Applications Of Continuous-Time Survival In Latent Variable Models For The Analysis Of Oncology Randomized Clinical Trial Data Using Mplus

Bengt Muthén
Tihomir Asparouhov
Mark Boye
Michelle Hackshaw
April Naegeli *

June 25, 2009

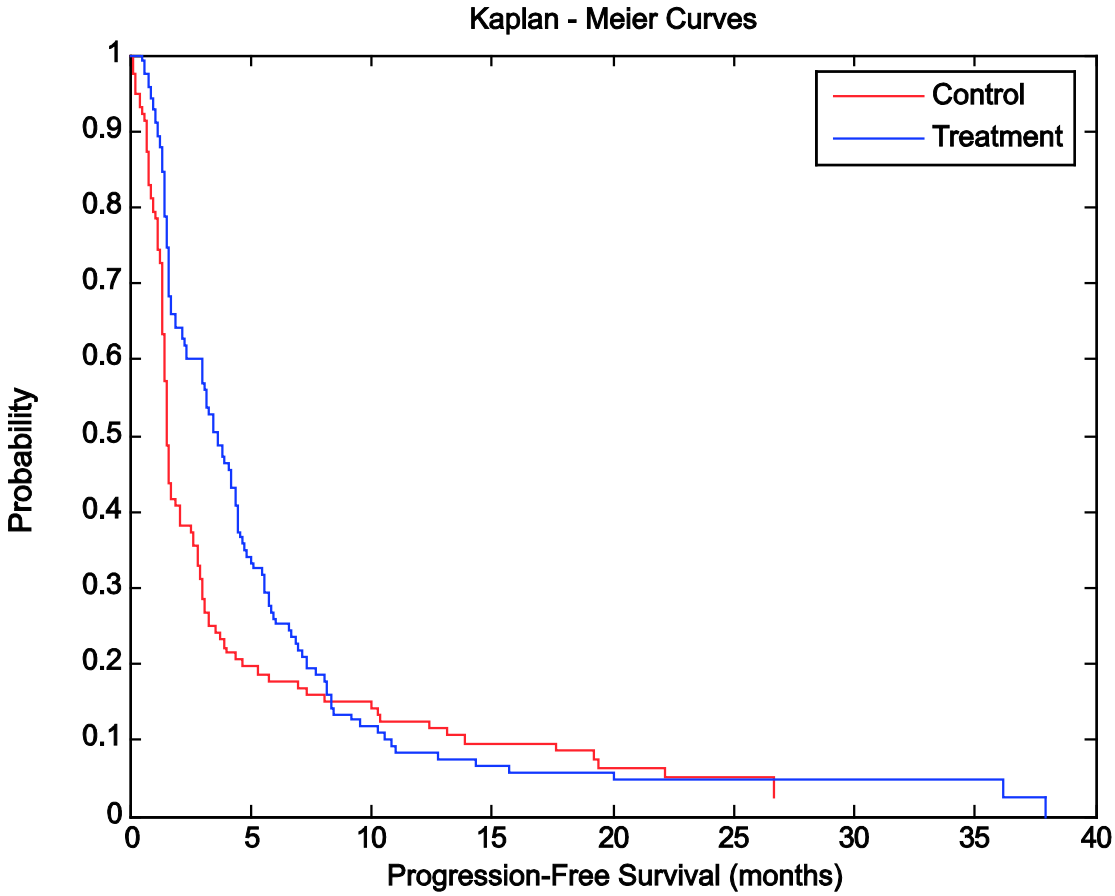
*This research was supported by Eli Lilly & Company Global Health Outcomes.

1 Introduction

This report discusses latent variable methods relevant to continuous-time survival analysis of clinical trial data. In depth analysis of data from a cancer trial is presented. The data are from a phase III study investigating the effects of second-line treatment for patients with advanced malignant pleural mesothelioma (MPM). 243 patients were randomly assigned to one of two arms: best supportive care (BSC) versus pemetrexed chemotherapy (P+BSC). Survival analysis results have been discussed in Jassem et al. (2008). The outcome to be studied in this report is progression-free survival for which Jassem et al (2008) reported a median survival time of 1.5 months for the BSC group and 3.6 months for the P+BSC group using Kaplan-Meier analysis. The Kaplan-Meier survival curves for the two groups are shown in Figure 1.

A question of great interest in cancer research is whether and how strongly patient-reported quality of life outcomes (PROs) are associated with survival. Do PROs interact with treatment in affecting survival? Do PROs measured at baseline predict survival? Do PROs have predictive utility also when controlling for the patient's cancer stage, prior treatment response, and clinician-rated performance status? Does treatment affect PROs? Does different development over time in PROs relate to differences in survival? Does PRO development worsen before disease progression? To attempt answers to these questions this report will draw on novel latent variable survival modeling techniques described in Asparouhov, Masyn, and Muthén (2006) and implemented in the Mplus software (Muthén & Muthén, 2008). Some of the latent variable survival models also draw on new software development to appear in the forthcoming Mplus version 6. A large variety of models will be explored, building up to the final model through a series of modeling steps.

The outline of the report is as follows. In Section 2 the data structure is briefly described. Section 3 presents Cox regression models, showing the need to allow for a non-proportional hazard model. Using the non-proportional hazard model, baseline information for the subjects is used to predict survival, exploring interactions with treatment arm. In Section 4, longitudinal information from one time-varying covariate at a time is used in a joint growth-survival analysis. Several alternative models for relating the two processes are compared. Section 5 discusses how the information from the PRO measures can be summarized using a variety of latent variable models. In Section 6 the best-fitting latent variable model is combined with the non-proportional hazard model to describe the effects on survival of PROs measured at baseline. Section 7 discusses latent variable survival



Median survival (months): Control (BSC) = 1.5, Treatment (P+BSC) = 3.6

Figure 1: Kaplan-Meier curves for progression-free survival

Table 1: Analysis variables (total sample size = 243)

Variable	Proportion
Tx: treatment arm	
0: best supportive care	0.494
1: pemetrexed/best supportive care	0.506
Pritreat: prior chemo response	
0: stable disease	0.416
1: progressive disease	0.300
2: partial response	0.202
3: unknown, missing	0.066
4: complete response	0.016
Dichotomized as 0+2+4	
(not treatment resistant) vs 1+3	0.617
Stage: Mesothelioma stage	
1: stage ib	0.021
2: stage ia	0.016
3: stage ii	0.074
4: stage iii	0.288
5: stage iv	0.601

modeling using time-varying information. Section 8 concludes with a summary of the substantive and statistical findings.

2 Mesothelioma trial data structure

Patients in the MPM trial were assessed at baseline (visit 0) and every 3 weeks thereafter for 24 weeks until progressive disease, death, or early discontinuation. For study design details, see Jassem et al. (2008). The set of covariates to be used in the analyses of this report is listed in Table 1 and Table 2.

Given the second-line treatment nature of the trial an important baseline back-

Table 2: Analysis variables (total sample size = 243), continued

	Mean	Std. Dev.
LCSS items		
appetite loss	31.1	28.1
fatigue	42.8	28.3
cough	21.1	26.2
dyspnea	37.1	28.7
hemoptysis	2.8	9.7
pain	32.0	29.2
overall symptoms	40.1	29.9
interference	47.6	29.3
quality of life	46.6	26.1
Karnofsky Performance Status	84.3	9.2
prior	0.6	0.5
stage	4.4	0.9
pfs	4.8	6.4
pfs censoring, BSC	0.09	0.08
pfs censoring, P+BSC	0.04	0.04

ground variable is response to prior chemotherapy. Using Cox regression, Jassem et al. (2008) found this to be the only significant treatment interaction factor. In line with the Jassem analysis, the prior experience is dichotomized as not treatment resistant versus treatment resistant (see definition in Table 1). 62% of the patients are classified as not treatment resistant. Cancer stage (stage i - stage iv) is treated as a continuous variable. 60% of the patients are at stage iv. The clinican-rated Karnofsky Performance Status (KPS) scale is shown in Table 3. A score of 100 refers to normal activity with 0 representing death.

The patient-reported quality-of-life assessment uses the Lung Cancer Symptom Scale-Mesothelioma (LCSS-Meso, or LCSS for short) shown in Table 4. The scale ranges from a score of 0 representing a good condition to a score of 100 representing a poor condition. The LCSS instrument has 9 items of which 6 refer to specific symptoms and 3 represent more global assessments (overall symptoms, interference, and quality of life). Table 2 shows that at baseline (visit 0) the mean of the item hemoptysis (coughing up blood) is very low and this item will not be included in the analysis. The cough item has the lowest mean of the remaining items at baseline. The LCSS and the KPS observations were made at the different visits up to a maximum of 11 time points. The LCSS items show a large degree of variation across time, whereas a majority of subjects obtain the same KPS score over time.

3 Survival analysis using proportional and non-proportional Cox regression modeling

3.1 Checking for non-proportionality

Survival analysis using the conventional proportional hazard model is specified as follows. Let the variable T_0 be a time-to-event variable such as death. Let I be the time when individual leaves the target cohort due to death or other types of censoring such as lost to follow up. The survival variable T and the censoring indicator δ are defined by

$$T = \min\{T_0,\, I\} \tag{1}$$

$$\delta = \begin{cases} 1 & \text{if } T_0 > I \\ 0 & \text{if } T_0 \leq I. \end{cases} \tag{2}$$

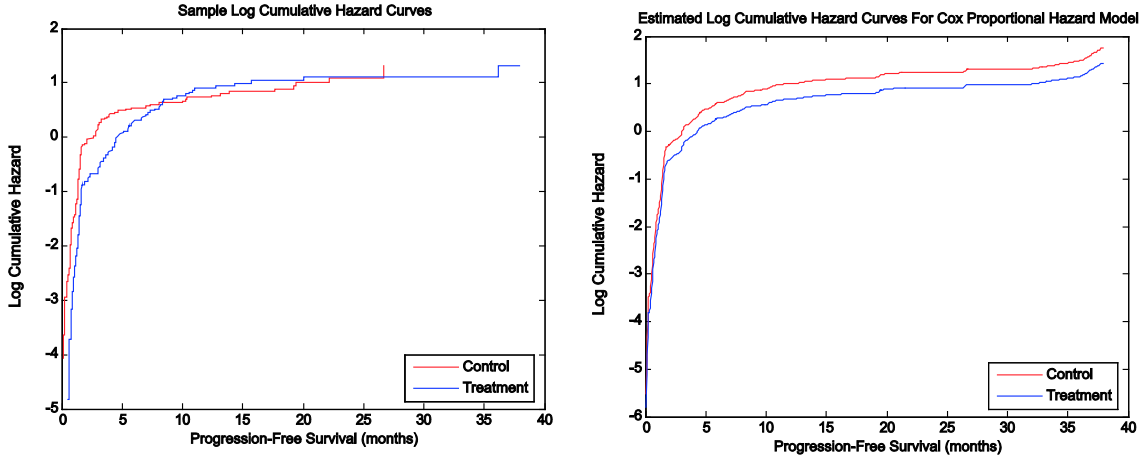
Table 3: Karnofsky Performance Status Scale

Activity Status	Point	Description
Normal Activity	100	Normal, with no complaints or evidence of disease
	90	Able to carry on normal activity but with minor signs or symptoms of disease present
	80	Normal activity but requiring effort; signs and symptoms of disease more prominent
Self-Care	70	Able to care for self, but unable to work or carry on other normal activities
	60	Able to care for most needs but requires occasional assistance
	50	Considerable assistance required, along with frequent medical care; some self-care still possible
Incapacitated	40	Disabled and requiring special care and assistance
	30	Severely disabled; hospitalization required but death from disease not imminent
	20	Extremely ill, supportive treatment, hospitalized care required
	10	Imminent death
	0	Dead

Table 4: Lung Cancer Symptom Scale-Meso (LCSS): Patient Scale

Directions: Please place a mark along each line where it would best describe the symptoms of your lung illness DURING THE PAST DAY (during the past 24 hours)	
1. How is your appetite?	
As good as it could be	As bad as it could be
2. How much fatigue do you have?	
None	As much as it could be
3. How much coughing do you have?	
None	As much as it could be
4. How much shortness of breath do you have?	
None	As much as it could be
5. How much blood do you see in your sputum?	
None	As much as it could be
6. How much pain do you have?	
None	As much as it could be
7. How bad are your symptoms from your lung illness?	
I have none	As bad as it could be
8. How much has your illness affected your ability to carry out normal activities?	
Not at all	So much that I can do nothing for myself
9. How would you rate the quality of your life today?	
Very high	Very Low

Figure 2: Log cumulative hazards for the sample and for the Cox proportional hazard model



Let X be an observed vector of covariate variables. The proportional hazard model specifies that the hazard function is proportional to the baseline hazard function,

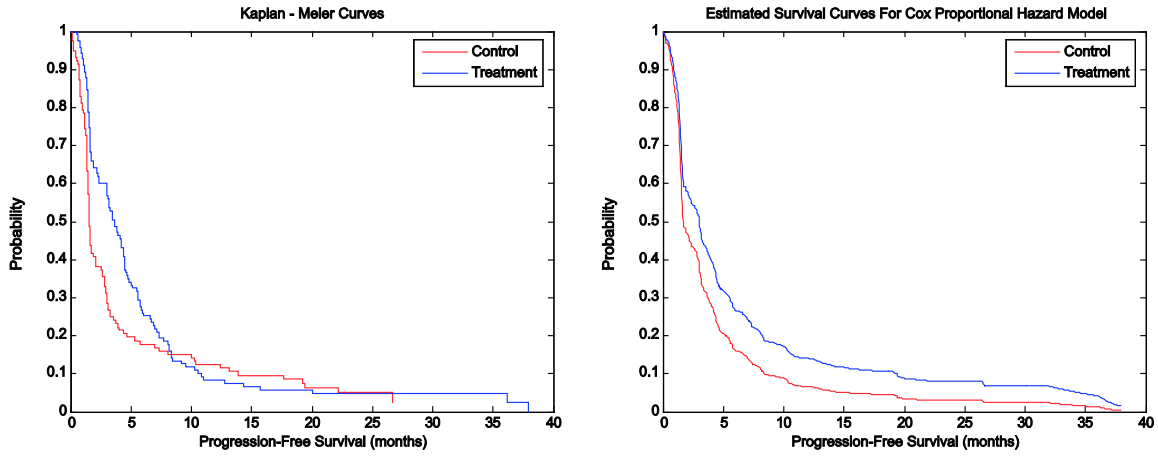
$$h(t) = \lambda(t)\, Exp(\beta X) \tag{3}$$

The Kaplan-Meier survival curves of Figure 1 provide a non-parametric representation of the survival data and are in this sense analogous to sample statistics obtained without modeling assumptions. A basic modeling aspect is if the assumption of proportionality of the hazard over time used in conventional Cox regression modeling is reasonable. In particular, it is important to know if the treatment effect is constant over time. To investigate this, it is useful to consider a plot of the sample log cumulative hazard for the treatment and control groups (cf Singer & Willett, 2003). If proportionality holds for the treatment effect this plot should show parallel curves at equal distance for all time points. The left panel of Figure 2 indicates that proportionality does not hold with the right panel showing the difference when using the estimated curve from the Cox proportional-hazard model. Figure 3 shows the differences in survival curves for Kaplan-Meier and the Cox proportional-hazard model.

3.1.1 Model alternatives

A non-proportional hazard model can be formulated by allowing an interaction between a covariate and time. Time can be broken up into different periods where

Figure 3: Survival curves for Kaplan-Meier versus Cox proportional hazard model



the hazard is different. For the MPM data time can be broken up into 9 periods corresponding to the timing of the different visits 3 weeks apart. The use of these 9 time periods is also motivated by them coinciding with the timing of the PRO measurements, so that the time-varying effects of PRO on survival can be described for each time period. Following is a set of models exploring non proportionality, starting from the conventional Cox regression assuming proportional hazards.

Let Z be a binary variable corresponding to treatment arm and X a vector of other covariate. Let Z take values 0 and 1.

Model 1

$$log(h(t|Z,Y)) = log(h_0(t)) + \alpha\, Z + \beta X$$

where h_0 is unrestricted non-parametric function. This is the standard Cox proportional hazard (CPH) model, rewriting (3) in log form. It can be estimated by regressing the survival variable on Z or by setting Z as a known latent class in mixture modeling using the new Mplus version. The parameter α is the regression coefficient for Z or the mean parameter of the survival variable.

Model 2

$$log(h(t|Z,Y)) = log(h_0(t)) + (\alpha + \gamma\, t)Z + \beta X$$

where h_0 is an unrestricted non-parametric function. The model shows an interaction between treatment arm and time. This is a model that can not be done in Mplus directly but it can be approximated by Model 3 below. The model is re-

garded as a natural extension to CPH and is a non-proportional hazard model.

Model 3

$$log(h(t|Z,Y)) = log(h_0(t)) + (\alpha + \gamma\, c[t/c])\, Z + \beta X$$

where h_0 is unrestricted non-parametric function and [ ] is the integer part function. The constant c can be any number. This model can be done in Mplus by splitting the time interval into subintervals of length c and creating separate survival variables for each individual and each interval. As $c->0$ Model 3 becomes equivalent to Model 2.

Model 4

$$log(h(t|Z,Y)) = log(h_0(t)) + \alpha_{[t/c]}\, Z + \beta X$$

where h_0 is unrestricted non-parametric function, [ ] is the integer part function and $\alpha_1, \alpha_2, ...$ are model parameters. This model is a generalization of Model 3 that relaxes the linear trend in the shift of the hazard function and is also estimated by splitting the interval into subintervals as in Model 3. Under the parameter constraints

$$\alpha_i = \alpha + \gamma\, i$$

Model 4 becomes equivalent to Model 3 and in fact that is how Model 3 is estimated, i.e., specifying Model 4 with the above parameter constraints.

Model 5

$$log(h(t|Z,Y)) = log(h_Z(t)) + \beta\, X$$

where h_1 and h_0 are both unrestricted non-parametric functions. This is the model that Mplus 5.2 will estimate by setting Z as known class. Model 5 can also be viewed as limit of Model 4 as $c->0$, i.e, Model 4 becomes equivalent to Model 5 as $c->0$.

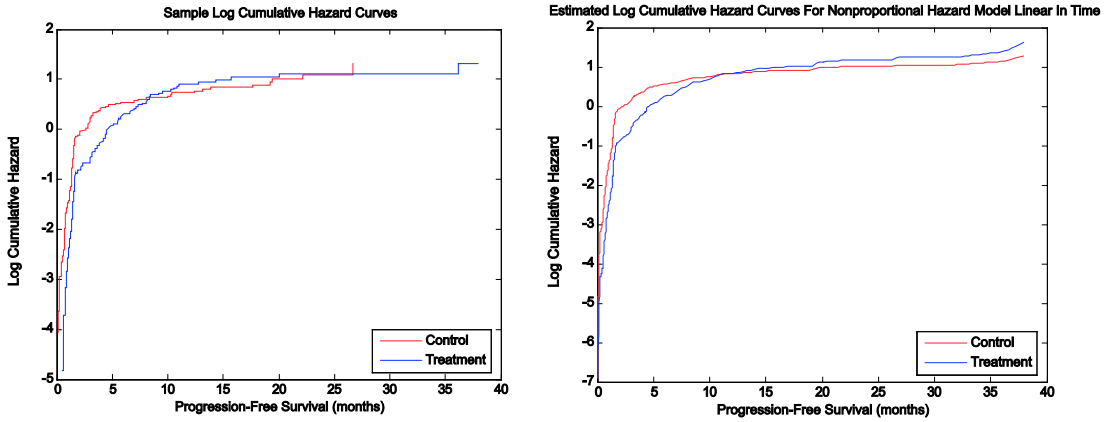
There is one complication in the above modeling. The time period is split into 8 intervals of size $c = 0.7$ corresponding to the treatment schedule. Only one more interval is included after these 8 intervals. Thus with abuse of notation $[t/c]$ is regarded to be 8 for all $t > 8c$.

The order of flexibility / generality of these models from least to most flexible is as follows.

Model 1

Model 2 (discrete) or Model 3 (continuous)

Figure 4: Log cumulative hazard curves for Kaplan-Meier versus non-proportional hazard model



Model 4 (discrete) or Model 5 (continuous)

(Note that the word "discrete" does not refer to discrete time survival modeling.)

3.1.2 Applying Model 4 to the treatment effect estimation

Log cumulative hazard functions estimated from the non-proportional, unrestricted Model 4 are shown in Figure 4 with the Kaplan-Meier curves given as comparison. The corresponding survival curves are shown in Figure 5 with the Kaplan-Meier curves as comparison. The figures show that Model 4 gives good agreement with the Kaplan-Meier curves.

Model 4 allows non-proportional hazard modeling using 9 different hazard parameters, one for each time period. Model 3 is a more parsimonious model where the log hazards are specified to be linear in time. In some settings a better fit may instead be obtained with log hazards that are linear in the log of time (Singer & Willett, 2003). These models are compared to the Cox proportional hazard model in Table 6. It is seen that the linear log hazard model has only a slightly worse loglikelihood than the unrestricted model and is preferable based on having the lowest BIC (Bayesian Information Criterion) value. This model will be used for further analyses.

Figure 5: Survival curves for Kaplan-Meier versus non-proportional hazard model

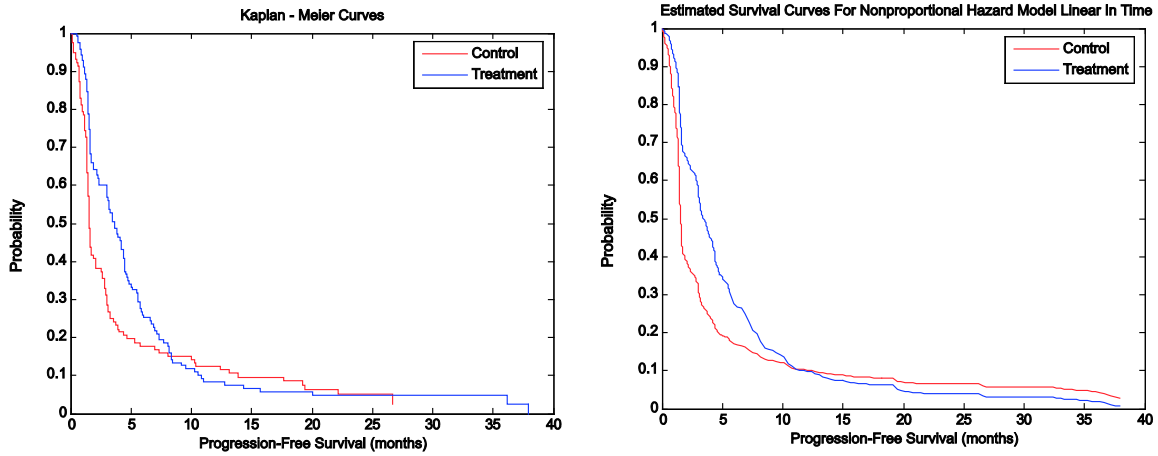


Table 6: Summary of hazard modeling of treatment effects

Model	Loglikelihood	#par.s	BIC
Proportional	-433	1	871
Unrestricted	-420	9	890
Linear	-422	2	856
Log linear	-424	2	858

3.2 Covariate effects

Model 3, the non-proportional hazard model linear in time, is used to explore the influence on survival of a set of covariates, including treatment arm, prior treatment success, cancer stage, the Karnofsky score at baseline, and the LCSS items at baseline (visit 0). The model allows for interaction between treatment arm and the covariate. The analyses use one covariate at a time. The results are listed in Table 7. The table entries describe the 5% significance of the hazard estimates at each of the 9 time periods. It is seen that for the BSC (control) group none of the covariates have a significant effect on the hazard during any of the periods. The interaction effect for the covariate prior is in line with the Jassem finding, here replicated for the non-proportional hazard model as significantly lower hazards for time periods 1 - 6. Cancer stage has a main effect in the P+BSC (treatment) group where a higher stage results in increased hazard during time periods 5-6. Most of the 9 LCSS items, and their sum, show an increased hazard for P+BSC for high LCSS values corresponding to low quality of life. In line with this, the Karnofsky score points to decreased hazard for higher scores, corresponding to higher activity levels.

3.3 Estimation

Mplus uses the maximum-likelihood estimation method for the estimation of all models described here. The estimation is based on a non-parametric baseline hazard function for the survival variables. This is accomplished by estimating a baseline hazard function as a step function that is constant between every two consecutive event times. This approach dates back to Breslow (1974) and is now referred to as the Breslow likelihood approach or the profile likelihood approach. The baseline hazard parameters are estimated as nuisance (unrestricted) parameters with this approach. The method was firmly established in the latent variable modeling area with the papers of Larsen (2004; 2005) and Asparouhov et al. (2006). Simulation studies in Asparouhov et al. (2006) show that the Mplus implementation of these methods yields consistent and asymptotically efficient results. In the case when the method is applied to the standard Cox regression model, the Mplus implementation agrees with the implementation in other statistical packages such as SAS and Stata.

There are two different ways to implement survival mixture analysis. The two methods are described in Asparouhov et al. (2006), which is implemented in Mplus 5.2, and Larsen (2004). In Larsen (2004) the baseline hazard function

Table 7: Non-proportional hazard modeling effects of each covariate

Covariate	Effect in BSC group?	Effect in P+BSC group?	Interaction (P+BSC-BSC)?
prior	no	no	1-6 (−)
stage	no	5-6 (+)	no
appetite loss	no	no	no
fatigue	no	1-5 (+)	no
cough	no	1-6 (+)	1-7 (+)
hemoptysis	no	1 (−), 4-9 (+)	1 (−), 5-9 (+)
dyspnea	no	1-5 (+)	no
pain	no	3-9 (+)	no
overall symptoms	no	1-7 (+)	3-7 (+)
interference	no	1-4 (+)	no
quality of life	no	1-7 (+)	5-7 (+)
LCSS sum	no	1-7 (+)	3-6 (+)
Karnofsky	no	1-6 (−)	1-4 (−)

Significant hazard estimates at 5% level.
Numbers refer to time period (1-9).
+: significantly higher hazard
-: significantly lower hazard

varies across classes only by a single multiplicative factor. Consider a two-class model where the baseline hazard functions in the two classes are $h_1(t)$ and $h_2(t)$. In Larsen (2004) $h_1(t)$ is estimated as a non-parametric step function while $h_2(t)$ is constrained by the following equation

$$h_2(t) = \alpha h_1(t) \tag{4}$$

where α is a parameter that is estimated. In contrast Mplus 5.2 will estimate both $h_1(t)$ and $h_2(t)$ as unconstrained non-parametric step functions. The advantage of Larsen's approach is that the class effect on the baseline can be explicitly estimated, while in the Mplus 5.2 approach the effect of the class variable is not obtained directly, simply because the two baseline function are completely unconstrained and no parameter summarizes the difference between the two baseline hazards. Essentially when using the Mplus 5.2 approach one knows that the baselines are different across classes but not how. The Mplus 5.2 approach is more flexible, but the extra flexibility was not needed for analyzing the Mesothelioma data and the Larsen (2004) approach is used for all models. The Larsen (2004) approach will be commercially available later this year with the release of Mplus Version 6.

4 Joint growth and survival analysis

The non-proportional hazard model is now expanded to include time-varying information on the LCSS items and the Karnofsky score. Random effect growth modeling is applied to a specific LCSS item. For an overview of joint random-effects growth and survival modeling, see, e.g., Diggle et al. (2008).

A key missing data issue arises in these analyses. In Table 8 and Table 9 the quality of life LCSS item is used to illustrate the strong decline in sample coverage over the visits. Data from only the first 9 visits are used here, corresponding to the regular scheduled visits. The attrition over time is determined by the progression-free survival time. Analyzing the LCSS item growth by itself draws heavily on the assumption of MAR (Little & Rubin, 2002), that is, assuming that the attrition is explained by previously observed outcomes and covariates. This assumption may not hold in which case non-ignorable missingness calls for different modeling. For an overview, see, e.g., Little (2008). Adding the survival model part to the growth model, however, alleviates the missing data problem and again makes it fall under MAR. This is the approach taken here.

Table 8: Proportion of data present

Covariance Coverage

	QOL_0	QOL_1	QOL_2	QOL_3	QOL_4	QOL_5	QOL_6	QOL_7
QOL_0	0.881							
QOL_1	0.053	0.053						
QOL_2	0.630	0.004	0.683					
QOL_3	0.519	0.004	0.531	0.564				
QOL_4	0.391	0.004	0.407	0.374	0.420			
QOL_5	0.337	0.004	0.342	0.325	0.325	0.358		
QOL_6	0.276	0.004	0.276	0.272	0.259	0.276	0.284	
QOL_7	0.230	0.004	0.230	0.222	0.214	0.226	0.218	0.235
QOL_8	0.169	0.000	0.169	0.156	0.169	0.169	0.152	0.148
QOL_9	0.123	0.000	0.119	0.119	0.119	0.123	0.115	0.115
QOL_10	0.025	0.000	0.025	0.025	0.025	0.025	0.025	0.025
QOL_11	0.021	0.000	0.021	0.021	0.021	0.021	0.021	0.021
PFS	0.881	0.053	0.683	0.564	0.420	0.358	0.284	0.235

Table 9: Proportion of data present, continued

Covariance Coverage

	QOL_8	QOL_9	QOL_10	QOL_11	PFS
QOL_8	0.177				
QOL_9	0.119	0.123			
QOL_10	0.025	0.025	0.025		
QOL_11	0.021	0.021	0.021	0.021	
PFS	0.177	0.123	0.025	0.021	1.000

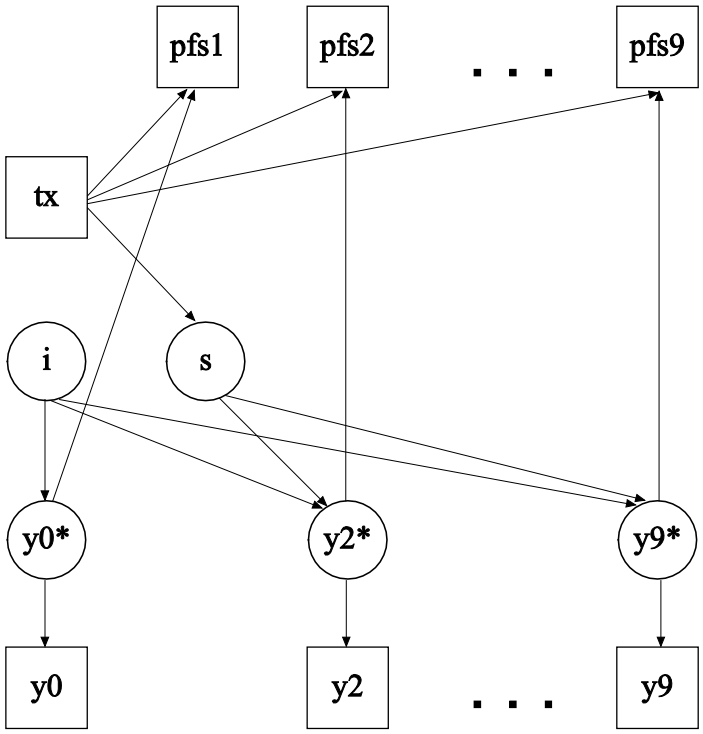


Figure 6: Predicting survival from trend (Xu-Zeger model)

For each of the time-varying variables, a random effect linear growth model is considered. Four competing models are considered for how the development over time in the variable influences survival: Predicting survival from the underlying growth trend (Xu & Zeger, 2001), predicting survival from the observed outcome, predicting survival from the random effects of the growth model, and predicting survival from latent trajectory classes (Asparouhov et al. 2006). The four models are shown in schematic form in Figure 6, Figure 7, Figure 8, and Figure 9.

The Xu and Zeger (2001) model is defined as follows. Let Y_{it} be an observed dependent variable for individual i at time t. Suppose that Y_{it} follows a linear growth model

$$Y_{it} = Y_{it}^{*} + \varepsilon_{it} \tag{5}$$

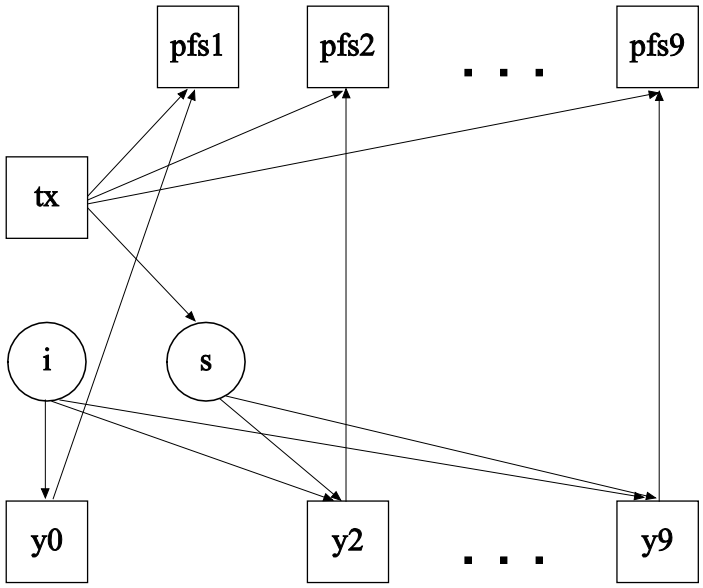


Figure 7: Predicting survival from observed outcome

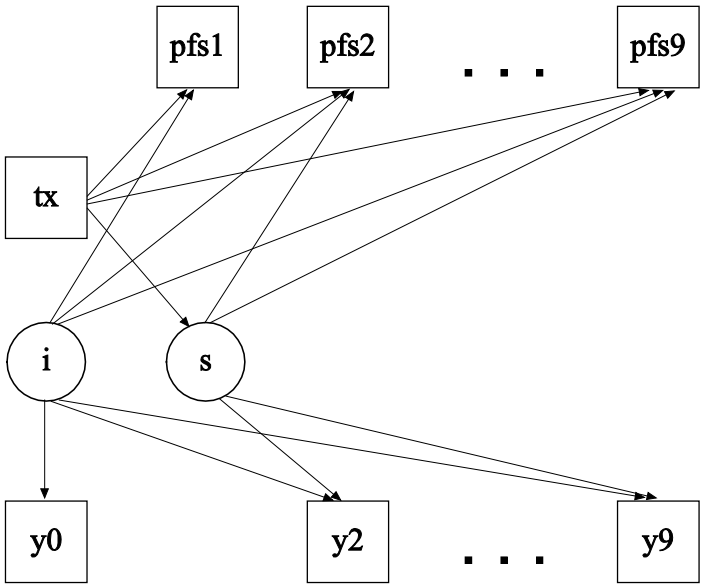


Figure 8: Predicting survival from random effects

$$Y_{it}^{*} = \alpha_i + \beta_i \, t \tag{6}$$

where α_i and β_i are normally distributed random effects.

Model 1 The Xu-Zeger model is given by

$$log(h_i(t)) = log(h_0(t)) + \gamma \, Y_{it}^{*} + \beta \, X_i. \tag{7}$$

This model can not be done in Mplus but is approximated by Model 2 below.

Model 2 For a constant c let

$$Y_{itc}^{*} = \alpha_i + \beta_i \, c[t/c] \tag{8}$$

The Mplus Xu-Zeger approximation model is given by

$$log(h_i(t)) = log(h_0(t)) + \gamma \, Y_{itc}^{*} + \beta \, X_i. \tag{9}$$

As $c->0$ Model 2 is equivalent to Model 1. Model 2 is implemented in Mplus by splitting the time interval and the survival variable into subintervals of length c.

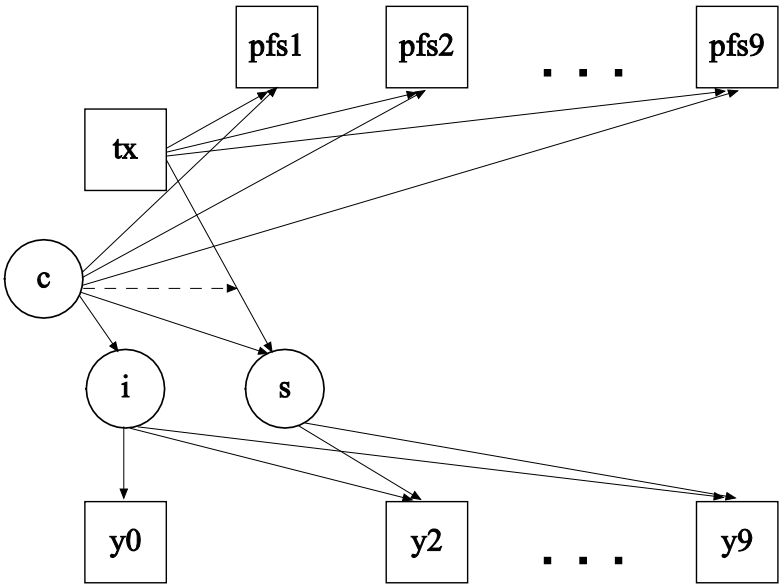


Figure 9: Predicting survival from growth mixture

Model 3 For a constant c define the alternate Xu-Zeger model which uses the actual observed values as predictors rather than their expected value.

$$log(h_i(t)) = log(h_0(t)) + \gamma\, Y_{i([t/c])} + \beta\, X_i. \tag{10}$$

Model 3 is implemented in Mplus also by splitting the time interval and the survival variable into subintervals of length c. The model is based on the assumption that the variables Y are observed at times c, 2c, 3c, ...

Description of a growth mixture model (Muthén & Asparouhov, 2008) alternative will be deferred to Section 6.

The analysis results are shown in Table 10. The first model predicts survival from the trend (referred to as Model 1, the Xu-Zeger model above); see Figure 6. The second model (referred to as Model 3 above) predicts survival from the time-varying observed LCSS outcome; see Figure 7. The third model predicts survival from the random effects α_i and β_i, referred to as i and s in Figure 8. The fourth model predicts survival from the latent trajectory classes; see Figure 9. Table 10 shows that overall the BIC differences between the models are small and no model is the best for all three items. It should be noted that for the quality of life item the 2-class growth mixture model does not have a better BIC than its 1-class counterpart (not reported), suggesting that there are not distinguishable latent trajectory classes for this LCSS item. For the other two items, there is only a small BIC advantage for the 2-class model. Because the growth mixture model needs a latent class variable with at least 2 classes to correlate the LCSS item development with survival, the growth mixture model appears to be less useful in this setting.

The last three columns refer to 5% significance of effects: Treatment effect on survival, treatment effect on the LCSS item development, and effect of LCSS item development on survival. The treatment effect on survival is significant and in the expected direction of lowering the hazard for all three items under all four models. The treatment effect on the LCSS item development is not significant for any of the items under any of the models. The point estimates are, however, of the expected sign and significance might have been obtained at a larger sample size. The effect of LCSS item development on survival is not significant under the Xu-Zeger model for any of the items but is significant under the other models. For all three items under the random effect model, it is interesting to note that survival is predicted by the trend s (β_i), so that higher trend value for a subject is associated with a higher hazard of disease progression. This finding suggests

Table 10: Survival analysis related to development in the three global LCSS items

Model	Log-Likelihood	Number of Parameters	BIC	Tx Effect on PFS	Tx Effect on LCSS	LCSS Effect on PFS
Quality of life						
Xu-Zeger	-4615	19	9334	Yes	No	No
Observed	-4610	19	9324	Yes	No	Yes
Random effects	-4611	20	9332	Yes	No	Yes (both i and s)
Growth mixture	-4598	24	9328	Yes	No	Yes
Interference						
Xu-Zeger	-4679	19	9463	Yes	No	No
Observed	-4675	19	9454	Yes	No	Yes
Random effects	-4653	20	9414	Yes	No	Yes (s only)
Growth mixture	-4674	24	9479	Yes	No	Yes
Overall symptoms						
Xu-Zeger	-4655	19	9414	Yes	No	No
Observed	-4650	19	9405	Yes	No	Yes
Random effects	-4652	20	9414	Yes	No	Yes (s only)
Growth mixture	-4639	24	9411	Yes	No	Yes

a positive answer to the question of whether disease progression is preceded by a PRO worsening.

To further study the joint growth-survival models the three key covariates of cancer stage, prior treatment success, and Karnofsky Performance status are added to the modeling. All three covariates are allowed to interact with the treatment variable in their influence on survival. Only the random effect model version is used. Table 11 shows that this model extension alters the three effect estimates reported in Table 10. The treatment no longer has a significant effect of lowering the hazard of progression-free survival, and only for the interference item is the trend *s* significantly predicting survival. Judging from the ratios of point estimates

Table 11: Survival analysis related to development in the three global LCSS items, adding covariates and using the random effect model

Tx Effect on PFS	Tx Effect on LCSS	LCSS Effect on PFS	Total Tx Effect
Quality of life			
No	No	Yes (i only)	No
Interference			
No	No	Yes (s only)	No
Overall symptoms			
No	No	No	No

to their standard errors, this finding may be due to lack of power with a relatively small sample size, where it should be noted that the Table 11 model has approximately 50% more parameters (33) as compared to the Table 10 models. Power analysis for latent variable survival models using Mplus is discussed in the Conclusions section. Note also, however, the last column of Table 11 which gives the total effect of the treatment on survival, specifically the hazard at time period 2, which is the first time period that is affected by the LCSS random effect trend s. Because survival is influenced by both treatment and the LCSS development, the direct treatment effect on survival may be "diluted". The total treatment effect is the sum of the direct effect of treatment on survival plus the product of the effect of treatment on the LCSS random effect s times the effect of s on survival. It is seen, however, that the total effect is not significant for any of the items.

A final note concerns the possibility of using information on the timing of the separate chemotherapy sessions. In version 5.21 of Mplus, modeling with time-varying covariates is accomplished by splitting the interval into subintervals and constructing survival models for the specific interval. In order to take advantage of individual-specific treatment administration one has to split the entire interval of 24 weeks into small intervals of 1 week or smaller. That in turn will produce 24

survival variables and 24x8 TX variables (8 Tx variables for each infusion multiplied by 24 for the 24 intervals that the variables are split into). That in turn will lead to a very large multivariate model that is difficult to write, navigate, compute, and debug. In addition, the deviations in the timing of the treatment administration in most cases were found to be less than a week off from the prescribed schedule. This suggests that the potential gain from this modeling attempt is quite limited. In a future Mplus version a more efficient algorithm will be available that can use time varying covariates in a more efficient manner. In addition to the above issues it became clear that searching for TX time specific effect for treatment is doomed to failure because there is not enough variation in the administration of the chemotherapy. In order to evaluate the effect of each dose TX1, TX2, ..., TX8 greater variation in the administration schedule is needed. In these data the effect of TX1 during week 4 is confounded with the effect of TX2 in its first week. Therefore it is not possible to separate these two effects.

5 Latent variable modeling of LCSS items

The LCSS items were given in Table 4 and their correlations at baseline (visit 0) are shown in Table 12 and Table 13.

The 9 LCSS items consist of 6 symptom-specific questions (appetite loss, fatigue, cough, dyspnea, hemoptysis, pain) and 3 questions of a more global nature (overall symptoms, interference, quality of life). Table 12 shows that the symptom items have moderate correlations among themselves and with the global items. The global items correlate somewhat more among themselves. All LCSS items have lower correlations with the Karnofsky Performance Status score (called kps in the table), indicating that the patient-reported outcomes carry different information than the clinician-reported score.

5.1 Alternative latent variable models

A single analysis using the combined information from all LCSS items may be more powerful than analyses based on one item at a time. The question is how to succinctly combine the information. To approach this, a variety of latent variable models will be used. Figure 10 shows a summary of 3 key models. For an overview discussion of these models, see Muthén (2008).

The top of Figure 10 shows a factor analysis model which summarizes the information in the items in terms of a continuous latent variable, labeled f as in

Table 12: Correlations

	anrx_0	ftg_0	cgh_0	dysp_0	pain_0	sx_0	intfr_0	qol_0
anrx_0	1.000							
ftg_0	0.456	1.000						
cgh_0	0.216	0.445	1.000					
dysp_0	0.325	0.649	0.518	1.000				
pain_0	0.362	0.464	0.285	0.356	1.000			
sx_0	0.411	0.572	0.320	0.426	0.537	1.000		
intfr_0	0.370	0.519	0.254	0.519	0.428	0.602	1.000	
qol_0	0.487	0.554	0.283	0.496	0.445	0.573	0.709	1.000
kps_0	-0.181	-0.281	-0.137	-0.266	-0.334	-0.365	-0.300	-0.315
tx	-0.050	-0.029	0.000	-0.076	0.029	0.027	0.037	0.024
prior	-0.022	-0.061	-0.038	0.001	-0.044	-0.043	0.022	0.001
stage	0.103	-0.005	-0.012	0.004	0.108	0.082	0.117	0.161
pfs	-0.082	-0.099	-0.078	-0.108	-0.185	-0.077	-0.047	-0.130

Table 13: Correlations, continued

	kps_0	tx	prior	stage	pfs
kps_0	1.000				
tx	-0.038	1.000			
prior	-0.024	-0.033	1.000		
stage	-0.033	0.056	-0.106	1.000	
pfs	0.110	0.095	-0.020	-0.213	1.000

factor. Different items contribute different amounts of information on this factor. For example, as shown in the left panel, compared to the other items item 4 discriminates much less between low and high factor values by having a lower slope (factor loading). More than one factor may be needed to capture the correlations among the items and this type of model can be approached using either exploratory analysis or confirmatory analysis. A strength of the factor analysis approach is that a continuous score is obtained for the subjects. A disadvantage is that the model provides no cut point on the factor(s) in order to designate an individual has having, for example, low versus high quality of life.

The middle segment of Figure 10 shows latent class analysis (also referred to as latent profile analysis when the items are continuous). In latent class analysis (LCA) the classes are defined as having different item mean profiles as shown on the left. The model assumes conditional independence of the items given the classes. The model has the advantage of providing a classification of subjects. The posterior probability of latent class membership can be estimated for each subject and the subject classified as belonging to the most likely class. The classification quality is assessed by the model entropy describing how clearly the posterior probabilities discriminate between the classes.

The bottom segment of Figure 10 shows a newer form of latent class analysis which is a hybrid LCA - factor analysis model. It is referred to as factor mixture analysis (FMA). In FMA, the strong assumption of conditional independence is relaxed in that the model allows for within-class correlation among the items. The strength of within-class correlation is determined by the factor and the item loadings on this factor. The factor may be interpreted as a severity dimension, letting subjects have different levels within a class, or may merely be used to get a better fit to the data and therefore a more trustworthy classification.

Figure 11 shows a special kind of factor analysis model. It is a confirmatory factor model with one general and one specific factor. The factors are uncorrelated. The figure shows how the model can be applied to the LCSS items, where the general factor influences all items and the specific factor influences only the symptom items. The general factor is expected to have the largest influence on the 3 global items. This model may be an improvement on a 1-factor model in that some of the items may have stronger correlations among them than a single factor can explain. For example, some of the symptom items may correlate more strongly than others. The LCSS instrument has a peculiar structure in that the overall symptom items in a sense incorporates the symptom-specific items. This direct relationship among some items may cause a rejection of the 1-factor model, but could be handled by the confirmatory model of Figure 11. The model is akin

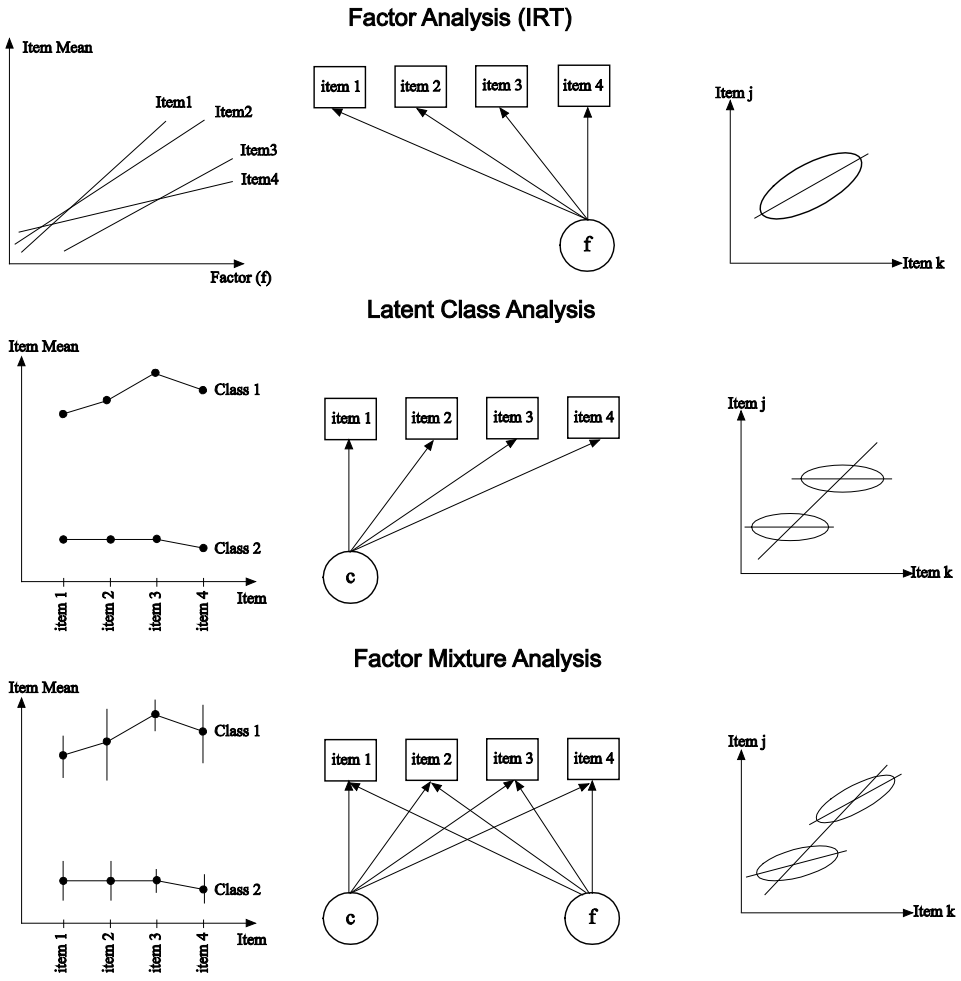


Figure 10: Overview of latent variable models

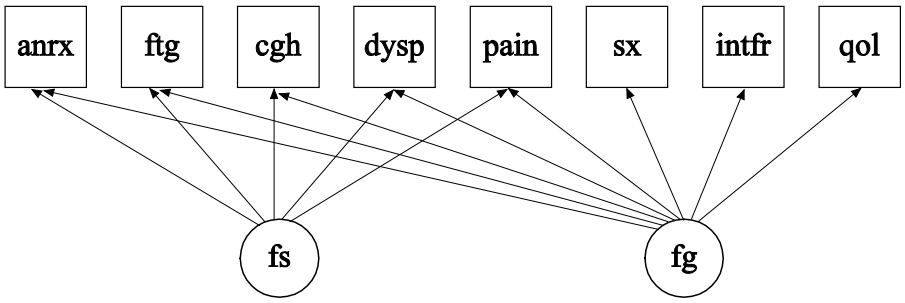


Figure 11: General-specific confirmatory factor model

to a 2-factor exploratory model where the factors are specified to be orthogonal, but also adds a specific structure of zero loadings for the 3 global items and the specific factor.

Figure 12 considers formative latent variable models and relate them to more conventional models, which are referred to as reflective. With formative indicators, the factor is influenced by the items instead of influencing the items. There is no model restriction applied to the covariances among the formative indicators. For an overview of modeling with formative indicators, see Bollen et al. (2009). Model 1 in Figure 12 shows 3 items influencing a formative factor f. Because the formative model is not identified by itself it has to be combined with other, observed variables - in this case y. For identification purposes, one slope is fixed and the residual variance is fixed at zero. In this way, the formative factor is a weighted sum of the formative indicators with relative weights being estimated. Model 2 shows how the formative factor f can be combined with a reflective factor fy. It should be noted, however, that Model 1 is equivalent to Model 3 and Model 2 is equivalent to Model 4. Model 3 is simply a regular regression model and Model 4 is a conventional structural equation model referred to as a MIMIC model.

The formative model idea may be of interest for modeling the LCSS items given the special structure of this measurement instrument referred to earlier. Figure 13 shows how the symptom-specific items can be seen to influence a quality of life factor which is in turn influencing the responses to the 3 global items. The model does not impose a structure on the covariances among the symptom-specific items, but only on their covariances with the global items and on covariances among the global items. It may be noted that an observed-variable counterpart to

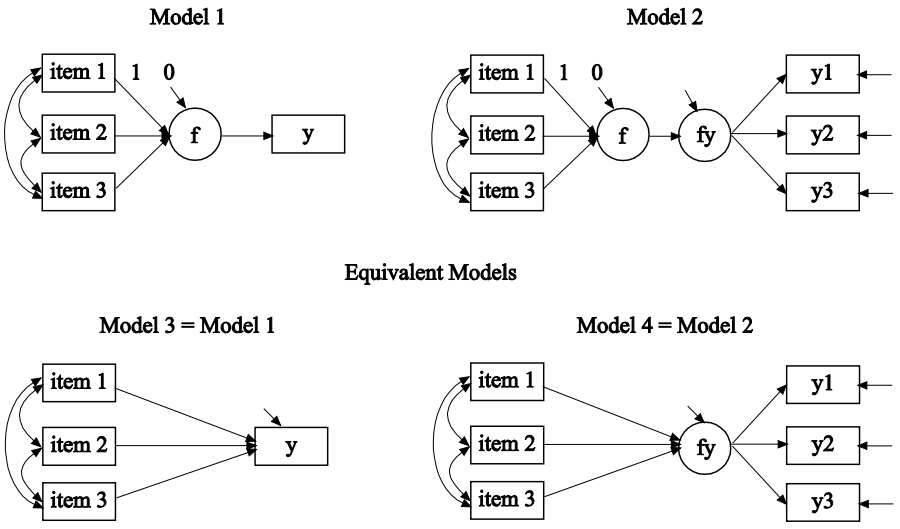


Figure 12: Formative latent variable models

this model was applied in the Hollen et al. (2004) validation of LCSS in that the 3 global items were summed and regressed on the symptom items to determine their predictive contributions.

5.2 Results of latent variable modeling for visit 0

5.3 Analysis of 8 LCSS items at visit 0

Table 14 shows the results of fitting the latent variable models discussed to the LCSS items at baseline (visit 0). Due to missing data the sample size is $n = 216$. Maximum-likelihood estimation is used. Due to the non-normality of the items, the analyses use non-normality robust χ^2 test of model fit and non-normality robust standard errors. The 1-factor model does not fit well based on a too high χ^2 and a too low CFI (comparative fit index). The exploratory 2-factor model is also rejected. The 3-factor model fits well but has only 1 item with a large loading for one of the factors and shows a Heywood case in that this item has a negative residual variance.

Model M4 is the general factor, specific factor confirmatory model of Figure 11. It fits marginally well in terms of CFI, although it is rejected by χ^2. M4 has the

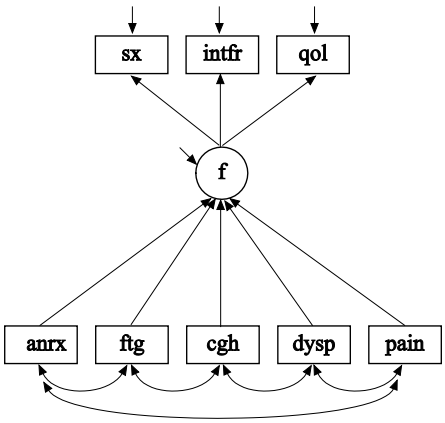


Figure 13: MIMIC model for the LCSS items

Table 14: Latent variable models for 8 LCSS items (not including hemoptysis) at visit 0, n=216

Model	Loglikelihood	#par.'s	BIC	Comments
Factor analysis				
M1: EFA 1f	-7847	24	15823	$\chi^2(20) = 68$, CFI = 0.90
M2: EFA 2f	-7817	31	15800	$\chi^2(13) = 39$, factor corr = 0.7
M3: EFA 3f	-7802	37	15803	$\chi^2(7) = 7$, Heywood
M4: CFA 1gf 1sf	-7818	29	15791	$\chi^2(15) = 30$, CFI = 0.97
M5: MIMIC 5x 3y	-7814	34	15810	$\chi^2(10) = 23$, CFI = 0.95
Latent class analysis				
M6: LCA 2c	-7911	25	15957	51% in high class
M7: LCA 3c	-7834	34	15850	
M8: LCA 4a	-7790	43	15812	
M9: LCA 5c	-7760	52	15799	
Factor mixture analysis				
M10: FMA 2c 1f	**-7772**	**33**	**15721**	**19% in high classs, entropy = 0.945**
M11: FMA 3c 1f	-7739	42	15705	14% (high), 17% (mid), 69%(low)

Table 15: Factor Solutions At Visit 0, Using 8 LCSS Items

	M1	M2		M4	
	f	f1	f2	fg	fs
anrx	0.552	0.149	0.430	0.547	0.054
ftg	0.768	0.666	0.191	0.694	0.404
cgh	0.464	0.797	-0.245	0.355	0.522
dysp	0.687	0.785	0.016	0.597	0.575
pain	0.597	0.207	0.414	0.583	0.070
sx	0.741	0.171	0.600	0.747	0.000
intfr	0.763	0.000	0.816	0.804	0.000
qol	0.783	-0.004	0.844	0.824	0.000
		1.000			
		0.715	1.000		

best (lowest) BIC among the 5 factor models. The solution is interesting as seen in Table 15, which gives the factor solutions for models M1, M2, and M4. The fg (general) factor of Model M4 is best measured by the 3 global items as expected, while the fs (specific) factor is measured best by fatigue, cough, and dyspnea. The 2-factor exploratory model M2 captures the same 3 items as measuring the factor f1, so these 3 items seem to have a common underlying feature. The M2 model shows this factor to be highly correlated with f2, whereas in model M4 the specific factor is defined as an uncorrelated, residual factor.

The MIMIC model M5 also has a borderline acceptable fit in terms of CFI. The BIC value is not as good as the multiple-factor models, which is perhaps due to the many parameters used to make the covariate part of the model unrestricted. The M5 solution is shown in Table 16. The top 3 rows show the factor loadings for the 3 global items and they are about equally strong indicators of the factor. The bottom 5 rows show the regression coefficients for the factors regressed on the 5 symptom-specific items. An interesting finding is that the cough items does not show a significant influence on the factor. As an aside, this is also found when each of the 3 global items is related to the 5 symptom-specific items in a regular linear regression. The finding is also in line with the Hollen et al. (2004) regression

Table 16: MIMIC Latent Variable Model For Visit 0

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
f BY				
sx_0	1.000	0.000	999.000	999.000
intfr_0	1.085	0.097	11.224	0.000
qol_0	0.980	0.095	10.321	0.000
f ON				
anrx_0	0.171	0.046	3.691	0.000
ftg_0	0.233	0.069	3.398	0.001
cgh_0	-0.041	0.052	-0.790	0.429
dysp_0	0.207	0.057	3.618	0.000
pain_0	0.208	0.049	4.227	0.000

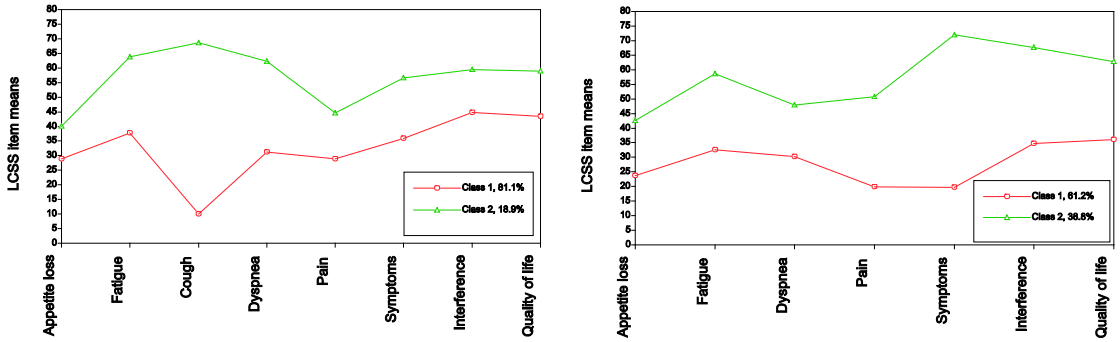
analysis using another data set. The cough item does not seem to contribute to overall quality of life. This finding will be returned to in the next step of modeling.

The latent class models M6 - M9 of Table 14 show that BIC continues to improve with increased number of latent classes, but does not show a minimum which would be pointing to the optimal number of classes. None of the latent class models has as good of a BIC value as M4.

The factor mixture models M10 and M11 give better BIC values than models M1-M9. M10 uses 2 latent classes and 1 factor, finding 19% in the high LCSS class corresponding to low quality of life. The entropy of 0.945 is excellent. The factor loadings (not reported) are highest for the 3 global items, while the cough item has a distinctly lower, albeit significant loading. The 3-class model M11 improves further on the BIC and results in 3 ordered classes.

The special behavior of the cough item seen in the MIMIC modeling is also seen in the latent class modeling. Deleting the cough item, the 7-item LCSS analyses are displayed in Table 17. It is seen that the factor models can no longer accommodate the second factor which was previously defined by fatigue, cough, and dyspnea, but inadmissible solutions with negative residual variances are ob-

Figure 14: Item mean profiles for the 8- and 7-item LCSS factor mixture models



tained. The factor mixture model M10 still has a better BIC than M1-M9. Model M11 with 3 classes has a better BIC, but one of the classes contains only 3% of the subjects.

Model M10 finds 39% in the high LCSS class as compared to the 19% of Table 14. To understand the reason for this, Figure 14 compares the item mean profiles of the 8-item latent class solution of Table 14 with those of the 7-item solution of Table 17. Figure 14 shows that the cough item dominates the creation of the latent classes in the 8-item analysis. For the 8-item solution, the mean difference across classes is the largest in standard deviation terms for the cough item. The mean differences across classes between the other items are smaller for the 8-item solution than for the 7-item solution. For example, the 8-item solution has a class separation between the means for the overall symptom item of about 2/3 of a standard deviation, while the 7-item solution has a separation of about 5/3 of a standard deviation.

The standard deviation of the factor mixture severity factor for the 7-item solution is estimated as 13. The severity factor variation can be used to consider to which extent there is overlap between subjects who are, say, 1 standard deviation below the mean in the high class and subjects who are 1 standard deviation above the mean in the low class. In the 7-item solution all but one item has no such overlap and the overall symptom item means are four such standard deviations apart, implying a good degree of class separation.

The latent variable analyses suggest that Model M10 using 7 items is the model of choice for the LCSS items. The entropy of 0.840 indicates a clear classification. The classification quality of model M10 is further investigated below.

Table 17: Latent variable models for 7 LCSS items (not including hemoptysis or cough) at visit 0, n=216

Model	Loglikelihood	#par.'s	BIC	Comments
Factor analysis				
M1: EFA 1f	-6857	21	13827	$\chi^2(14) = 63$, CFI = 0.92
M2: EFA 2f	-6836	27	13818	$\chi^2(8) = 22$, Heywood
M3: EFA 3f	-6827	32	13827	$\chi^2(3) = 31$, Heywood
M4: CFA 1gf 1sf	-6840	25	13814	$\chi^2(10) = 31$, CFI = 0.97, Heywood
M5: MIMIC 4x 3y	-6839	27	13824	$\chi^2(8) = 20$, CFI = 0.95
Latent class analysis				
M6: LCA 2c	-6915	22	13947	52% in high class
M7: LCA 3c	-6843	30	13848	
M8: LCA 4a	-6815	38	13835	
M9: LCA 5c	-6798	46	13843	
Factor mixture analysis				
M10: FMA 2c 1f	**-6823**	**29**	**13802**	**39% in high class, entropy = 0.840**
M11: FMA 3c 1f	-6796	37	13791	only 3% in one class

The classification provided by the factor mixture model M10 for the 8- and 7-item analyses is further described in Table 18. The high entropy is reflected by most likely class membership counts agreeing closely with estimated class counts. It is also reflected by the low off-diagonal values of the classification table. Here, each row corresponds to individuals most likely belonging to that class. The entries are the average posterior probability for subjects in different classes.

The class membership of the 7-item solution can also be related to other variables. It was found that there are no significant differences with respect to gender, age, cancer stage, or prior treatment success. The Karnofsky Performance Status score was significantly different in the two classes, with a mean of 86 in the low class and 81 in the high class, representing about half a standard deviation difference.

As a final step in the latent variable analysis, it is of interest to descriptively relate the latent variable modeling of visit 0 LCSS to the logarithm of progression-free survival (PFS) time and to also relate the factor modeling results to the factor mixture modeling results. Both goals are met in Figure 15 using the 7-item approach and considering the treatment group (P+BSC). For simplicity, the 1-factor model is used. First, it is seen that the estimated factor scores correlate with PFS time. The lower the factor value (the higher the quality of life), the higher on the whole is the PFS time. Second, it is seen that latent class membership correlates with PFS time in that the lower LCSS class (higher quality of life) has on the whole higher PFS times than the higher LCSS class. Third, it is seen that the factor scores and the latent classes do not have a 1-to-1 correspondence. Considering the factor score axis, it is not the case that the low LCSS class subjects are all to the left with the high class LCSS subjects all to the right. Because of this, the choice between these two latent variable models can have an impact when incorporating the visit 0 LCSS items in a latent variable survival model.

6 Predicting survival from baseline LCSS

The visit 0 LCSS factor mixture model with 2 classes and 1 factor will now be used to predict progression-free survival using the non-proportional hazard model linear in time. The 7-item version of the latent variable model is used. The key question is if the LCSS items can contribute significant information beyond that already contributed by the baseline covariates of prior treatment success, cancer stage, and Karnofsky Performance Status. The model is complex and will be

Table 18: Classification Of Patients Using Visit 0 Factor Mixture Analysis

5+3 = Using 8 LCSS items (excluding hemoptysis),
4+3 = Using 7 LCSS items (excluding hemoptysis and cough)

Final class counts and proportions for the latent classes based on the estimated model

Latent classes

	5+3		4+3	
1	175.13706	0.81082	132.24459	0.61224
2	40.86297	0.18918	83.75541	0.38776

Classification quality

	5+3	4+3
Entropy	0.945	0.838

Class counts and proportions

Latent classes

	5+3		4+3	
1	175	0.81019	132	0.61111
2	41	0.18981	84	0.38889

Classification table

Average latent class probabilities for most likely latent class membership (row)
by latent class (column)

	5+3		4+3	
	1	2	1	2
1	0.991	0.009	0.963	0.037
2	0.040	0.960	0.061	0.939

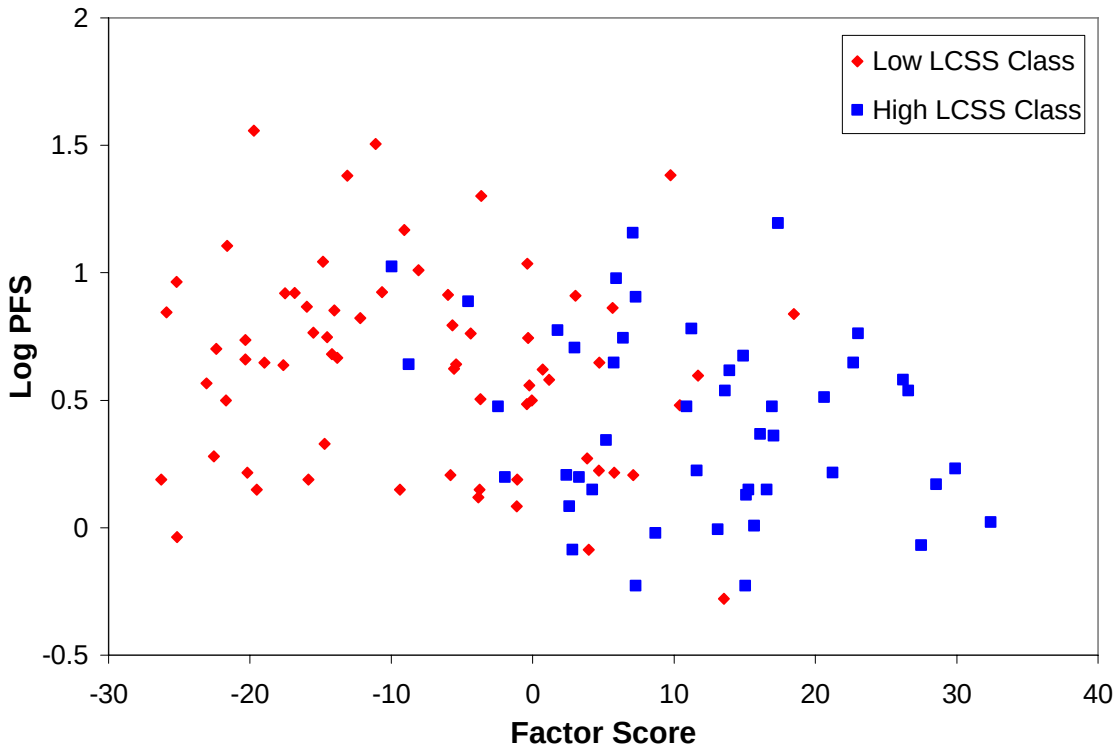


Figure 15: Relating latent variables at visit 0 to survival time, comparing the factor scores of the 1-factor model to the classification of the factor mixture model

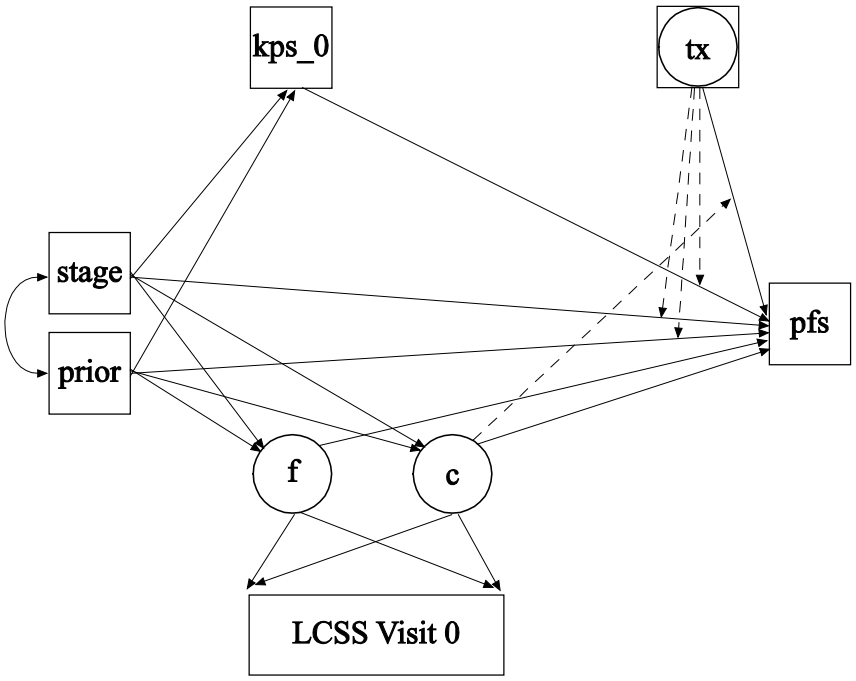


Figure 16: Predicting survival using a latent variable model for LCSS items at visit 0

shown in both diagram form and statistically. The model diagram is shown in Figure 16. The treatment dummy variable tx is here captured by a latent class variable with known class membership in order to allow maximum modeling flexibility. The treatment variable influences pfs (solid arrow) and also influences the influence of stage, prior, and kps on pfs (broken arrows), allowing for treatment interaction for these three covariates. The interaction between latent class and treatment is shown as a broken arrow from c to the arrow from tx to pfs.

6.1 Survival modeling with a latent class variable

Formally, the model can be described as follows. Let Z be a binary treatment variable. Let Z take values 0 (control) and 1 (treatment). Let C be a latent class binary variable that will correlate the LCSS observations at visit 0 and the survival PFS variable. Let Z take values 0 and 1. Let $Y_1, ..., Y_p$ denote the LCSS observations at

visit 0. The distribution of Y is described by the following equations

$$Y_i|C = \mu_{ic} + \lambda_i \, \eta + \varepsilon_i$$

where η is a standard normal random variable and ε_i are independent residuals with variance θ_i. An alternative way to write this model is

$$Y_i|C = \mu_i + \beta_i \, C + \lambda_i \, \eta + \varepsilon_i.$$

The distribution of the survival variable is described by the following hazard model (c is the constant 0.7 corresponding to 3 weeks periods)

$$log(h(t|Z,X,C)) = log(h_0(t)) + (\alpha_1 + \gamma_1 \, [t/c])Z + (\alpha_2 + \gamma_2 \, [t/c])C$$

$$+(\alpha_3 + \gamma_3[t/c])ZC + (\alpha_4 + \gamma_4 \, [t/c]) \, X_1$$

$$+(\alpha_4 + \gamma_4[t/c])X_1 Z + \beta_2 \, X_2 + \beta_3 \, X_3$$

where X_1 is the KPS variable, X_2 and X_3 are the prior treatment success and cancer stage variables and $h_0(t)$ is the non-parametric baseline hazard function.

Finally the distribution of C is given by the following equation

$$P(C = 1) = \frac{1}{1 + exp(b_0 + b_2 X_2 + b_3 X_3)}.$$

6.2 Analysis results

The model estimates show that for the control group (BSC) only the partial effect of the covariate prior on survival has a significant hazard estimate, increasing the hazard for the non treatment-resistant subjects. This significant effect is in contrast to the insignificant effect that was found in Table 7 when the covariate was entered alone. For the treatment group (P+BSC), stage has a significant hazard estimate, increasing the hazard for subjects at a more severe stage. Increasing Karnofsky Performance Status score (kps) significantly lowers the hazard in the treatment group.

The effects of LCSS latent class membership controlling for stage, prior, and Karnofsky are as follows. To vizualize the estimated model, the resulting survival curves are shown with curves evaluated at the most frequent cancer stage iv (referred to as stage=5 in Figure 17-19), for non-resistant treatment response, and for

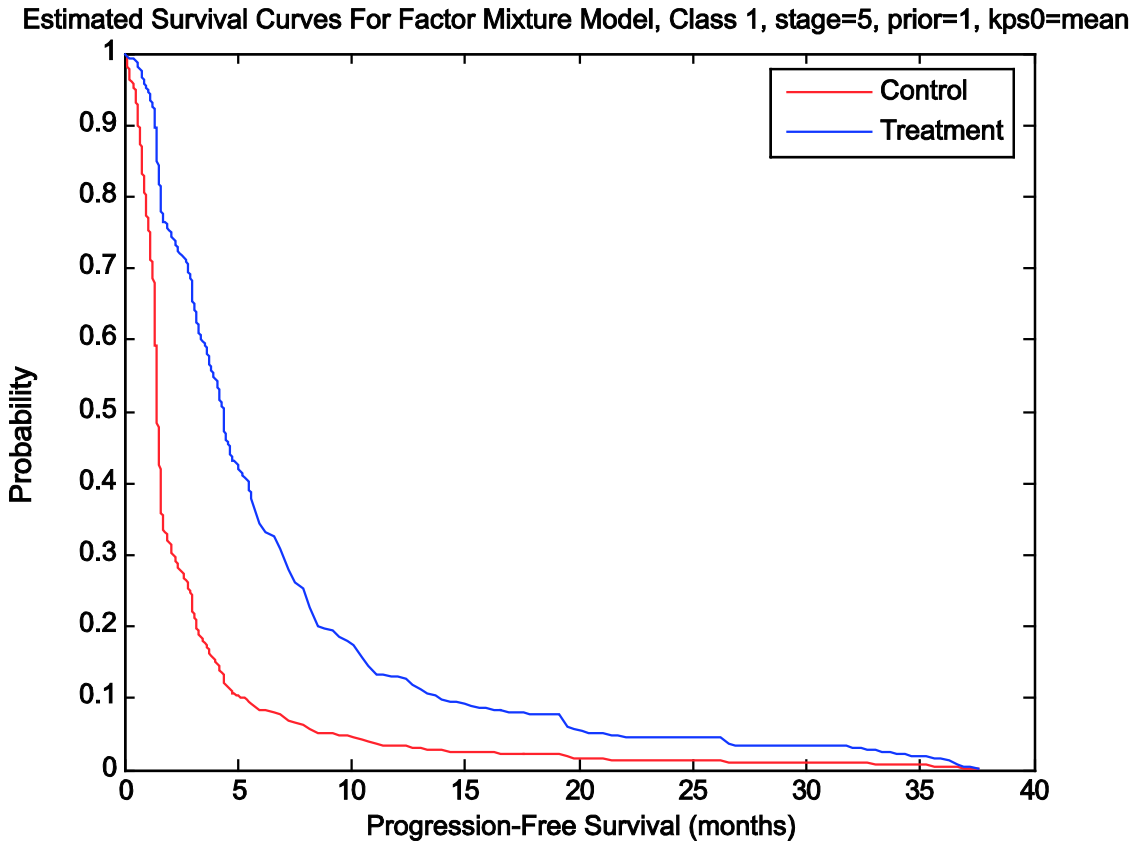


Figure 17: Factor mixture survival, low LCSS class, treatment vs control groups

average kps value. In the 61% low LCSS class (class 1; high quality of life), the hazard estimates are significant negative (lower) in the treatment group for time periods 1-3. Figure 17 shows the corresponding estimated survival curves and it is seen that the treatment and control curves show a larger difference than the Kaplan-Meier curves of Figure 1. Furthermore, the curves are not crossing. The median survival times for control and treatment groups are 1.5 and 4.3, respectively.

In the 39% high LCSS class (low quality of life), there is a significant negative treatment-control group hazard difference only for the first time period. Figure 18 shows the estimated survival curves. The treatment-control curves are closer and do cross. The median survival times for control and treatment groups are 1.4 and

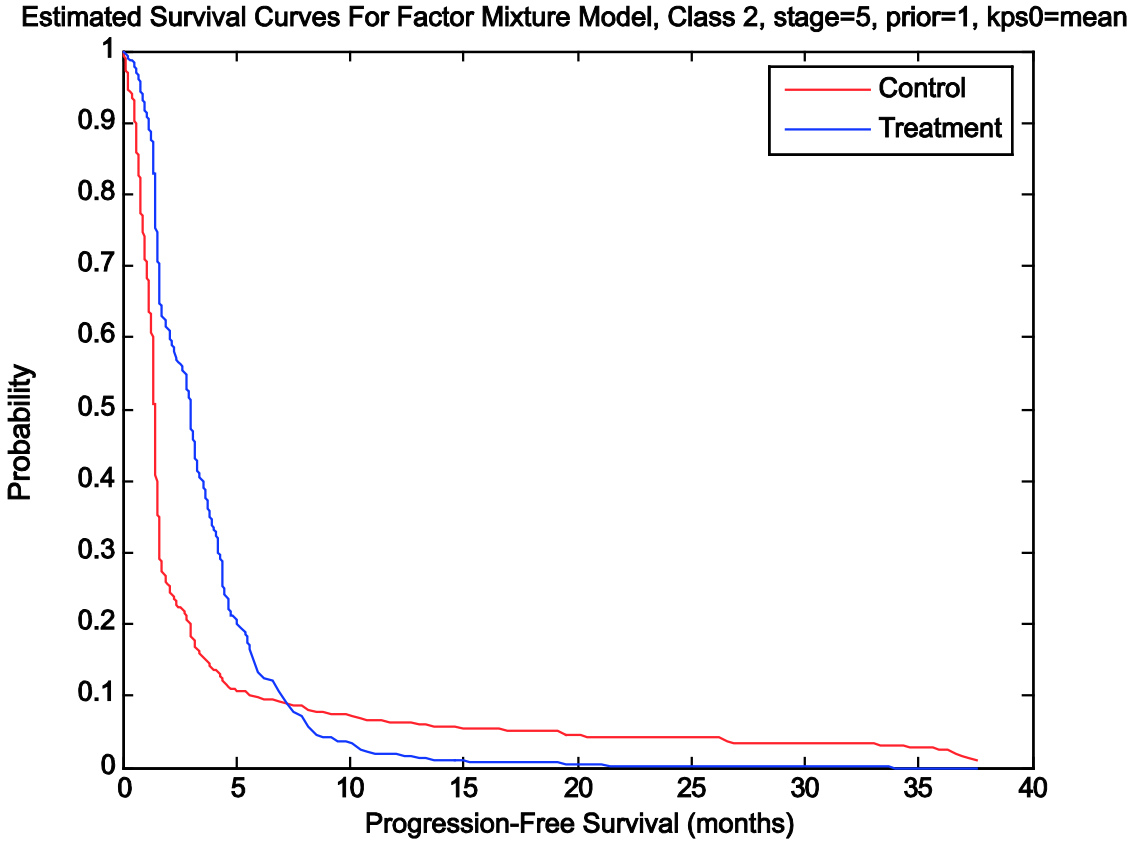


Figure 18: Factor mixture survival, high LCSS class, treatment vs control groups

3.0, respectively. The difference between Figure 17 and Figure 18 indicate the magnitude of the LCSS effect on survival.

Figure 19 and 20 further explore the effect of LCSS class membership. Figure 19 shows the estimated survival curves for high and low LCSS class among treatment group subjects. Here, the hazard estimates are significantly lower in the low LCSS class for time periods 3-9. Figure 20 shows the corresponding control group curves. Here, there are no significant differences in the hazard estimates. Taken together, the figures illustrate the interaction between treatment and LCSS class in influencing survival, controlling for stage, prior, and kps which are also allowed to interact with treatment. In terms of survival, the latent class distinction has little importance in the control group, but is important in the treatment group

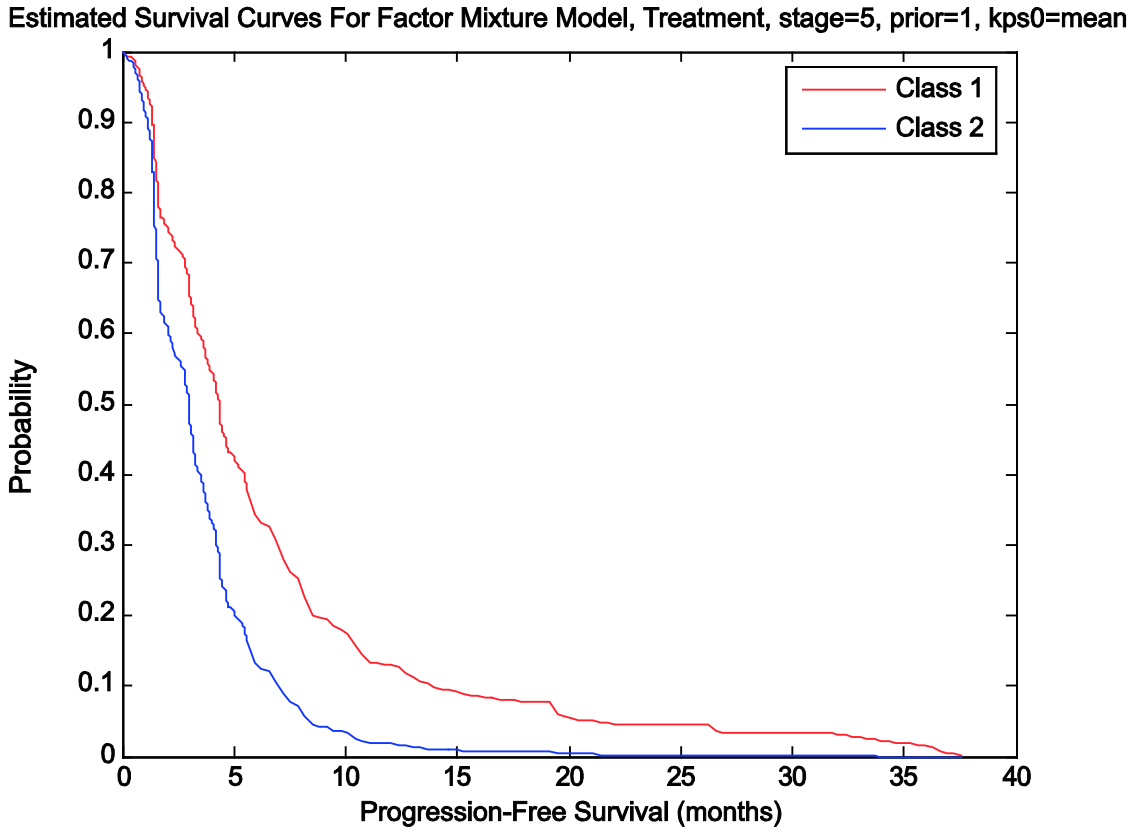


Figure 19: Factor mixture survival, treatment group, high and low LCSS classes

in that subjects in the low LCSS class (high quality of life) benefit more from the treatment.

7 Modeling survival using a latent variable model for longitudinal LCSS item information

7.1 Latent transition analysis over the first two visits

Because the factor mixture model was chosen as the best model for the LCSS items, the question arises how modeling of the longitudinal development of the

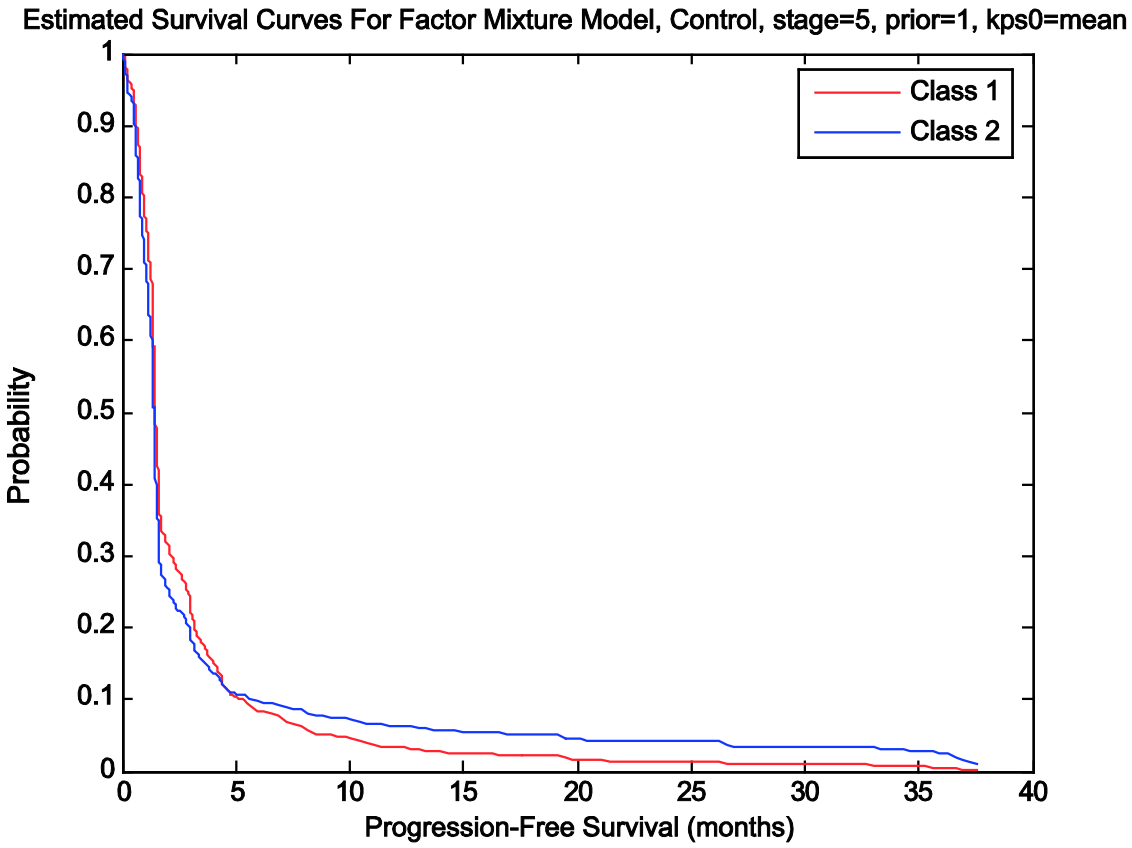


Figure 20: Factor mixture survival, control group, high and low LCSS classes

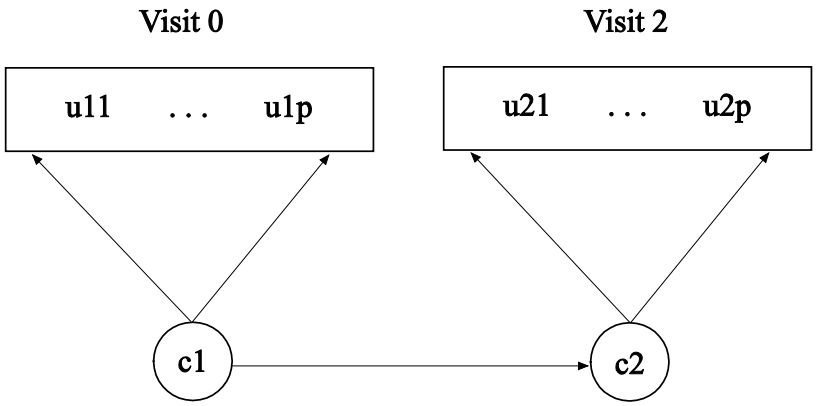


Figure 21: Latent transition analysis

LCSS items should be formulated. In Section 4 a random effect growth model was used, but this does not account for the latent LCSS classes. The growth mixture model (Muthén & Asparouhov, 2008) that was applied in Section 4 concerned a continuous outcome. In contrast, when using the factor mixture model the latent class variable is the primary dependent variable. In such a case, changes over time can be described by latent (hidden) Markov modeling, also referred to as latent transition analysis (LTA). The idea behind LTA is shown in Figure 21 for visits 0 and 2. At each time point, a latent class measurement model is used. Measurement invariance over time is applied. The visit 0 latent class variable c1 influences c2 via a multinomial logistic regression. This produces a transition table containing the probability of changing class or staying in the same class over time.

Muthén (2008) proposed a modified LTA using a factor mixture measurement model at each time point. This model is shown in Figure 22. The changes in LCSS over time will be described using this FMA-LTA model. The estimated class probabilities and transition probabilities are shown in Table 19. The high LCSS class (low quality of life) is estimated at 43% at visit 0, only slightly higher than the 39% in earlier analyses. The high LCSS class at visit 2 is estimated at 45%. The transition table shows a high degree of stability in latent class membership over the two visits with only 11% moving from the low class (class 1) to the high class (class 2) and 9% moving from high to low class.

Table 19: Latent Transition Analysis Of Visit 0 And Visit 2

Class counts and probabilities

c0	Low LCSS class
138.98404	0.57195
	High LCSS class
104.01595	0.42805
c2	Low LCSS class
133.48444	0.54932
	High LCSS class
109.51556	0.45068

Latent transition probabilities based on the estimated model

c0 classes (rows) by c2 classes (columns)

	1	2
1	0.894	0.106
2	0.089	0.911

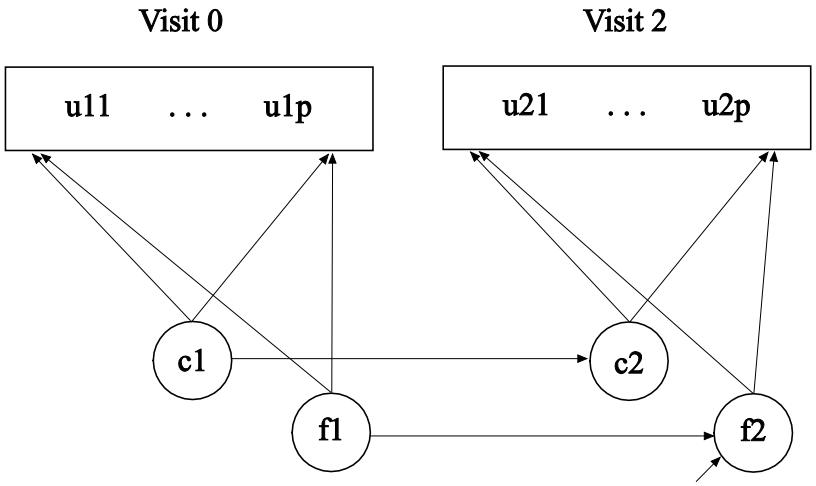


Figure 22: Factor mixture latent transition analysis

7.2 Joint latent transition - survival analysis

As a next analysis step, the FMA-LTA is combined with a survival model using the same non-proportional hazard representation used earlier. Figure 23 shows the model in diagram form with the latent class variables influencing pfs.

The estimated survival curves are shown in Figures 24-27. Figure 24 shows the difference between treatment and control survival curves for subjects who are in the low LCSS class at both visits. Comparing this figure with the visit 0 low LCSS class survival curves of Figure 17 suggests that subjects have no extra survival advantage of staying in the low LCSS latent class membership at visit 2. It appears that the initial LCSS status at visit 0 is the key factor in treatment affecting survival.

7.3 Joint latent variable growth-survival analysis of all time points

As a final modeling step, six models are considered for the LCSS items at visits 0 - 9. The six models are shown in Figures 28-33.

Model M1, shown in Figure 28, predicts survival using a Xu-Zeger trend model based on a single-factor model for the LCSS items. The single-factor model

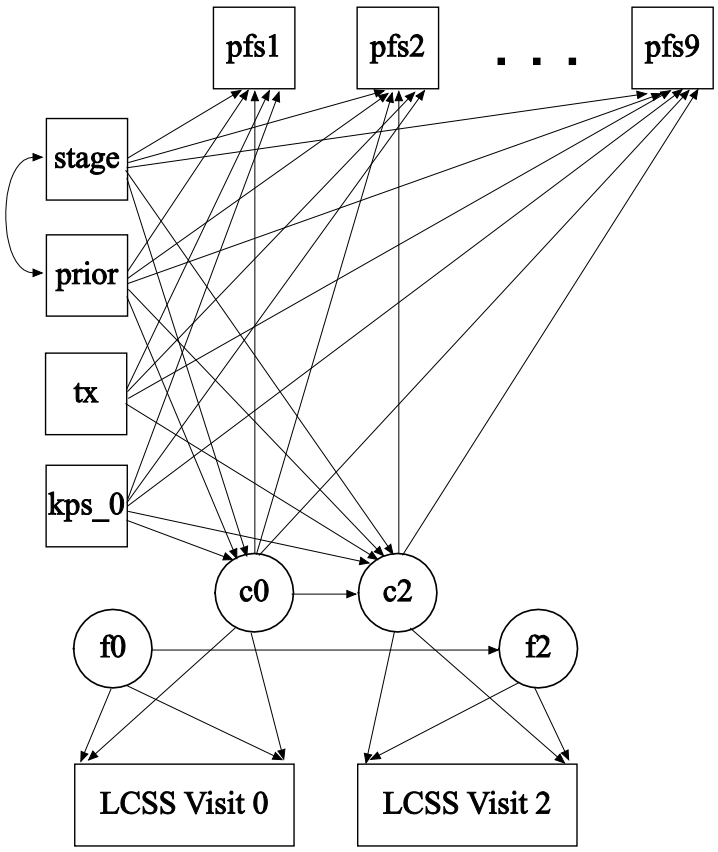


Figure 23: Joint factor mixture latent transition analysis - survival analysis

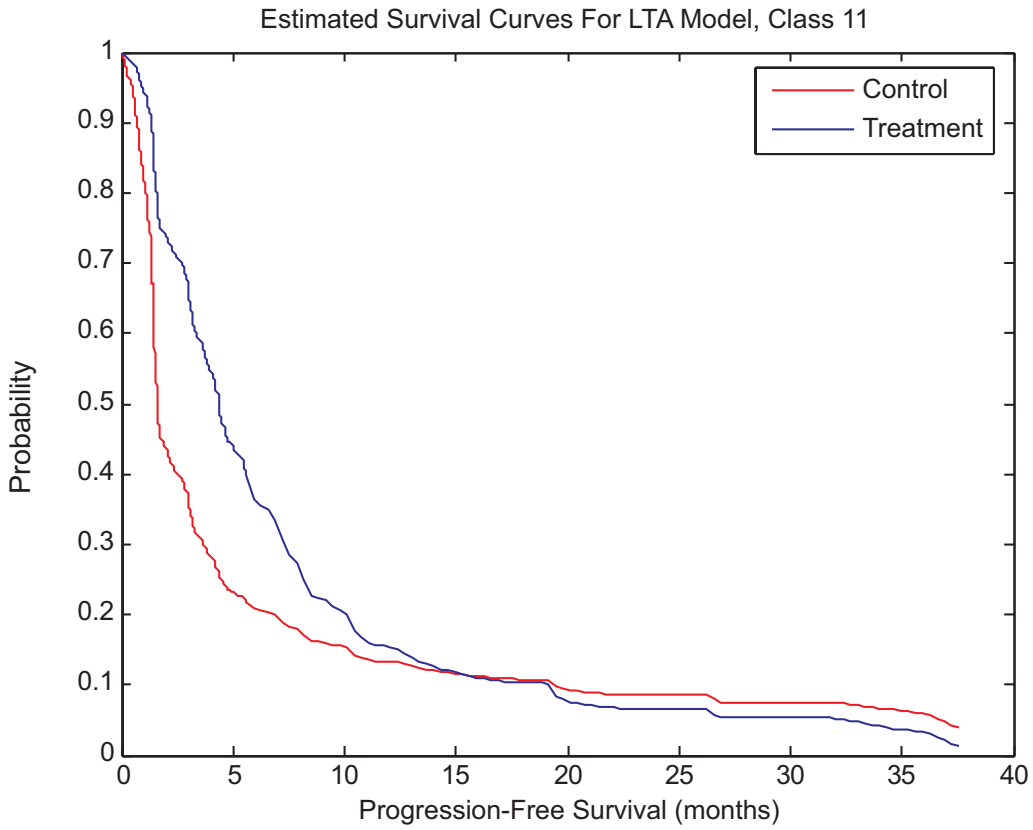


Figure 24: Joint factor mixture latent transition analysis - survival analysis: Low, low LCSS class

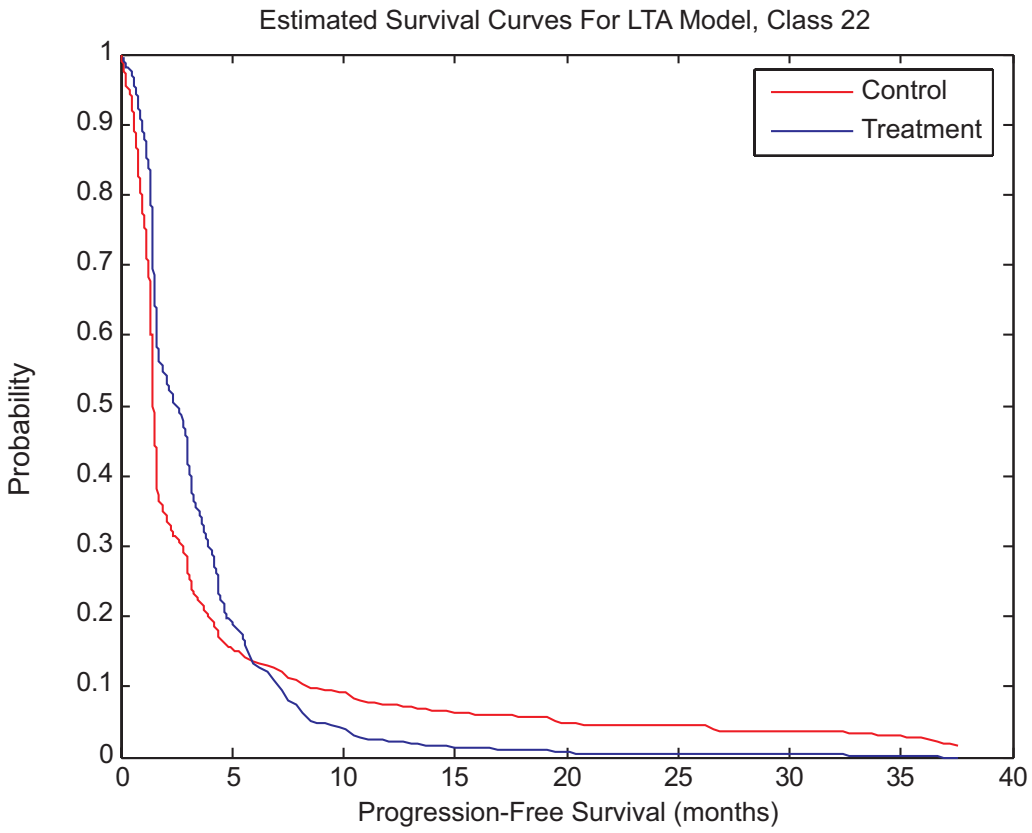


Figure 25: Joint factor mixture latent transition analysis - survival analysis: High, high LCSS class

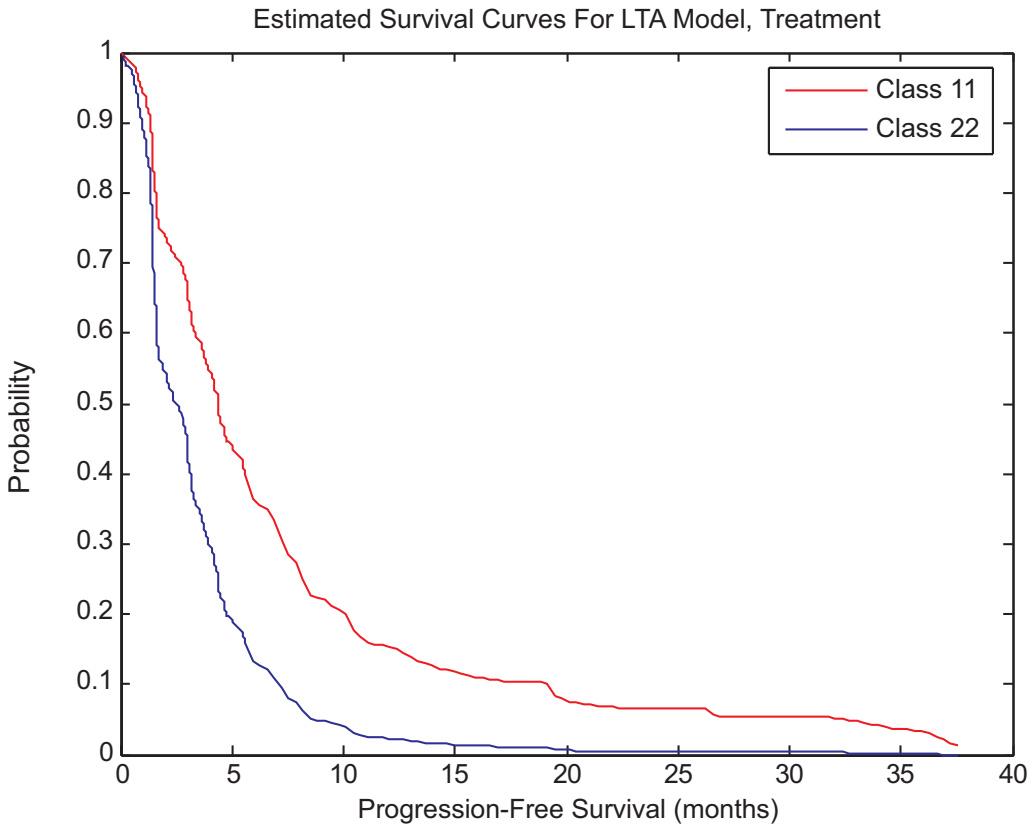


Figure 26: Joint factor mixture latent transition analysis - survival analysis: Treatment group

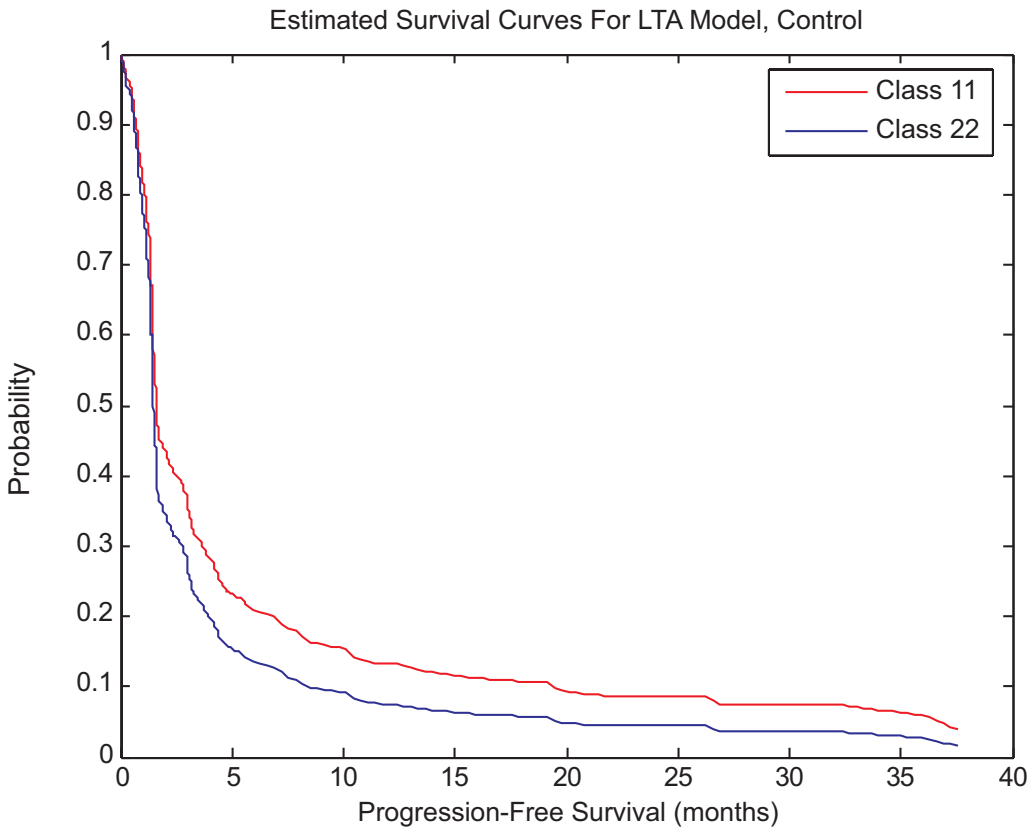


Figure 27: Joint factor mixture latent transition analysis - survival analysis: Control group

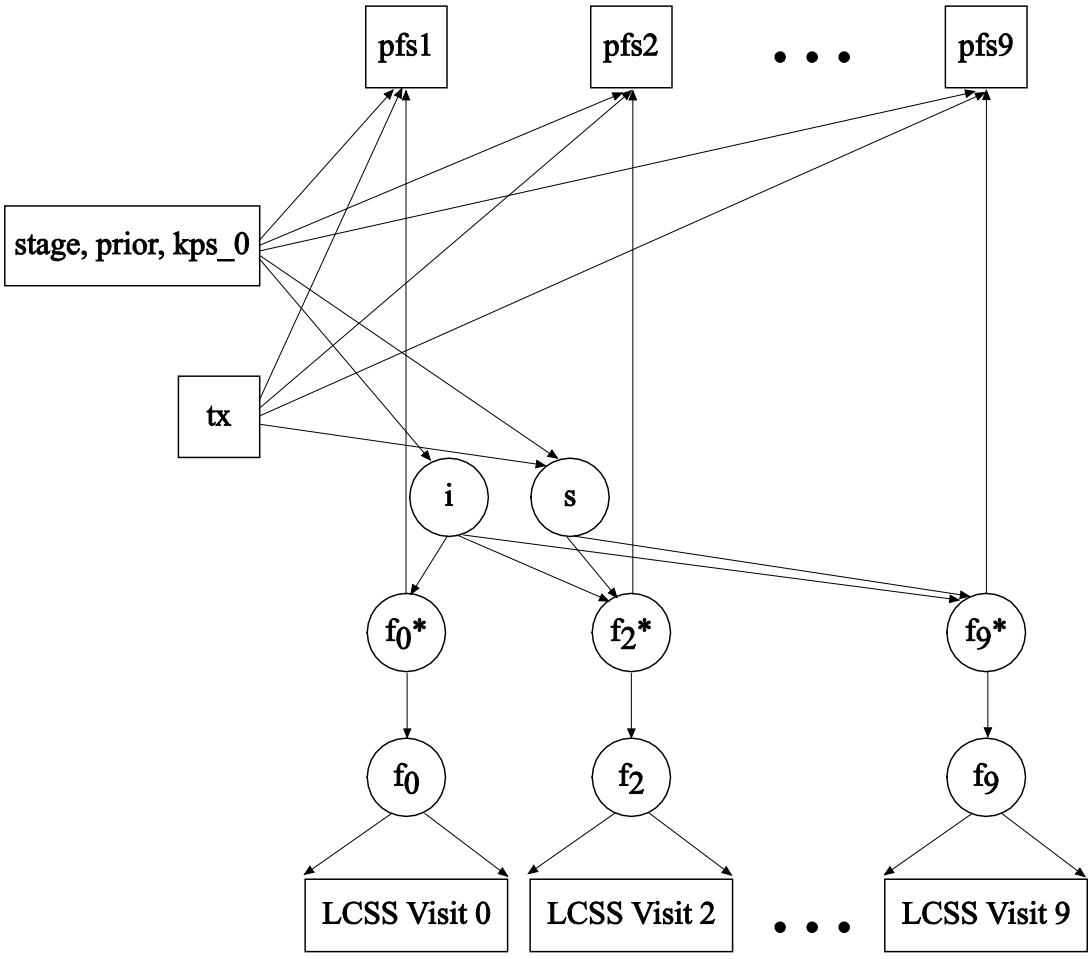


Figure 28: M1 - Predicting survival by a Xu-Zeger model with a single-factor model for LCSS items

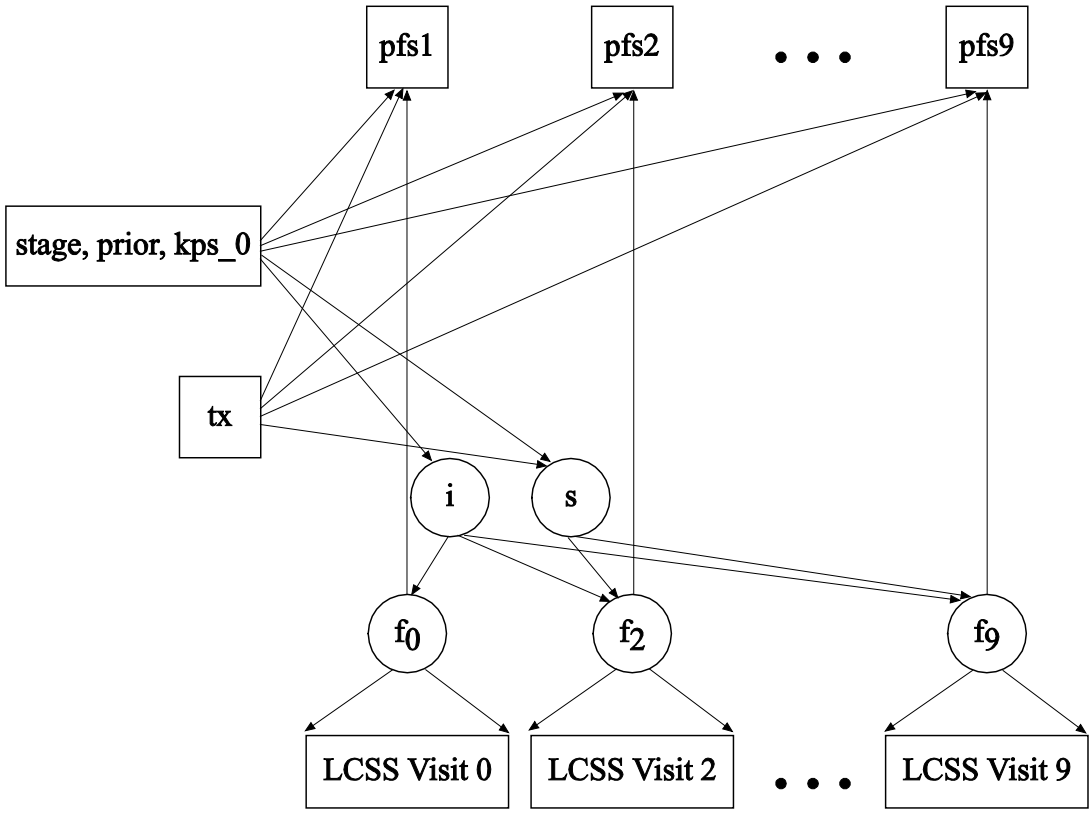


Figure 29: M2 - Predicting survival from factors of single-factor model for LCSS items

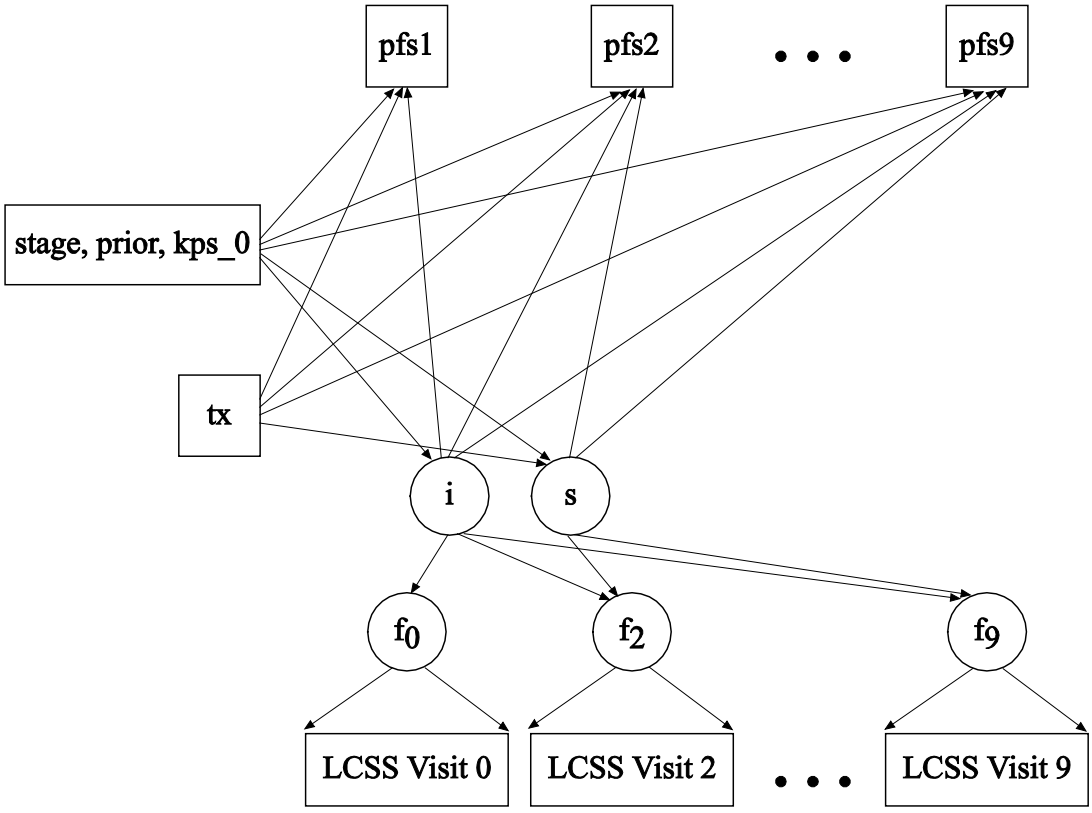


Figure 30: M3 - Predicting survival by random effects with a single-factor model for LCSS items

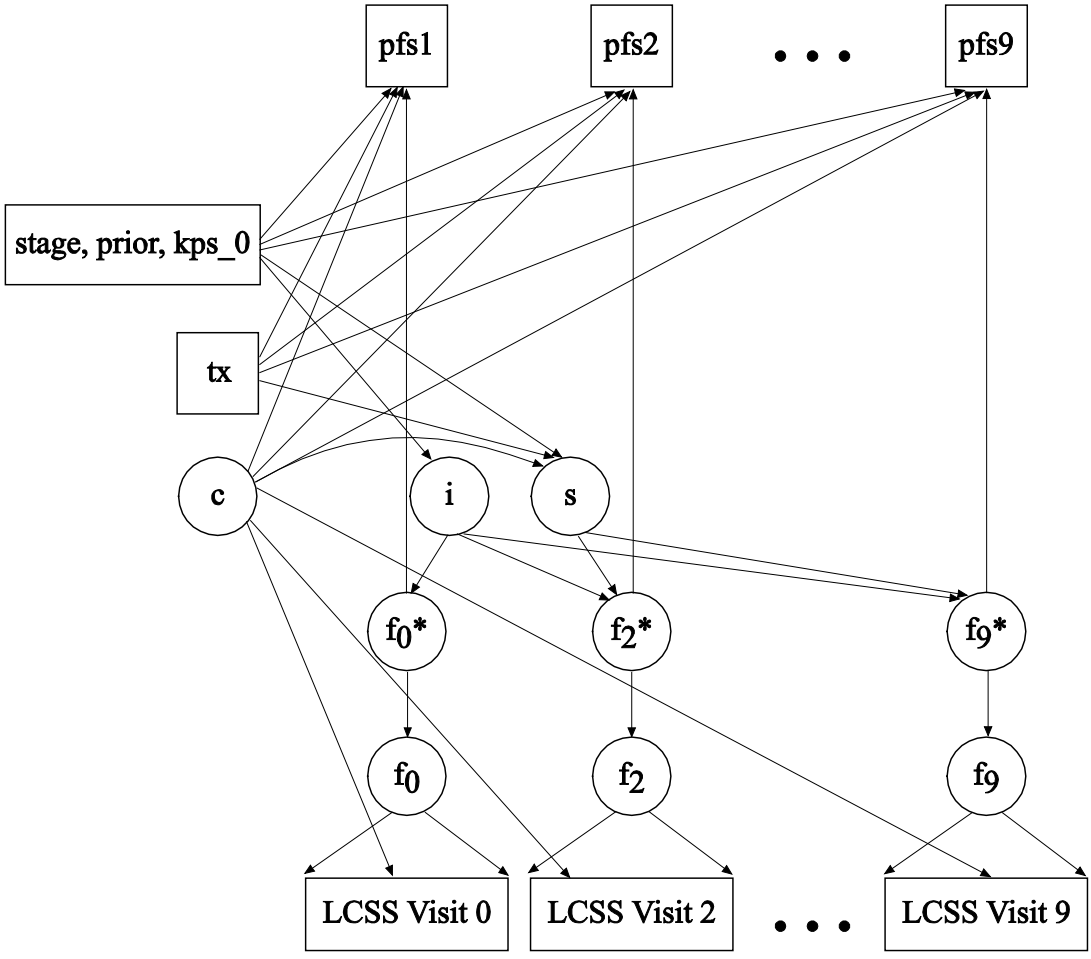


Figure 31: M4 - Predicting survival by a growth mixture Xu-Zeger model with a factor mixture model for LCSS items

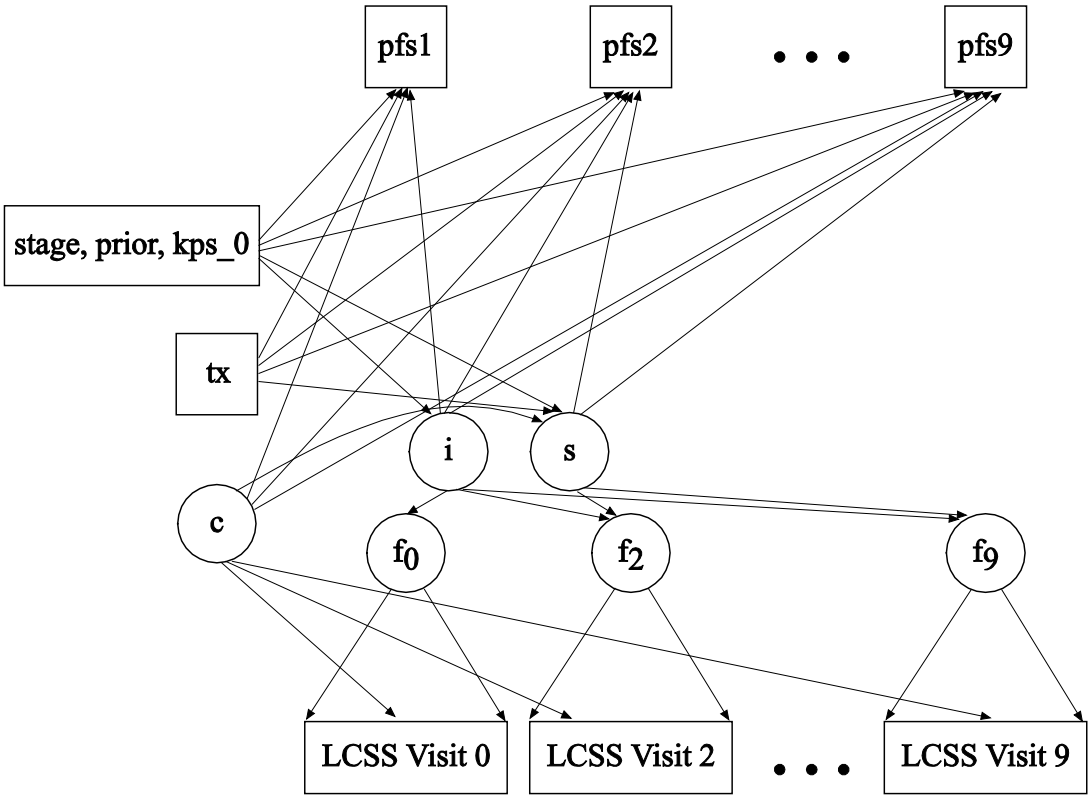


Figure 32: M5 - Predicting survival by random effects of a growth mixture model with a factor mixture model for LCSS items

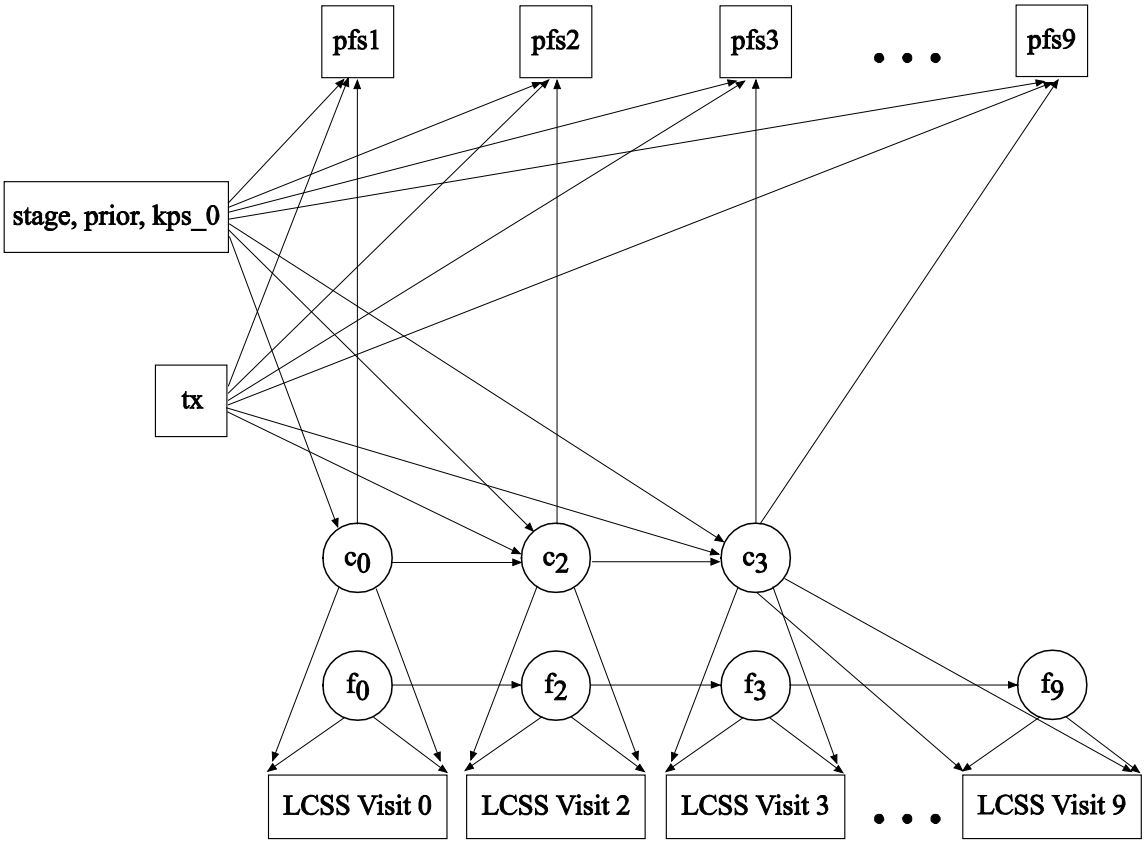


Figure 33: M6 - Predicting survival by latent transition model with a factor mixture model for LCSS items

was not found to be the best of the models investigated in Section 5, but is used for illustrative purposes here in order to compare its performance to the other models. Based on the results of Section 5, it can be applied to the three global LCSS items. Model M2, shown in Figure 29, uses the same latent variable model and is the latent variable counterpart to the Table 10 Figure 7 model using the observed LCSS item as a predictor of survival. Model M3, shown in Figure 30, bases the prediction of survival on the random effects i and s. Model M4, shown in Figure 31, uses a growth mixture model (Muthén & Asparouhov, 2008) and bases the survival prediction on a Xu-Zeger trend for the within-class severity factor. Model M5, shown in Figure 32, takes the same approach as in Model M4, but replaces the prediction from the trend with prediction from the random effects. Model M6, shown in Figure 33, extends the previously discussed latent transition factor mixture model to several time points and predicts survival from time-specific latent class membership.

Table 20 shows that the single-class models M1 and M3 are outperformed in terms of BIC by the mixture models M4, M5, and M6. The BIC differences among M4, M5, and M6 are not large. It is seen that only the mixture models give significant effects of the treatment on survival. As seen in earlier analyses, none of the models show a significant effect of treatment on LCSS development. The effect of LCSS development on survival is significant in all models.

8 Conclusions

8.1 Summary of substantive and statistical findings

From a substantive point of view, the latent variable survival analyses show that the information for the patient-reported items of the Lung Cancer Symptom Scale-Meso instrument is useful in predicting progression-free survival in a second-line treatment study. This is the case also when controlling for cancer stage, prior treatment, and the clinician-rated Karnofsky Performance Status score. The latent variable analyses of LCSS suggest that a model with two latent classes is preferable, classifying subjects into a high and a low LCSS class. Controlling for the covariates, the class membership has no significant effect on survival in the control group (BSC arm), but has a significant effect on survival in the treatment group (P+BSC arm). Although subjects in either class benefit from the treatment, subjects in the low LCSS class, corresponding to high quality of life, have a significant added benefit from the treatment in terms of longer survival.

Table 20: Joint latent variable growth-survival analysis of all time points

Model	Log-Likelihood	Number of Parameters	BIC	Tx Effect on PFS	Tx Effect on LCSS	LCSS Effect on PFS
M1	-11642	110	23883	No	No	Yes
M2	too	heavy	computations			
M3	-11639	113	23894	No	No	Yes (i only)
M4	-11546	67	23456	Yes (low class only)	No	Yes
M5	-11544	69	23464	Yes (low class only)	No	Yes (both i and s)
M6	-11567	60	23462	Yes	No	Yes

Model description:
M1: Survival predicted by a Xu-Zeger model with a single-factor model for LCSS items (Figure 28).
M2: Survival predicted from factors of single-factor model for LCSS items (Figure 29).
M3: Survival predicted by random effects with a single-factor model for LCSS items (Figure 30).
M4: Survival predicted by a growth mixture Xu-Zeger model with a factor mixture model for LCSS items (Figure 31).
M5: Survival predicted by random effects of a growth mixture model with a factor mixture model for LCSS items (Figure 32).
M6: Survival predicted by latent transition model with a factor mixture model for LCSS items (Figure 33).

The analysis results suggest that the baseline LCSS information collected at visit 0 provides important predictive information regarding potential treatment success. In this study, about 40% of the subjects are classified into the high LCSS class corresponding to poor quality of life. This portion of patients benefit less from the treatment in terms of progression-free survival.

Treatment effects on PRO in the form of the LCSS items were not found to be significant. This may, however, be due to insufficient power given the relatively small sample size of $n = 243$. The development of the LCSS items over time were found to provide significant prediction of survival.

Statistically, the report shows that continuous-time survival modeling with latent variables is now feasible. This is the case when using latent classes, latent variables in the form of random effects, and combinations of the two kinds of latent variables. The modeling in Section 6, predicting survival from a baseline factor mixture model, is not difficult or time-consuming to compute. The joint latent transition - survival modeling of Section 7 is more challenging but still feasible. A more difficult task is to combine a mixture model with time-varying information such as from LCSS items. This calls for predicting survival in different time intervals using either a time-varying observed variable or a factor. Because of the special type of data, however, where the dependent variable and he predictors are missing simultaneously in most cases, it is possible to estimate the model with a small number of integration points. This makes the computation still feasible when predicting from LCSS items.

8.2 Future topics

Based on an estimated model such as in Section 6 it is of great interest to compare the survival rates under treatment and control for a new patient with specific background variables as well as responses to the initial LCSS evaluation. This is not a straightforward task, however, because the treatment variable affects many parts of this model in contrast to a simple regression model of survival. Thus sufficient care should be taken to use the full advantage and flexibility of these models. In addition the task is complicated somewhat by the various ways the actual specific question is posed. For example, you can change the variables that PFS is conditioned on, i.e., you can simply consider the distribution of [PFS | TX, stage] or [PFS | TX, stage, LCSS0] or [PFS | TX, stage, LCSS]. The easiest to compute is [PFS | all other variables, TX=1] and [PFS | all other variables, TX=0] as this simply amounts to plugging in these values into the Cox regression. If however you condition only on a subset of all the variables that are in the model such as

LCSS0 for example one has to derive a new model with only these variables that is deduced from the full model with all the variables. This would involve the distribution of [missing variables | observed variables] which should be already in the model. In some cases the total TX effect on survival is decomposed as a TX effect on LCSS variables which in turn affects survival as well as the direct TX effect on survival. Additional complications arise when continuous and categorical latent variables are involved in the model as these will typically need to be integrated out of the model as well. Standard errors can also be included into various final results. Typically these will be constructed with the delta method following the point estimate derivation. Using the non-parametric baseline methodology, however, it is currently not possible to derive confidence limits for anything that involves the baseline curve. One feasible alternative is to use a stepwise parametric model as a substitute for the non-parametric baseline model. For these reasons, the topic of estimating treatment benefits under different baseline conditions deserves further research.

Power analysis using latent variable survival models is too complex to be expressed via explicit formulas, but can be carried out using the Mplus Monte Carlo feature. Such Monte Carlo data generation however is fairly complicated for several reasons and so should be done carefully. The first step in the data generation process is to realize that the generation of PFS can not be done as one variable. Instead it has to be done as in the model, i.e., by generating PFS1, PFS2, ... , PFS9 as well as the corresponding censoring indicators C1,C2, ... ,C9. Mplus can generate censored survival data based on an exponential distribution for the censoring process. But these variables have two types of censoring, a deterministic one and an exponential random censoring. The second type can be implemented with Mplus Monte Carlo by choosing an appropriate level of the censoring hazard parameter. The first type, however, is deterministic and has to be implemented with the DEFINE statement, by simply truncating at 0.7. Because the DEFINE statements are not available in internal Monte Carlo that means that the generation and analysis has to be done with "External" Monte Carlo. This is done as follows. In Step 1 multiple data sets are generated and saved. In Step 2 data are further manipulated and serially analyzed. The deterministic censoring has to be done in the second step. Missing data specification is also critical in these analyses. Here missing data also has two sources, random and "deterministic". The deterministic source simply says that all variables are missing after the first PFSj variable that is smaller than 0.7. This type of missingness has to be imposed also in the DEFINE command, i.e., in step 2. The first type of missingness, the random kind, has to be generated in the first step with the internal Monte Carlo generation. Note that

both missingness proportions and censoring proportions should be computed separately since these quantities are not part of the model. Also to generate survival data in Mplus you have to specify a baseline hazard function which is also not given in this model. You can specify that function as a constant with the appropriate constant taken from a similar model assuming that the PFSj variables have a constant baseline. The generation of the covariates in the model should also be considered carefully, for example the variable stage is a categorical variable and it should be generated as such. Because the model typically does not involve any modeling for the covariates such modeling should be done separately and used in the data generation process. This means preserving not only the univariate distribution of the covariates but also the multivariate, i.e., a separate multivariate model for the covariates should be constructed.

References

[1] Asparouhov, T., Masyn, K. and Muthén, B. (2006). Continuous time survival in latent variable models. Proceedings of the Joint Statistical Meeting in Seattle, August 2006. ASA section on Biometrics, 180-187.

[2] Beunckens, C., Molenberghs, G., Verbeke, G., and Mallinckrodt, C. (2008). A latent-class mixture model for incomplete longitudinal Gaussian data. Biometrics, 64, 96-105.

[3] Bollen, K.A., Lennox, R.D., and Dahly, D.L. (2009). Practical application of the vanishing tetrad test for causal indicators measurement models: An example from health-related quality of life. Statistics in Medicine, 28, 1524-1536.

[4] Breslow, N.E. (1974). Covariance analysis of censored survival Data. Biometrics, **30**, 89-99.

[5] Diggle, P., Henderson, R. and Philipson, P. (2008). Random-effects models for joint analysis of repeated-measurement and time-to-event outcomes. In Fitzmaurice, G., Davidian, M., Verbeke, G. & Molenberghs, G. (eds.), Longitudinal Data Analysis, pp. 349-366. Boca Raton: Chapman & Hall/CRC Press.

[6] Hollen, P.J., Gralla, R.J., Liepa, A.M., Symanowski, J.T., and Rusthoven, J.J. (2004). Adapting the Lung Cancer Symptom Scale (LCSS) to mesothelioma. Cancer, 101, 587-595.

[7] Jassem, J., Ramlau, R., Santoro, A., Schuette, W., Chemaissani, A., Hong, S., Blatter, J., Adachi, S., Hanauske, A., and Manegold, C. (2008). Phase III trial of pemetrexed plus best supportive care compared with best supportive care in previoulsy treated patients with advanced lmalignant pleural mesothelioma. Journal of Clinical Oncology, 26, 1698-1704.

[8] Larsen, K. (2004). Joint analysis of time-to-event and multiple binary indicators of latent classes. Biometrics, **60(1)**, 85-92.

[9] Larsen, K. (2005). The Cox proportional hazards model with a continuous latent variable measured by multiple binary indicators. Biometrics, **61(4)**, 1049-1055.

[10] Lin, H., McCullogh, C.E., and Rosenheck, R.A. (2004). Latent pattern mixture models for informative intermittent missing data in longitudinal studies. Biometrics, 60, 295-305.

[11] Little, R.J. (2008). Selection and pattern-mixture models. In Fitzmaurice, G., Davidian, M., Verbeke, G. & Molenberghs, G. (eds.), Longitudinal Data Analysis, pp. 409-431. Boca Raton: Chapman & Hall/CRC Press.

[12] Little, R. J. and Rubin, D. B. (2002). *Statistical analysis with missing data*. Second edition. New York: John Wiley and Sons.

[13] McLachlan, G. J. and Peel, D. (2000). *Finite mixture models*. New York: Wiley and Sons.

[14] Muthén, B. (2008). Latent variable hybrids: Overview of old and new models. In Hancock, G. R., & Samuelsen, K. M. (Eds.), Advances in latent variable mixture models, pp. 1-24. Charlotte, NC: Information Age Publishing, Inc.

[15] Muthén B. and Asparouhov, T. (2008). Growth mixture modeling: Analysis with non-Gaussian random effects. In Fitzmaurice, G., Davidian, M., Verbeke, G. & Molenberghs, G. (eds.), Longitudinal Data Analysis, pp. 143-165. Boca Raton: Chapman & Hall/CRC Press.

[16] Muthén, B. and Muthén, L. (2008). Mplus User's Guide. Los Angeles, CA: Muthén & Muthén.

[17] Singer, J.D. and Willett, J.B. (2003). Applied longitudinal data analysis. Modeling change and event occurrence. New York: Oxford University Press.

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