Estimating time to disease progression comparing transition models and survival methods

Micha Mandel, Francois Mercier, Benjamin Eckert, Peter Chin, and Rebecca A. Betensky

August 2012 / Bergen

Introduction

- Multiple Sclerosis (MS) - a relapsing remitting disease
- Expanded disability status scale (EDSS) - ranges from 0 (normal) to 10 (death due to MS) in 0.5 points

Introduction

- Multiple Sclerosis (MS) - a relapsing remitting disease
- Expanded disability status scale (EDSS) - ranges from 0 (normal) to 10 (death due to MS) in 0.5 points
- Data: two multi-national parallel-group phase 3 trials evaluating oral fingolimod (FTY)
  - The TRANSFORMS trial (12 months, active control, $n = 1292$)
  - The FREEDOMS trial (24 months, placebo control, $n = 1272$)
  - Treatment significantly decreased relapse rate and was approved by FDA and EMA in 2010

Questions of interest

Scientific

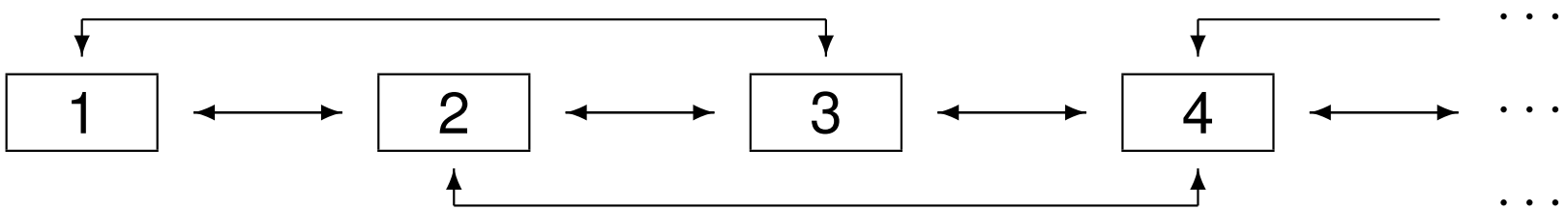
what is the effect of FTY on the EDSS and especially on the probability of confirmed progression (increase in 1 point that lasts for 3 months)?

Methods

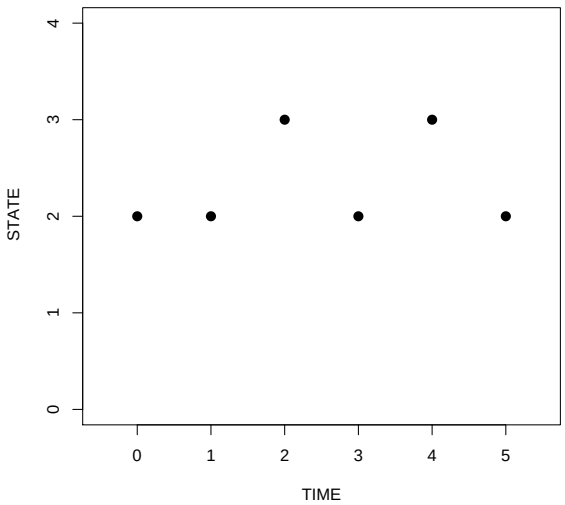
Survival methods vs. transition models

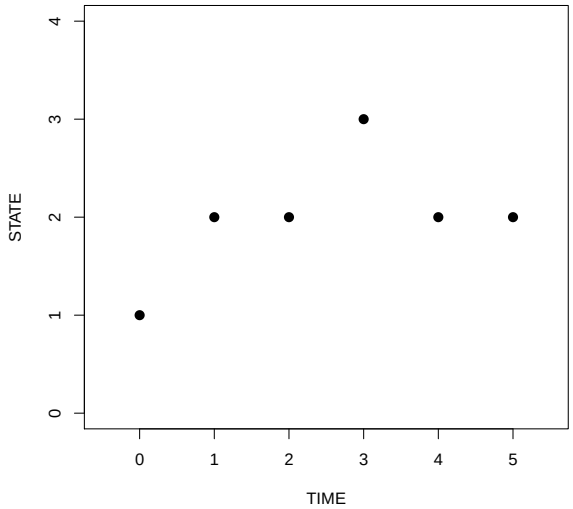
EDSS Data and endpoint

EDSS as a multi-state process (5-band model)



Three-month EDSS data; Dist. of time to confirmed progression (discrete)





Notation

- State space $\{1, 2, \ldots, J\}$
- Y_t state at time t
- $t_0 = 0, \ldots, t_n = n$ observed times (equally spaced)
- X vector of measured covariates

Data for subject i: $(y_{it_0}, y_{it_1}, \ldots, y_{it_{n_i}}, x_i)$

Time to confirmed progression:

$$T_i = \min\{j \mid y_{it_{j-1}} > y_{it_0} \ , \ y_{it_j} > y_{it_0}\}$$

Survival Analysis Methods

Typically, time of follow-up varies between subjects - censoring

T = time of event $\qquad$ C = end of follow-up

data = [$\min(T, C)$, $I\{T \leq C\}$]

Survival Analysis Methods

Typically, time of follow-up varies between subjects - censoring

T = time of event $\qquad$ C = end of follow-up

data = [$\min(T, C)$, $I\{T \leq C\}$]

Estimation and testing using survival methods:

- Kaplan-Meier Product-limit: $\widehat{\text{hazard}}(t) = \frac{\#\text{events at } t}{\#\{T \geq t\}}$
- Cox Proportional hazard: $\widehat{\text{hazard}}(t) = \hat{h}_0(t) \exp(x\hat{\beta})$
- $1 - F(t) = \prod_{s \leq t}[1 - h(s)] \quad \left(= \exp\{-\int_{s \leq t} h(s)ds\} \right)$

Survival Analysis Methods

Pros:

- Estimate the parameter of interest directly - a natural model
- Non/semi-parametric - consistent - robust

Survival Analysis Methods

Pros:

- Estimate the parameter of interest directly - a natural model
- Non/semi-parametric - consistent - robust

Cons:

- For an ordinal measure, $1 \rightarrow 2 \neq 2 \rightarrow 3$ - stratification or modeling are required
- Censored processes are informative
- Non-progressive process - difficulty with missing values

Markov Transition Models

We suggest to model the EDSS process using the Markov assumption - $P(Y_t|\text{History}, X) = P(Y_t|Y_{t-1}, X)$
Time-to-event $\leftrightarrow$ hitting time.

Markov Transition Models

We suggest to model the EDSS process using the Markov assumption - $P(Y_t|\text{History}, X) = P(Y_t|Y_{t-1}, X)$
Time-to-event $\leftrightarrow$ hitting time.

Pros:

- $1 \rightarrow 2$ and $2 \rightarrow 3$ are directly estimated
- Exploits information in censored processes
- Can deal with (non-informative) missing values

Cons:

- The parameter of interest is a derived quantity
- Uses rather strong assumptions - less robust

Discrete Markov Model

Transition probabilities between consecutive times are given by

$$P(x,\theta) = \begin{pmatrix} p_{1,1}(x,\theta) & \cdots & p_{1,J}(x,\theta) \\ \vdots & \vdots & \vdots \\ p_{J,1}(x,\theta) & \cdots & p_{J,J}(x,\theta) \end{pmatrix}$$

Example of a model (partial proportional odds):

$$p_{j,k}(x,\theta) = \frac{\exp(\alpha_{jk} + \beta x)}{1 + \exp(\alpha_{jk} + \beta x)} - \frac{\exp(\alpha_{jk-1} + \beta x)}{1 + \exp(\alpha_{jk-1} + \beta x)}$$

$j, k = 1, \ldots, J,$ $\theta = (\alpha_{11}, \ldots, \alpha_{JJ-1}, \beta),$ $\alpha_{j0} = 0,\ \alpha_{jJ} = \infty$

5-band matrix

We assume $p_{j,k} = 0$ for $|j - k| > 2$

s-step Transition Probabilities

The s-step transition probabilities are given by

$$P(Y_s = k | Y_0 = j, x) = P^s(x, \theta)_{j,k}$$

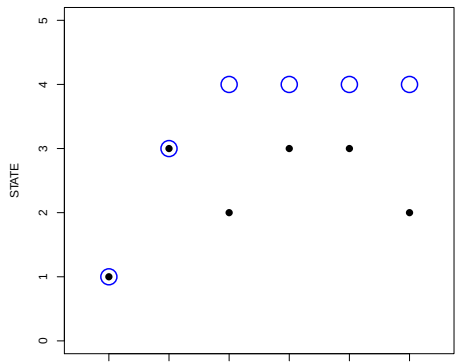
Calculating time-to-events probabilities

by modifying the transition matrix:

Time to $(\{j+1,\ldots,J\},\{j+1,\ldots,J\})$

$$\tilde{P}_j(x,\theta) = \begin{pmatrix} p_{11} & \cdots & p_{1j} & p_{1(j+1)} & \cdots & p_{1J} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{j1} & \cdots & p_{jj} & p_{j(j+1)} & \cdots & p_{jJ} & 0 \\ p_{(j+1)1} & \cdots & p_{(j+1)j} & 0 & \cdots & 0 & \sum_k p_{(j+1)k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{J1} & \cdots & p_{Jj} & 0 & \cdots & 0 & \sum_k p_{Jk} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Prediction: $P(\text{event during } s \text{ steps} | Y_0 = j, x) = \tilde{P}_j^s(x,\theta)_{j,J+1}$



Estimation

Estimation of θ by maximum likelihood

$$\hat{\theta} = \text{argmax} \prod_{i} \prod_{v=1}^{n} p_{y_{iv-1}, y_{iv}}(x_i, \theta)$$

Practically speaking, consider Y_{v-1} as a covariate for Y_v and assume that all transitions are independent. (Really practically speaking, use NLMIXED in SAS or clmm2 in R).

Prediction – by plugging-in $\hat{\theta}$.

Confidence intervals – by delta method or bootstrap-like method.

Results

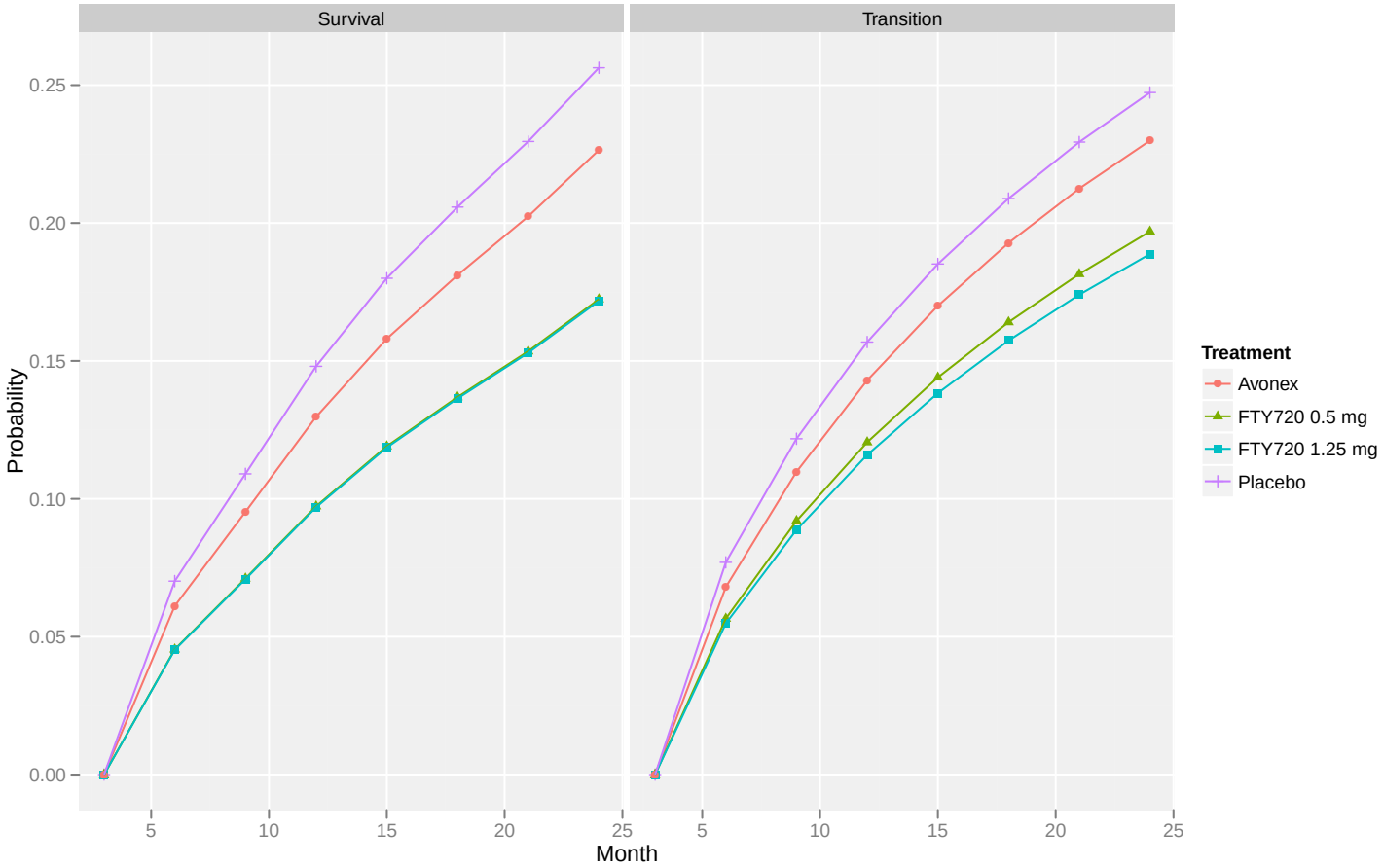
Two trials combined (2,348 subjects, 12,334 transitions)

	Cox Model HR	Markov Mixed OR
MS duration $<$ 5y	0.855	0.626
Low T2.vol base	0.910	0.763
FTY 0.5mg	0.639	0.708
FTY 1.25mg	0.636	0.653
Avonex	0.867	0.879
EDSS.base=[5, 5.5]	1.218	$>$100
EDSS.base=[4, 4.5]	0.636	$>$100
EDSS.base=[3, 3.5]	0.600	$>$100
EDSS.base=[2, 2.5]	0.636	19.1

P-value$<$0.05.

Probability curves

Estimated CDFs of time to sustained progression; Cox model (left) and Random effects Markov model (right).



Covariates: EDSS at baseline 0-1.5, disease duration $<$5 years, lesion volume $<$median

Survival vs. Markov - Significant level

Simulation (thanks to Tom Hope) - 2 groups, 200/200 subjects
$J = 3$ states 9 transitions per subject: $(y_1, y_2, ..., y_9)$

Empirical size of 5% tests for treatment effect

data generated by	Log-Rank	Markov Fixed	Markov Random
1-step Markov-fixed	.062	.052	.048
1-step Markov-random $\sigma^2 = 1$	.044	.086	.044
2-step Markov-fixed	.052	.038	.048
2-step Markov-random $\sigma^2 = 1$	.044	.072	.036
permutation real data	.050	.030	.030

$*$ Transition matrices according to CLIMB - a natural history study on multiple sclerosis.

Survival vs. Markov - power

Treatment effect - $\beta = 0.5$ in PPO model

Simulation (thanks to Tom Hope) - 2 groups, 200/200 subjects
$J = 3$ states 9 transitions for subject: $(y_1, y_2, ..., y_9)$

Empirical power of 5% tests for treatment effect

data generated by	Log-Rank	Markov Fixed	Markov Random
1-step Markov-fixed	.493	.978	.976
1-step Markov-random $\sigma^2 = 1$	.308	.766	.662
2-step Markov-fixed	.336	.884	.840
2-step Markov-random $\sigma^2 = 1$	.276	.678	.606
permutation RR=.55	.691	.066	.066

Summary and Remarks

- Markov model and survival analysis give similar estimates for progression curves, both indicate positive effect of FTY compare to placebo.
- Markov models provide information on the EDSS process, but this depends on more assumptions.
- Power and size of tests depend on the underlying EDSS process.
- Open challenges: flexible modeling of the EDSS process, comparison of power between survival and transition models, joint analysis of relapse and EDSS, curing MS.