Developing a postestimation command for joint models in `merlin`

Nuzhat B Ashra & Michael J Crowther

Biostatistics Research Group, University of Leicester, UK

London Stata Conference 2019

Motivation

- Simultaneously model longitudinally measured urinary hCG and time to miscarriage
- Current software for joint models `stjm` (Crowther et al, 2013) and `merlin` (Crowther, 2018 [submitted])
- Predict conditional survival probabilities from these joint models to allow real-time tracking of pregnancy progress (Rizopoulos, 2011)
- Evaluate model prediction capabilities

GCC dataset

- 368 women, aged 18-45, trying to conceive
- Collected early morning urine samples from the first day of their cycle to up to seven days of the next cycle if they did not become pregnant or up until day 60 if they did become pregnant
- 288 viable pregnancies and 80 miscarried pregnancies

hCG trajectories
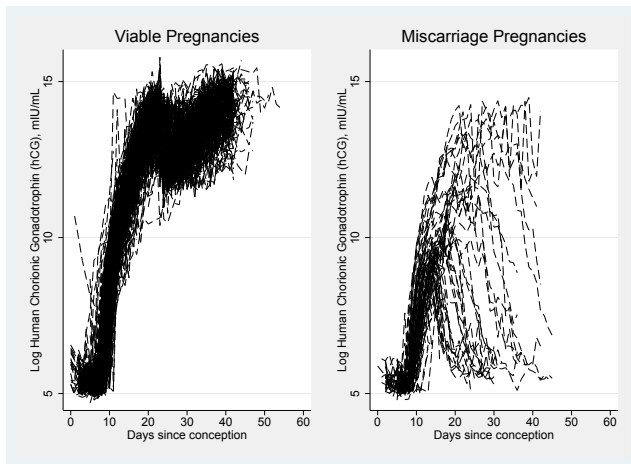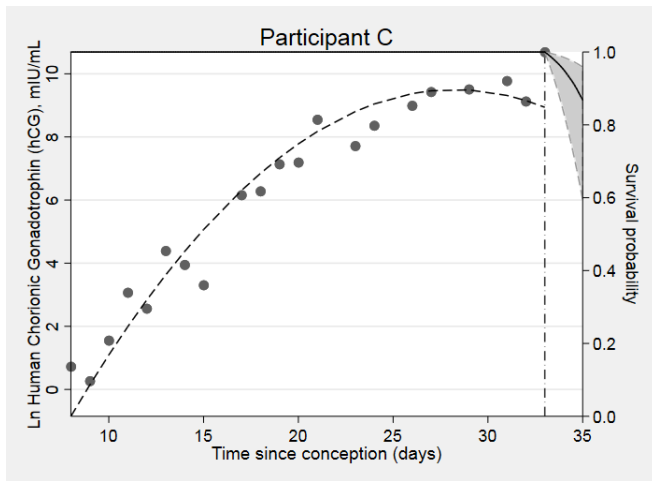
merlin

Unified modelling framework for data analysis

- Designed to be as flexible and general as possible
- Any number of outcome models can be specified, linked in any number of ways

Find examples of models which can be fit by `merlin` at `https://www.mjcrowther.co.uk/software/merlin/tutorials_stata`

Fitting a joint model with `merlin`

```
. merlin (stime trt EV[logb], family(weibull, failure(died))  ///
> timevar(stime))        ///
> (logb fp(time,pow(1)) fp(time,pow(1))#M2[id]@1 M1[id]@1,   ///
> family(gaussian) timevar(time)) ///
> , covariance(unstructured) restartvalues(M2 0.1)
```

$$h_i(t|M_i(t)) = h_0(t) \exp[\phi_1 trt + \alpha m_i(t)]$$

$$m_i(t) = logbilirubin_i(t) = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t$$

Fitting a joint model with `merlin`

```
Mixed effects regression model                    Number of obs    =      1,945
Log likelihood = -1919.2164
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
stime:						
trt	.0441737	.1790899	0.25	0.805	-.3068362	.3951835
EV[]	1.240676	.0932792	13.30	0.000	1.057852	1.4235
_cons	-4.411849	.2741419	-16.09	0.000	-4.949157	-3.874541
log(gamma)	.0193141	.0825814	0.23	0.815	-.1425425	.1811706
logb:						
fp()	.1850394	.0133236	13.89	0.000	.1589256	.2111532
fp()#M2[id]	1	.	.	.	.	.
M1[id]	1	.	.	.	.	.
_cons	.4929444	.0582791	8.46	0.000	.3787195	.6071693
sd(resid.)	.3471211	.0066724			.3342868	.3604481
id:						
sd(M1)	1.002467	.0426595			.9222474	1.089664
sd(M2)	.1808176	.0123978			.1580803	.2068252
corr(M2,M1)	.4252257	.0729127			.2725388	.5570211

.
- HR 3.458 (95% CI: 2.880, 4.152)

Model Discrimination

- Correctly identify those individuals who will experience an event in a defined time period and those who will not
- Discrimination can be assessed by extending the use of the area under the receiver operating characteristic curve (ROC AUC) to the joint model setting (Andrinopoulou et al, 2018; Ferrer et al, 2017)

ROC AUC

Given a randomly selected pair of individuals l_1, l_2 we define the ROC AUC as,

$$AUC(t, \Delta t) = Pr \Big[\pi_{l_1}(t, \Delta t) < \pi_{l_2}(t, \Delta t) \mid \{T_{l_1}^* \in (t, t + \Delta t]\} \cap$$
$$\{T_{l_2}^* > t + \Delta t\} \Big]$$

where

$$\pi_l(t, \Delta t) = Pr(T_i^* \geq t + \Delta t \mid T_i^* > t, \tilde{y}_l(t), \mathbf{D_n}),$$

Identify all pairs

To calculate this we need to identify all pairs of patients,

- l_1 experience an event in the time-frame, l_2 did not
- l_1 was censored during the time-frame, l_2 did not experience an event during the time-frame
- l_1 experienced an event during the time-frame and l_2 was censored after l_1 but before or in the time-frame of interest
- l_1 was censored in the time-frame and l_2 was censored after l_1 but before or in the time-frame of interest

ROC AUC

$$\widehat{AUC_w}(t, \Delta t) = \frac{\sum_{l_1=1}^{n} \sum_{l_2=1; l_2 \neq l_1}^{n} I\{\widehat{\pi_{l_1}}(t, \Delta t) < \widehat{\pi_{l_2}}(t, \Delta t)\} \times I\{\Omega_{l_1 l_2}^{(w)}(t)\} \times \widehat{K_w}}{\sum_{l_1=1}^{n} \sum_{l_2=1; l_2 \neq l_1}^{n} I\{\Omega_{l_1 l_2}^{(w)}(t)\} \times \widehat{K_w}}$$

- For any pair of subjects we want the model to correctly predict a higher survival probability for the individual who did not experience the event when compared to the individual who did

Output from postestimation command

```
.
. /*Time-frame for prediction*/
. gen t0=8

. gen fu=9

.
. merlin_p2 rocauc, rocauc tstart(t0) fu(fu) id(id)

ROC AUC = .71487892

.
```

ROC curve

- Calculate the sensitivity and specificity for various survival probability cut-offs and output a table (`roctab`)
- Produce a graph which plots sensitivity against 1-specificity

Calibration

- The calibration assesses the accuracy of the model, i.e. how well the model predicts the event rates in the data
- Estimate the mean squared prediction error, compares predicted probability of survival of patient to observed event status at time t for each individual and then takes average of the sum (Andrinopoulou et al, 2018; Henderson et al, 2000)

Prediction error

$$\widehat{PE}(t, \Delta t) = \{R(t)\}^{-1} \sum_{I:T \geq t} \Big\{ I(T_I > t + \Delta t)\{1 - \widehat{\pi}_I(t, \Delta t)\}^2$$
$$+ \delta_I I(T_I < t + \Delta t)\{0 - \widehat{\pi}_I(t, \Delta t)\}^2 + (1 - \delta_I)I(T_I$$
$$< t + \Delta t) \times [\widehat{\pi}_I(T_I, \Delta t)\{1 - \widehat{\pi}_I(t, \Delta t)\}^2$$
$$+ \{1 - \widehat{\pi}_I(T_I, \Delta t)\}\{0 - \widehat{\pi}_I \times (t, \Delta t)\}^2] \Big\}$$

- Red number of subjects at risk at t
- Blue are those event free after $t, \Delta t$
- Green experienced the event before $t, \Delta t$
- The final part denotes those censored in $t, \Delta t$

Output from postestimation command

```
.
. merlin_p2 pe, prederror tstart(t0) fu(fu)

Prediction Error = .10149774

.
```

Next Steps

- ROC curve output
- Improve efficiency, move into mata
- Rewrite `stjmcsurv` for `merlin`
- Current focus joint models - make predictions valid for arbitrary `merlin` model, incorporate into `predict`
- Stata Journal paper - `merlin` postestimation

Selected References

[1] Andrinopoulou, Eleni-Rosalina and Eilers, Paul H. C. and Takkenberg, Johanna J. M. and Rizopoulos, Dimitris. 2018. Improved dynamic predictions from joint models of longitudinal and survival data with time-varying effects using P-splines. Biometrics. 72(2):685-693

[2] Crowther, Michael J., and Abrams, Keith R., and Lambert, Paul C. 2013. Joint modeling of longitudinal and survival data. Stata J 13(1):165-184

[3] Crowther, Michael J. 2018. merlin-a unified modelling framework for data analysis and methods development in Stata. arXiv preprint arXiv:1806.01615 [submitted]

[4] Ferrer, Loïc and Putter, Hein and Proust-Lima, Cécile. 2017. Individual dynamic predictions using landmarking and joint modelling: validation of estimators and robustness assessment. Statistical methods in medical research.

Selected References 2

[5]   Henderson, Robin and Diggle, Peter and Dobson, Angela. 2000. Joint modelling of longitudinal measurements and event time data. Biostatistics 1(4):465-480

[6]   Rizopoulos, Dimitris. 2011. Dynamic Predictions and Prospective accuracy in joint models for longitudinal and time-to-event data. Biometrics 67(3):819-829