0.7508	
			8	0.7941	
			16	0.6865	