Joint Modelling of Survival and Growth

Tihomir Asparouhov

June 2, 2010

Articles and Resources

- Mplus
- www.statmodel.com
- Muthén, B., Asparouhov, T., Boye, M., Hackshaw, M. & Naegeli, A. (2009). Applications of continuous-time survival in latent variable models for the analysis of oncology randomized clinical trial data using Mplus. Technical Report.
- Asparouhov, T., Masyn, K. & 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.

Overview

- Research questions
- Mesothelioma trial data
- Mplus framework
- Survival analysis of treatment effects: proportional versus non-proportional hazard modeling
- Joint growth-survival modeling

Research Questions

- Substantive questions
  - Are patient-reported QOL outcomes associated with survival?
  - Do QOL outcomes interact with treatment in affecting survival?
  - Do QOL outcomes have predictive power also when controlling for traditional covariates (stage, prior, Karnofsky)?
  - Do QOL outcomes measured at baseline predict survival?
  - Does QOL development relate to differences in survival?
- Statistical questions
  - Choice of basic survival model
  - Choice of latent variable and growth model
  - Choice of joint growth-survival model

Mesothelioma Trial Data

- Lung Cancer Symptom Scale (LCSS). Patient Rated Scale - 9 items measuring a latent factor QOL
- QOL (Quality of Life) - latent variable measured at 9 different times
- Karnofsky Scale: Doctor Rated Scale - Time Varying Covariate
- Progression Free Survival (PFS) : Survival Variable
- Treatment: Tx
- Additional covariates: prior chemo response, Mesothelioma stage

Mplus Framework

Mplus Statistical Framework For Time-to-Event Variables

- Cox Proportional Hazard Model
  - Nonparametric baseline hazard
  - Baseline hazard treated as nuisance parameters: the profile likelihood
  - Maximum-likelihood estimation
  - Can be embedded in the Mplus framework
  - Unique to Mplus
- Parametric Hazard Model
  - Stepwise baseline hazard
  - Stepwise baseline hazard with model constraints: approximation for any other parametric model
- Baseline Hazard
  - Estimated explicitly
  - Estimated as nuisance parameter
- Baseline Hazard In Mixture Models
  - Equal Across Class: Estimating single class effect
  - Unequal Across Class: Totally unconstraint

Mplus Statistical Framework For Time-to-Event Variables

- Frailty and Multilevel Models
  - Latent factor as a predictor of Time-to-Event Variables
  - Estimation via numerical integration
- Multivariate Time-to-Event Models
  - Multiple Time-to-Event Variables correlated via regressions on latent variables
  - Single Time-to-Event Variables converted to a Series of Time-to-Event Variables: Survival Series

$$T_k = \begin{cases} d_k - d_{k-1} & \text{if } d_k < T \\ missing & \text{if } T < d_{k-1} \\ T - d_{k-1} & \text{otherwise} \end{cases} \quad (1)$$

$$\delta_k = \begin{cases} 1 & \text{if } d_k < T \\ missing & \text{if } T < d_{k-1} \\ \delta & \text{otherwise} \end{cases} \quad (2)$$

where δ_k is the censoring indicator of T_k.

Mplus Statistical Framework For Time-to-Event Variables

Survival Series T_1,...,T_K

- T_i is survival during the i-th interval.
- The likelihood of T is equivalent to the likelihood of T_1,...,T_K.
- Non-proportional hazard modeling: varying the regression coefficient
- Time varying covariates: varying the predictor
- Latent growth process as predictor.
- Mplus produces joint survival curves $T_1, \ldots, T_K$.

The Proportional Hazard Model: Cox Regression

T: time-to-event variable such as death

$$\text{Hazard function } h(t) = h_0(t)\, Exp(\beta X) \tag{3}$$

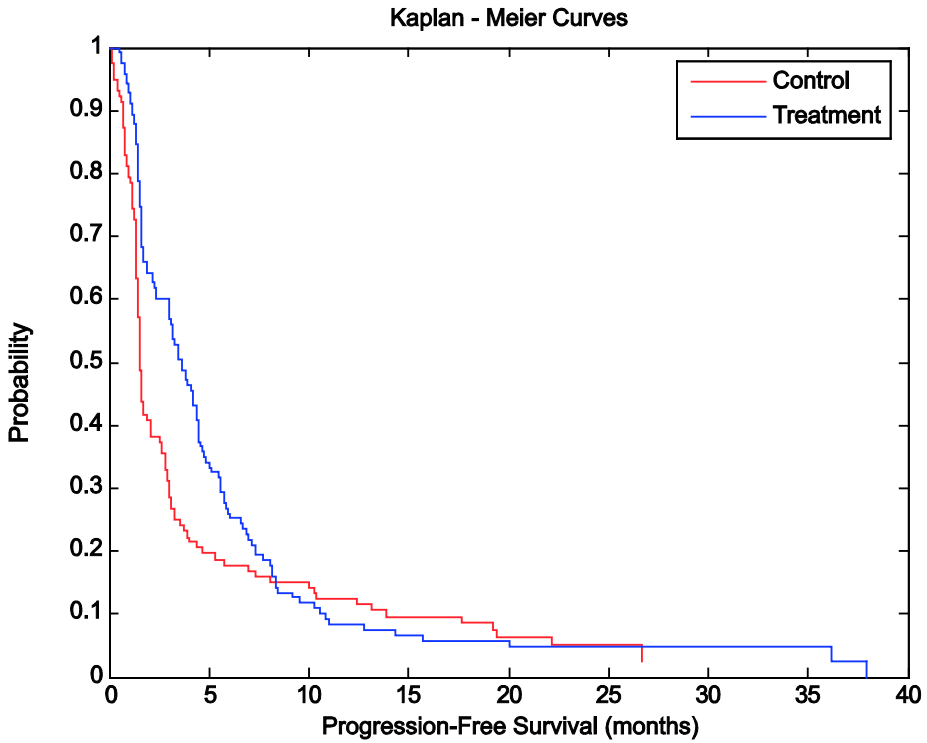
$$\text{Commulative hazard function } H(t) = \int_0^t h(s) d\, s \tag{4}$$

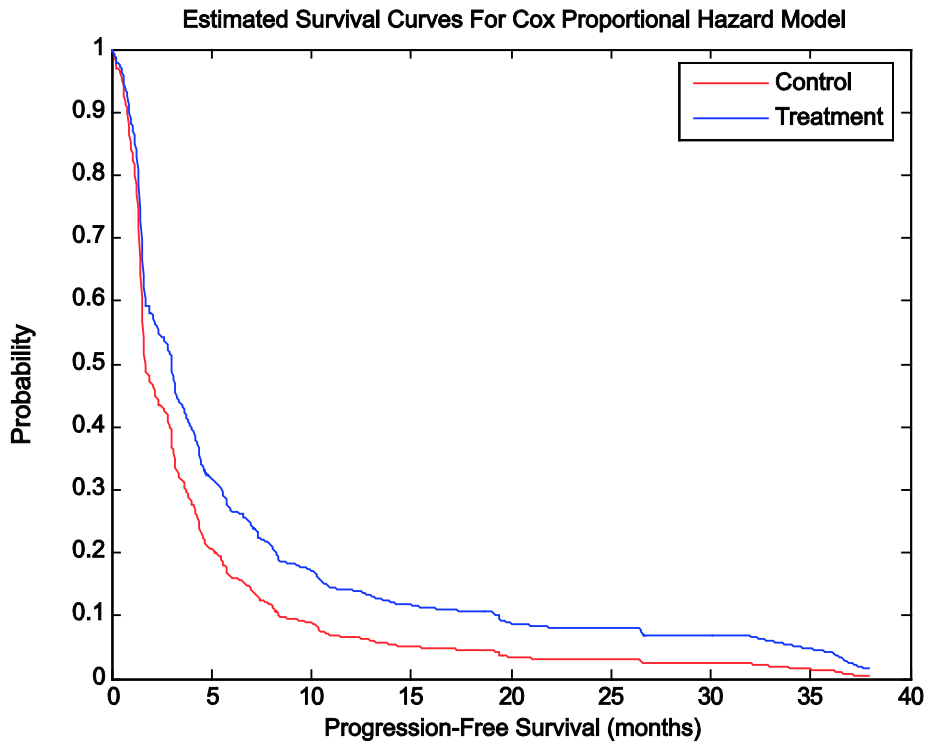
$$\text{Survival function } S(t) = P(T > t) = Exp(-H(t)) \tag{5}$$

$$S(t) = S_0(t)^{Exp(\beta X)} \tag{6}$$

Survival Analysis Of Treatment Effects

Kaplan-Meier vs Cox Proportional Hazard Model





Alternative hazard models

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, Cox proportional hazard model

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

Model 2, Linear non-proportional hazard model

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

where h_0 is an unrestricted non-parametric function. The model shows an interaction between treatment arm and time.

Model 3, Linear non-proportional hazard model: Survival Series

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

where h_0 is unrestricted non-parametric function and [ ] is the integer part function. The constant c can be any number. Modeled as Survival Series. As $c -> 0$ Model 3 becomes equivalent to Model 2.

Model 4, Unrestricted non-proportional hazard model: Survival Series

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

where h_0 is unrestricted non-parametric function, [ ] is the integer part function and $\alpha_1, \alpha_2, \ldots$ 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 as Survival Series. Under the parameter constraints

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

Model 4 becomes equivalent to Model 3. Using these constraints in Model Constraints is how Model 3 is estimated in Mplus.

Model 5, Unrestricted non-proportional hazard model

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

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

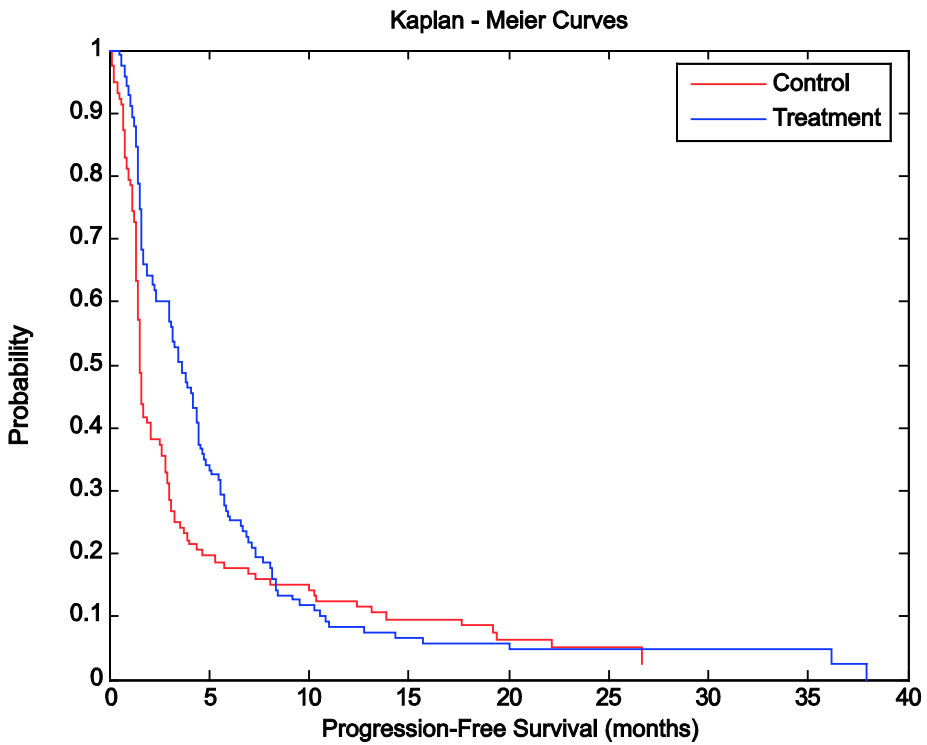
Model 1, 3, 4, 5 can be estimated in Mplus directly. Model 2 is approximated by Model 3.

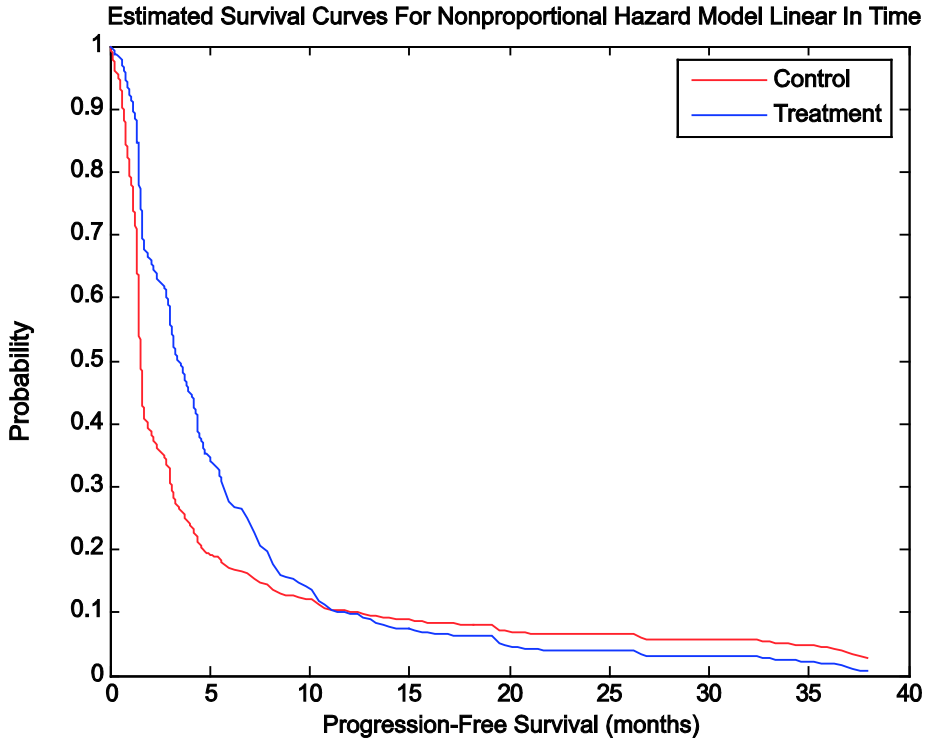
Table: Summary of hazard modeling of treatment effects for Mesothelioma trial data

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

LRT testing for nested models.

Kaplan-Meier vs Non-Proportional Hazard Model





Joint Growth-Survival Modeling

Choice Of Joint Growth-Survival Models

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

$$Y_{it} = Y_{it}^{*} + \varepsilon_{it} \tag{7}$$

$$Y_{it}^{*} = \alpha_i + \beta_i \, t \tag{8}$$

where α_i and β_i are normally distributed random effects.
Model 1. The Xu-Zeger Model. The Xu-Zeger model is given by

$$log(h(t)) = log(h_0(t)) + \gamma \, Y_{it}^{*} + \beta \, X. \tag{9}$$

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

Choice Of Joint Growth-Survival Models, Continued

Model 2. The Approximate Xu-Zeger. For a constant c let

$$Y^*_{itc} = \alpha_i + \beta_i \, c[t/c] \tag{10}$$

The Mplus Xu-Zeger approximation model is given by

$$log(h(t)) = log(h_0(t)) + \gamma \, Y^*_{itc} + \beta \, X. \tag{11}$$

As $c- > 0$ Model 2 is equivalent to Model 1. Model 2 is implemented in Mplus by modeling a Survival Series.

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

$$log(h(t)) = log(h_0(t)) + \gamma \, Y_{i([t/c]c)} + \beta \, X. \tag{12}$$

Model 3 is implemented in Mplus by modeling a Survival Series. The model is based on the assumption that the variables Y are observed at times c, 2c, 3c, ... or approximately so.

Model 4. Growth Mixtures. For a two-class Mixture model, assuming the class variable C takes values 0 and 1, the model is given by

$$\alpha_i|C = \gamma_0 + \gamma_1 C + \gamma_2 X + \gamma_3 C\,X + \varepsilon_{1,i} \tag{13}$$

$$\beta_i|C = \gamma_4 + \gamma_5 C + \gamma_6 X + \gamma_7 C\,X + \varepsilon_{2,i} \tag{14}$$

$$log(h(t)) = log(h_0(t)) + \gamma_8\ C + \gamma_9\ X + \gamma_{10}\ C\,X. \tag{15}$$

The correlation between the growth model and the latent variable is entirely through the latent class variable.

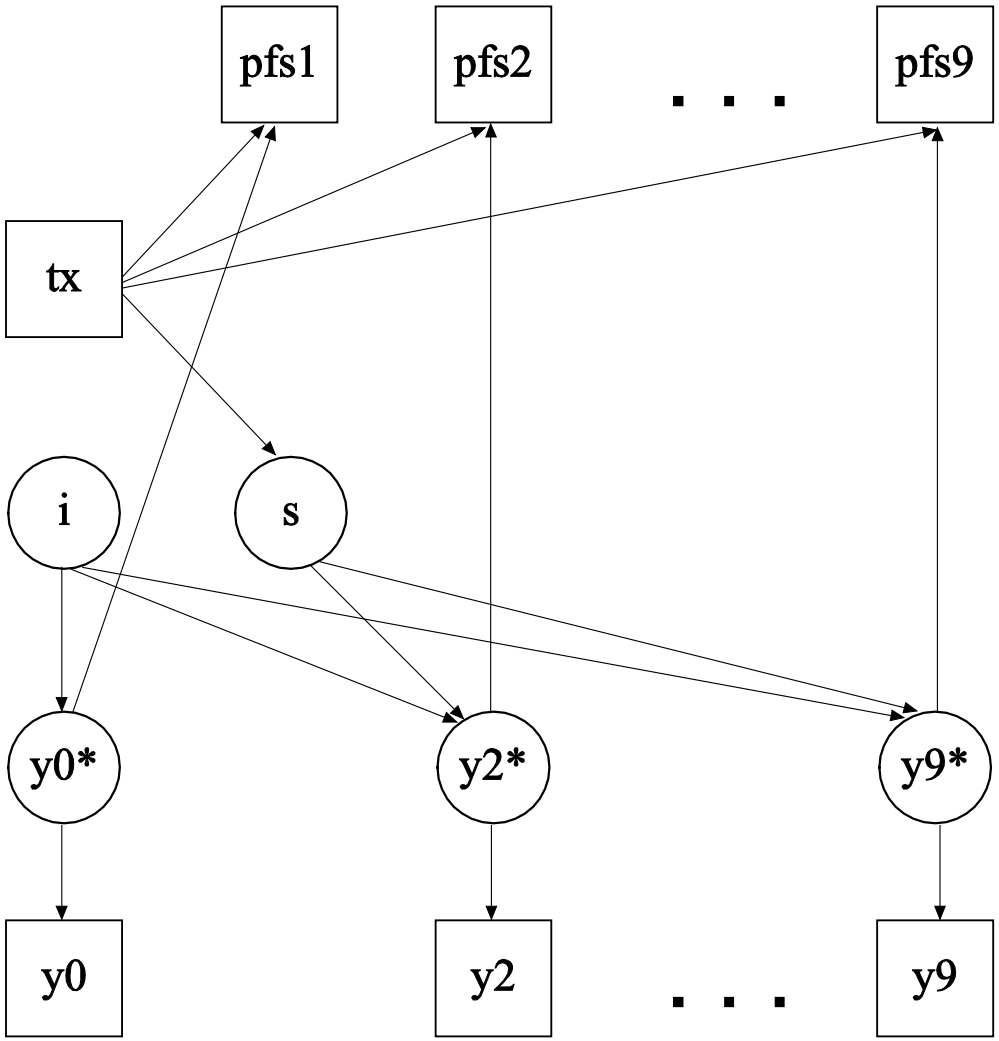
Choice Of Joint Growth-Survival Models, Continued

Computational Aspects

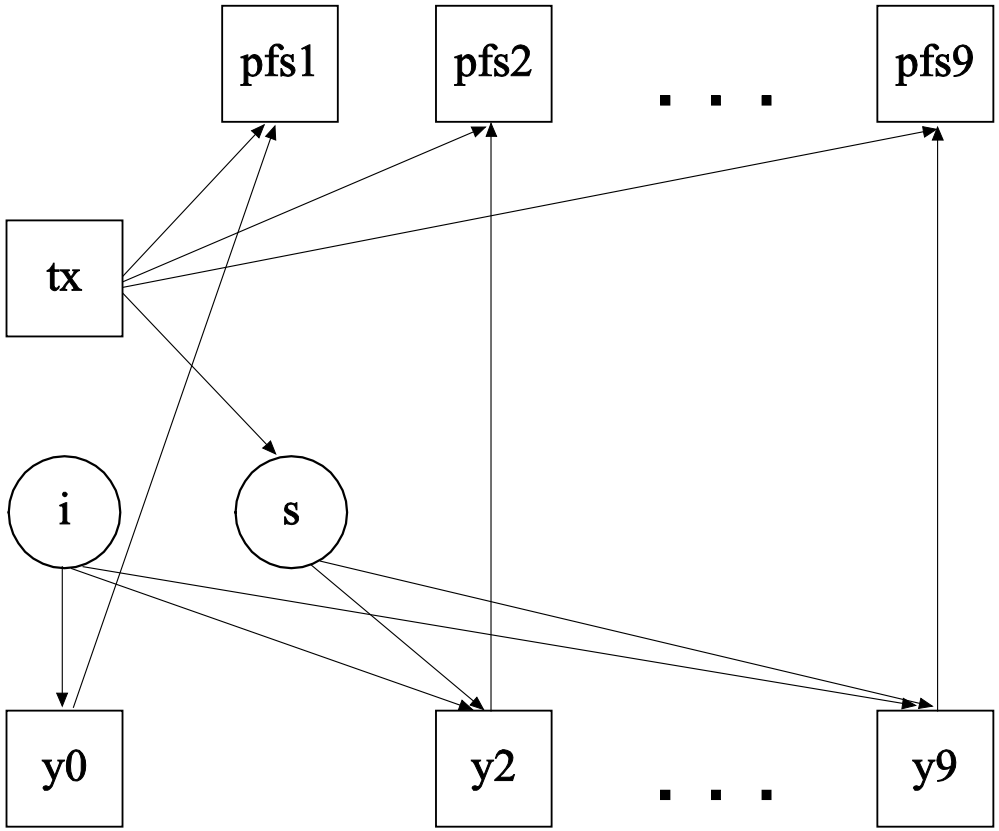
- Model 2. Uses 2 dimensional integration.
- Model 3. Uses Montecarlo integration for intermittent missing values.
- Model 4. No numerical integration.

Model 3 fastest, and Model 2 is the slowest. All three are quite easy to estimates within a couple of minutes.

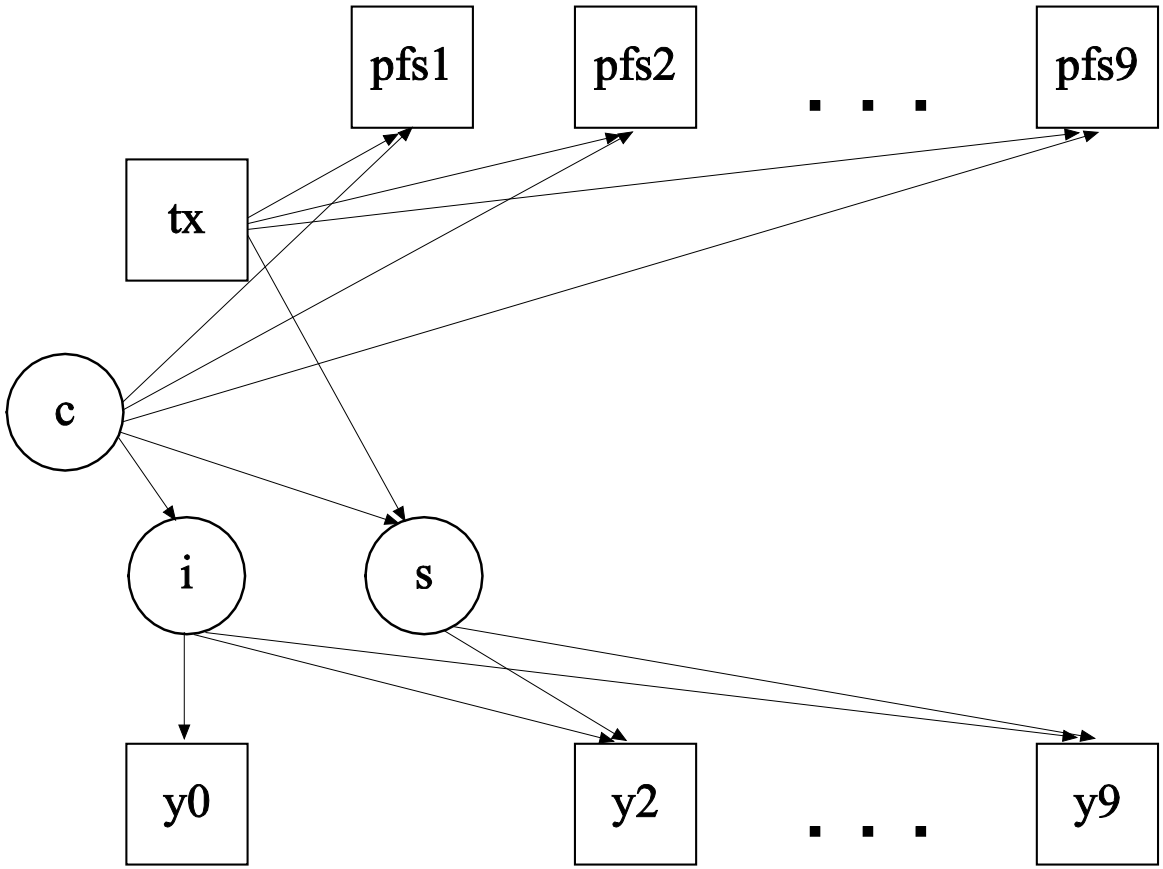
Joint Growth - Survival Mode: Approximate Xu-Zeger



Joint Growth - Survival Model: Observed Xu-Zeger



Joint Growth - Survival Model: Growth Mixtures

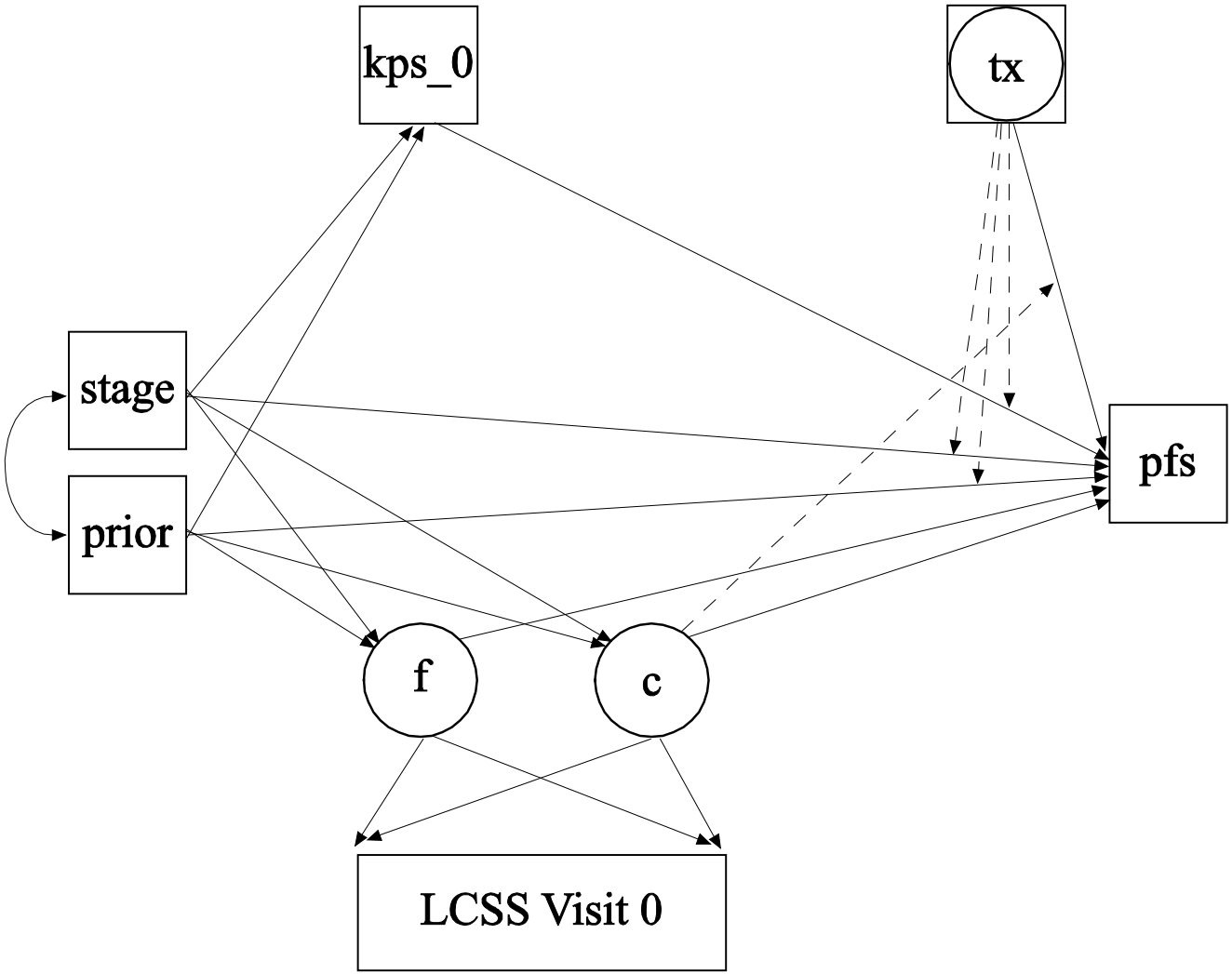


Survival analysis related to development in the three global LCSS items

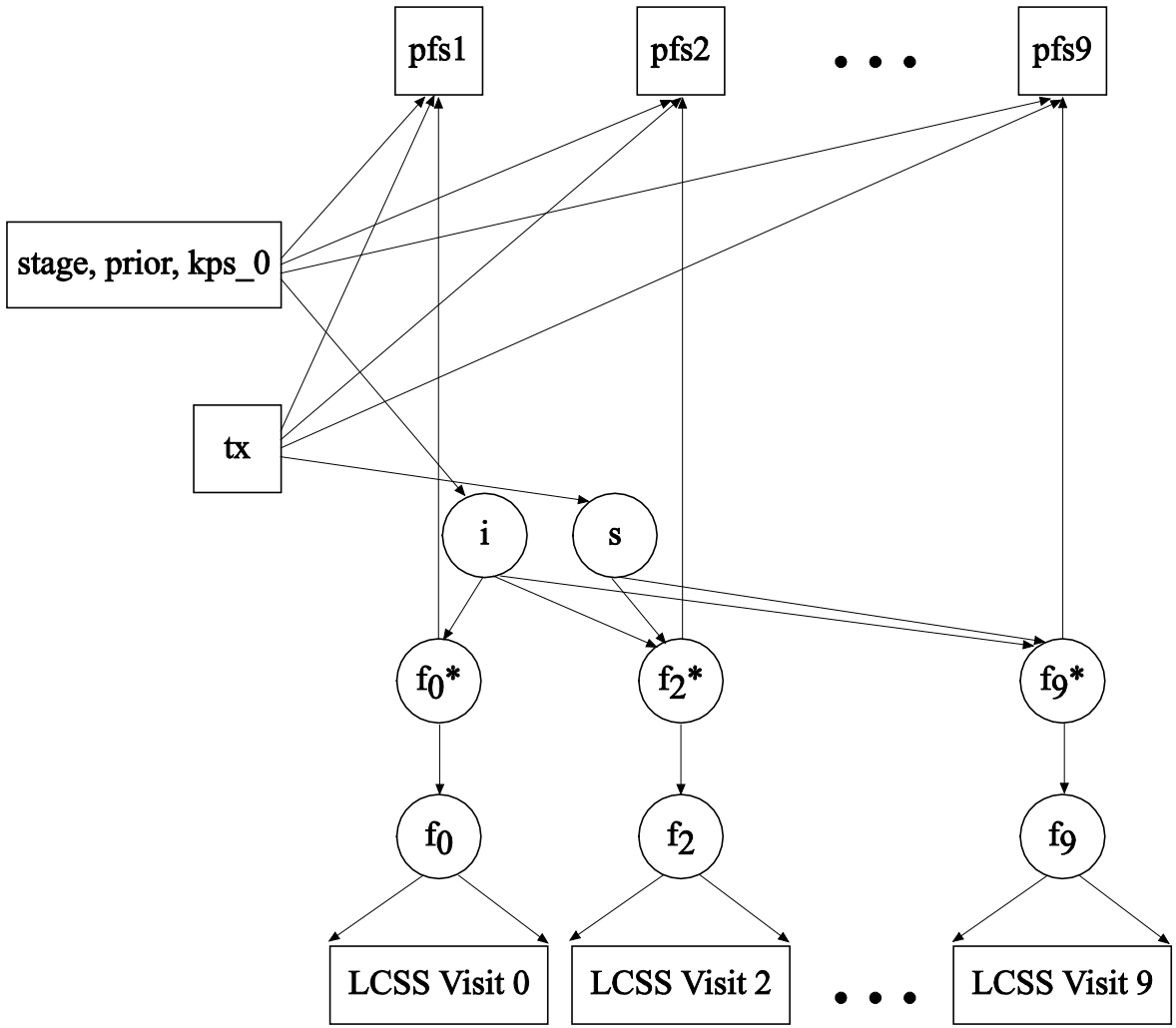
Model	Log-Likelihood	Number of Parameters	BIC	Tx Effect on PFS	Tx Effect on LCSS	LCSS Effect on PFS
Quality of life						
Xu-Zeger	-4615	19	9334	Yes	No	No
Observed	-4610	19	9324	Yes	No	Yes
Growth mixture	-4598	24	9328	Yes	No	Yes
Interference						
Xu-Zeger	-4679	19	9463	Yes	No	No
Observed	-4675	19	9454	Yes	No	Yes
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
Growth mixture	-4639	24	9411	Yes	No	Yes

Latent Variable Survival Modeling

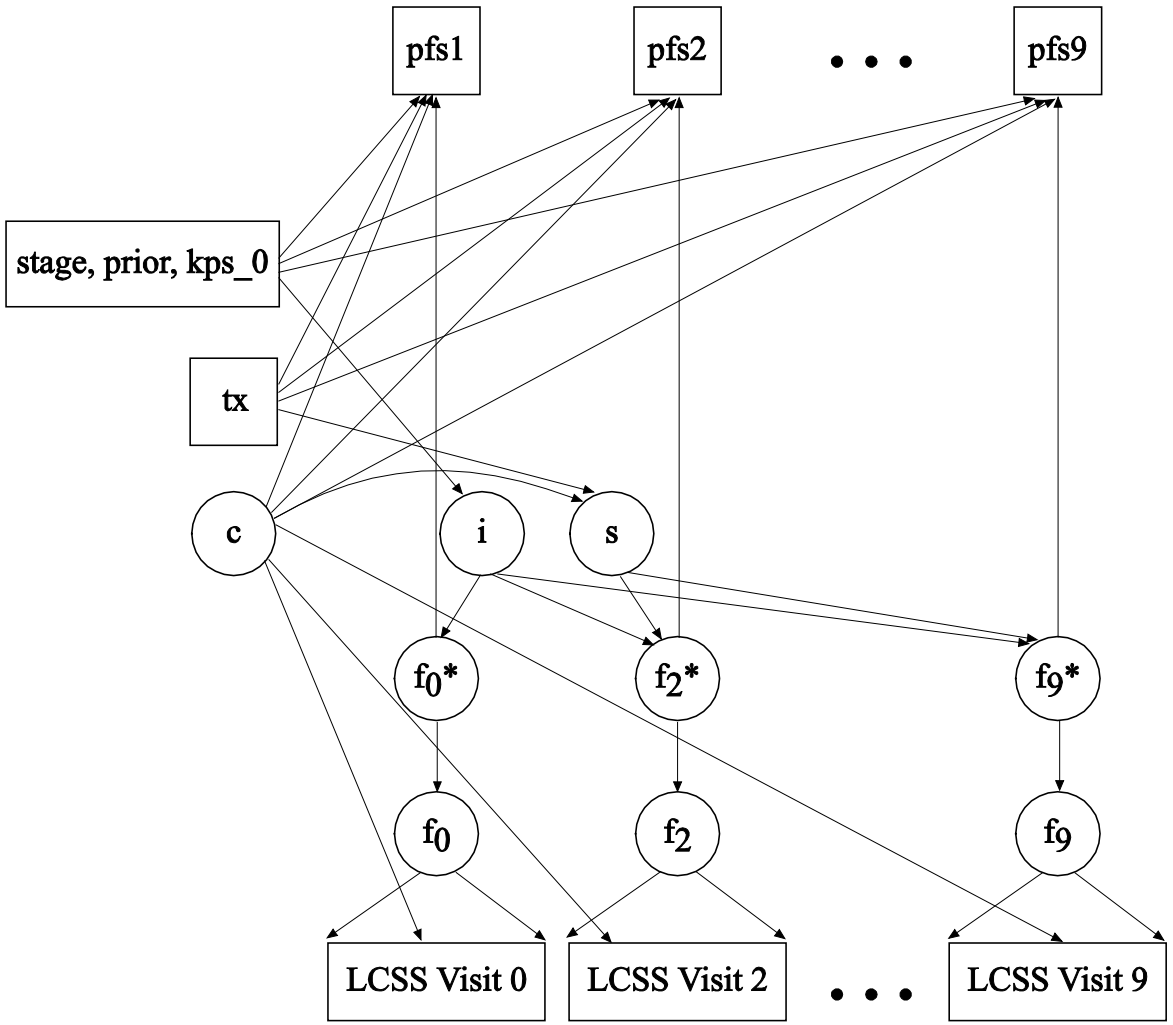
Model 0. Predicting Survival From Visit 0 Using a Factor Mixture Model For LCSS Items



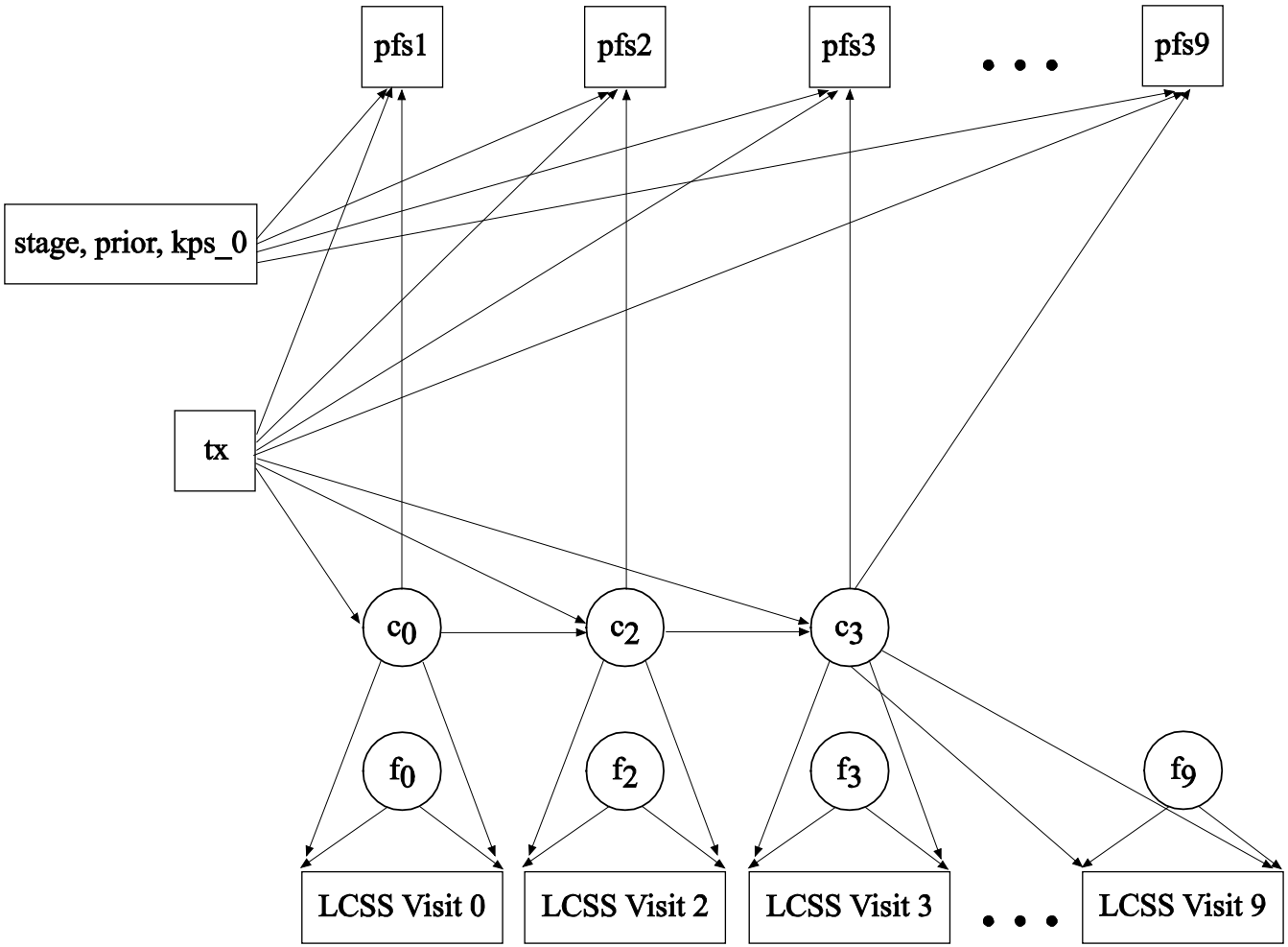
Model 1: Joint Latent Factor Growth Modeling And Survival Analysis



Model 2: Joint Latent Factor Growth Mixture Modeling And Survival Analysis



Model 3: Joint Factor Mixture Latent Transition Analysis And Survival Analysis With Attenuation At Time 3



Conclusions

- Analysis results
  - LCSS useful in predicting progression-free survival
  - LCSS contributes information beyond stage, prior, and Karnofsky
  - Patients with high baseline QOL (low LCSS score) benefit more from treatment
  - LCSS information beyond the baseline is predictive of progression-free survival
- Statistical results
  - Latent variable survival modeling possible in practical applications using Mplus
  - Visit 0 factor mixture model prediction of survival: easy
  - Joint latent growth - survival modeling: easy
  - Joint multiple-indicator factor latent growth - survival modeling: a bit harder
  - Joint latent transition - survival modeling: harder
  - Final thought: Estimated latent variable survival model used as survival prediction instrument for new patients