Joint Modeling Tumor Burden and Time to Event Data in Oncology Trials

Ye Shen, Aparna Anderson, Ritwik Sinha, and Yang Li

ASCPT 2018 Annual Meeting

Orlando, FL

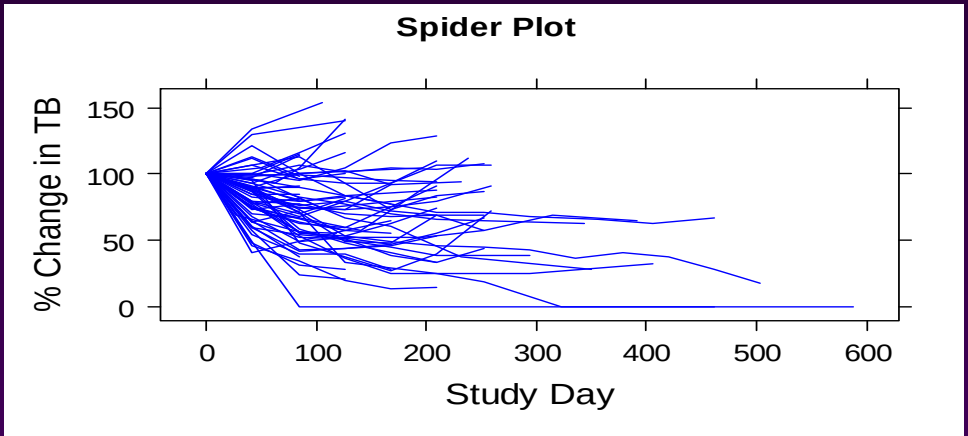
2018.3.23

Introduction

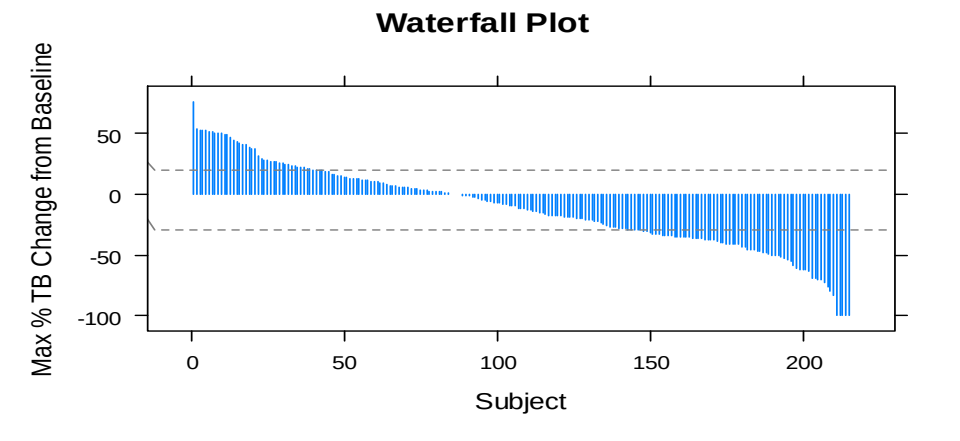
- In oncology, overall survival is the ideal measure of treatment benefit
- However, the mechanistic and biologic effects of a therapeutic agent are generally described in terms of **tumor burden (TB)**, measured repeatedly at protocol-specified time intervals
- Tumor burden is usually categorized (e.g., per RECIST) for the purpose of analyzing objective response rate or **progression-free survival (PFS)**
- The loss of information due to categorization may result in a misrepresentation of the true association between treatment and change in TB

- **Question: How can we fully exploit the complete longitudinal tumor burden data to characterize biological treatment effects?**

Commonly applied methods have important limitations



- Patterns obscured if too many subjects
- Qualitative; no formal inference
- Bias due to lack of follow-up in those who progress or die (non-random missingness)



- A different measure of change (e.g., max % change vs. nadir) might reveal a different pattern
- Qualitative; no formal inference
- Ignores timing, durability, and survival

The nature of missing TB data affects estimation and interpretability

Missing completely at random (MCAR):
Missing value does not depend on observed or unobserved measurements of interest

Example:
Subjects are more likely to miss tumor assessment visits during the holiday season

Loss of power

Missing at random (MAR):
Missing value does not depend on unobserved value itself, but possibly on other factors

Example:
Subjects have missing TB data following death or evidence of disease progression

Bias

Loss of power

Missing not at random (MNAR):
Missing value depends on unobserved value itself

Example:
Subjects with high TB are more likely to miss tumor assessment visits due to impaired mobility

Loss of efficiency

A more robust method is required

- We want a model that
  - uses the **longitudinal TB data**
  - accounts for **MAR and possible MNAR** due respectively to PFS and a possible latent missing mechanism
  - allows **formal inference** about the association between treatment and TB

- We want to evaluate how such a model performs
  - with respect to bias and variance
  - under linear and nonlinear tumor burden distributions

  as compared to
  - a gold standard model fitted to fully observed data (i.e., no missingness)
  - a simpler model that requires missing data to be (at worst) MAR

Construct a model that jointly captures TB and PFS

- **Linear mixed effects (LME)** model for longitudinal TB data

  $TB(t) = covariate(t)*\beta_1 + RE_1(t) + error$

- **Event time (ET)** model for progression-free survival

  $hazard(t) = \exp\{covariate(t)*\beta_2 + RE_2\}$

  PFS is defined as the time from study entry to disease progression or death, whichever occurs first

- **Joint model (JM)**: Introduce correlation between random effects (RE_1, RE_2):

  $RE_1(t) = U_1 + U_2*t$ and $RE_2 = \pi_1*U_1 + \pi_2*U_2$

  where U_1=random intercept, U_2=random slope, and π_1,π_2 = fixed coefficients

  NOTE: Non-zero π_1 and π_2 produces MNAR data

LME alone offers a simple alternative to the more complex JM, but LME gives biased results if MNAR conditions exist (Schafer and Graham, 2000).

Estimate joint model parameters

- A joint likelihood function can be constructed for the LME and ET components of the JM
  - L_Y = likelihood of longitudinal TB process
  - $L_{N|Y}$ = conditional likelihood of event process for progression-free survival, given TB data
  - $L = L_Y \times L_{N|Y}$ = joint likelihood function (A) (Henderson et al., 2000)

- Apply algorithm and SAS program developed by Guo and Carlin (2004) to maximize joint likelihood

Simulation 1: JM operating characteristics under linear TB

- Apply (B) to simulate data:

$$TB(t) \sim b_{L0} + b_{L1}*t + b_{L2}*t*Trt + U_1 + U_2*t + error$$
$$Log(PFS\ hazard) \sim b_{S0} + b_{S1}*Trt + p_1*U_1 + p_2*U_2 \qquad (B)$$

where $U_1 \sim N(0, v_{11})$, $U_2 \sim N(0, v_{22})$, $error \sim N(0, 1)$, and
b_{L0}, b_{L1}, b_{L2}, b_{S0}, b_{S1} are fixed at 0.4, 0.1, -0.2, -4, -1, respectively

- TB values q6wks in t=[0,48], n=300, replications=200

- Missing mechanism for data generated from (B):
  - If event occurs on or before t=48, set TB data as missing after event time
  - If event occurs after t=48, no missing TB and censor PFS at t=48

Simulation 1: Linear TB, missing data are MAR

	Parameter estimates (SD) for b_{L0}, b_{L1}, and b_{L2} from TB model					
	FLME	PLME	PJM	FLME	PLME	PJM
True parameters	$p_1 = 0$, $p_2 = 0$ (MAR)					
	$v_{11} = 0.1$, $v_{22} = 0.05$			$v_{11} = 1$, $v_{22} = 0.05$		
$b_{L0} = 0.4$	0.40(0.029)	**0.40(0.028)**	**0.39(0.033)**	0.40(0.070)	**0.40(0.071)**	**0.40(0.072)**
$b_{L1} = 0.1$	0.10(0.005)	**0.10(0.006)**	**0.10(0.006)**	0.10(0.005)	**0.10(0.006)**	**0.10(0.007)**
$b_{L2} = -0.2$	-0.20(0.005)	**-0.20(0.006)**	**-0.20(0.007)**	-0.20(0.005)	**-0.20(0.006)**	**-0.20(0.007)**
	$v_{11} = 0.1$, $v_{22} = 0.5$			$v_{11} = 1$, $v_{22} = 0.5$		
$b_{L0} = 0.4$	0.40(0.039)	**0.40(0.038)**	**0.40(0.038)**	0.40(0.061)	**0.40(0.066)**	**0.41(0.068)**
$b_{L1} = 0.1$	0.09(0.055)	**0.09(0.057)**	**0.09(0.057)**	0.09(0.061)	**0.08(0.059)**	**0.09(0.060)**
$b_{L2} = -0.2$	-0.19(0.036)	**-0.19(0.037)**	**-0.19(0.038)**	-0.20(0.054)	**-0.20(0.051)**	**-0.20(0.053)**

FLME = LME fitted on full data (ideal) PLME = LME fitted on partial data due to MAR PJM = Joint Model fitted on partial data

<u>Conclusion</u>: Under MAR, JM produces unbiased estimates and shows minimal loss of efficiency compared to correctly specified LME approach.

Simulation 1: Linear TB, missing data are MNAR

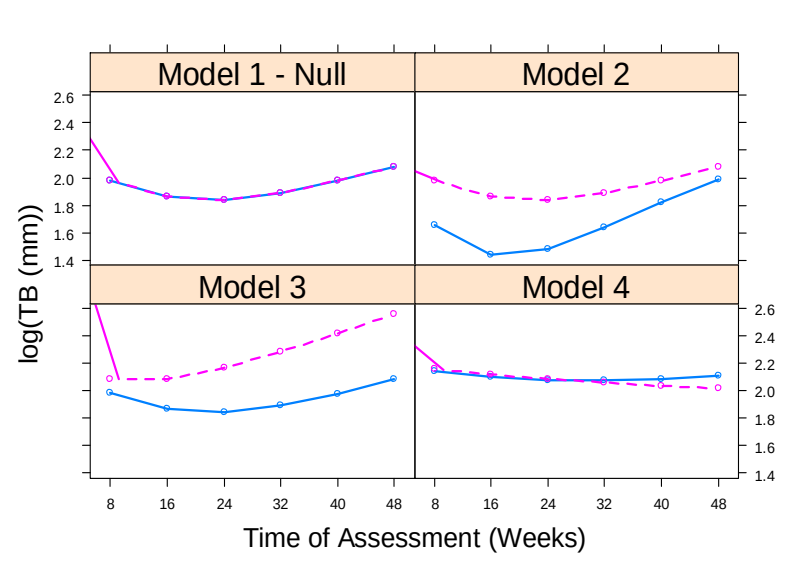
	Parameter estimates (SD) for b_{L0}, b_{L1}, and b_{L2} from TB model					
	FLME	PLME	PJM	FLME	PLME	PJM
True parameters	$p_1 = -1.5$, $p_2 = 2$ (MNAR)					
	$v_{11} = 0.1$, $v_{22} = 0.05$			$v_{11} = 1$, $v_{22} = 0.05$		
$b_{L0} = 0.4$	0.40(0.033)	**0.39(0.037)**	**0.39(0.039)**	0.39(0.070)	**0.33(0.071)**	**0.40(0.072)**
$b_{L1} = 0.1$	0.10(0.006)	**0.10(0.006)**	**0.10(0.007)**	0.10(0.007)	**0.10(0.008)**	**0.10(0.008)**
$b_{L2} = -0.2$	-0.20(0.004)	**-0.20(0.005)**	**-0.20(0.006)**	-0.20(0.004)	**-0.20(0.005)**	**-0.20(0.006)**
	$v_{11} = 0.1$, $v_{22} = 0.5$			$v_{11} = 1$, $v_{22} = 0.5$		
$b_{L0} = 0.4$	0.40(0.035)	**0.39(0.038)**	**0.40(0.039)**	0.40(0.062)	**0.29(0.066)**	**0.40(0.069)**
$b_{L1} = 0.1$	0.10(0.054)	**0.05(0.056)**	**0.09(0.061)**	0.10(0.052)	**0.06(0.059)**	**0.10(0.060)**
$b_{L2} = -0.2$	-0.19(0.038)	**-0.11(0.040)**	**-0.19(0.044)**	-0.20(0.042)	**-0.09(0.046)**	**-0.20(0.052)**

FLME = LME fitted on full data (ideal) PLME = LME fitted on partial data due to MNAR PJM = Joint Model fitted on partial data

<u>Conclusion</u>: Under MNAR and large inter-subject variability, JM reduces estimation bias compared to misspecified LME model.

Simulation 2: JM operating characteristics under non-linear TB

- Consider non-linear TB distribution (Wang et al., 2009)
- Non-linear TB curves reflect an initial dip due to tumor shrinkage, followed by an increase (progression) as the treatment loses its effectiveness
- MNAR mechanism; TB values q8wks in t=[0,48]; n=300; 200 replications



TB data generated under 4 scenarios:

- Model 1: No treatment effect (null model)
- Model 2: TB shrinks more quickly in one arm, but progression occurs at the same rate in both arms
- Model 3: Both arms have the same shrinkage rate, but different rates of progression
- Model 4: One arm has slow steady decrease in TB (phenomenon seen with immunotherapies), while other has progression

Simulation 2:
Non-linear TB, missing data are MNAR

Parameter estimates (Bias) for b_{L0}, b_{L1}, and b_{L2} from TB model						
	FLME †	Bias(PLME)	Bias(PJM)	FLME †	Bias(PLME)	Bias(PJM)
	Model 1 (Null condition – no difference)			**Model 2 (Different shrinkage rates)**		
$\hat{b}_{L0}$	1.37	**1.17**	**0.80**	1.24	**0.75**	**0.51**
$\hat{b}_{L1}$	-0.002	**0.137**	**0.089**	-0.002	**0.060**	**0.059**
$\hat{b}_{L2}$	0	**0.017**	**0.004**	-0.010	**0.048**	**0.019**
	Model 3 (Different progression rates)			**Model 4 (Stability vs. progression)**		
$\hat{b}_{L0}$	1.45	**1.24**	**0.84**	1.00	**0.90**	**0.65**
$\hat{b}_{L1}$	0.005	**0.151**	**0.095**	-0.006	**0.160**	**0.084**
$\hat{b}_{L2}$	-0.001	**0.021**	**0.004**	0.014	**0.112**	**0.022**

† FLME is a misspecified model in Simulation 2 and therefore not the ideal reference. Convergence issues with non-linear methods prevented calculation of gold standard estimates.

Conclusion: With estimates from FLME as points of reference, JM introduces less bias than LME by accounting for latent MNAR process.

Two-stage piecewise linear mixed effects model

- Longitudinal data generated following the non-linear TB component suggested by Wang et. al.(2009). Event defined as 20% increase in TB from baseline or nadir.
- To account for the non-linear trajectory in TB, the two stages were suggested by previous simulation results from the figure.

- **Stage 1: t = [0, 24]**

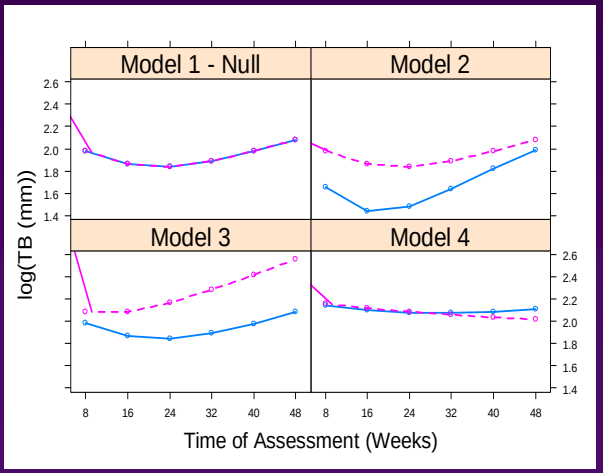
  $TB(t) \sim b_{L01} + b_{L11}*t + b_{L21}*t*Trt + U_{11} + U_{21}*t + error$

  $Log(PFS\ hazard) \sim b_{S01} + b_{S11}*Trt + p_{11}*U_{11} + p_{21}*U_{21}$

Stage 2: t = (24, 48]

$TB(t) \sim b_{L02} + b_{L12}*t + b_{L22}*t*Trt + U_{12} + U_{22}*t + error$

$Log(PFS\ hazard) \sim b_{S02} + b_{S12}*Trt + p_{12}*U_{12} + p_{22}*U_{22}$



- In practice, one may need to check the longitudinal process pattern before making a decision on whether a piecewise linear approximation is needed..

Simulation results from the Two-stage model under MNAR (Stage1)

Parameter estimates (Bias) for b_{L0}, b_{L1}, and b_{L2} from TB model						
	FLME †	Bias(PLME)	Bias(PJM)	FLME †	Bias(PLME)	Bias(PJM)
	Model 1 (Null condition – no difference)			**Model 2 (Different shrinkage rates)**		
$\hat{b}_{L01}$	2.136	**0.017**	**0.011**	2.111	**0.023**	**0.017**
$\hat{b}_{L11}$	-0.016	**4.7E-3**	**3.2E-3**	-0.016	**5.9E-3**	**4.0E-3**
$\hat{b}_{L21}$	-6.7E-4	**8.2E-4**	**6.4E-4**	-6.2E-3	**1.3E-3**	**1.2E-3**
	Model 3 (Different progression rates)			**Model 4 (Stability vs. progression)**		
$\hat{b}_{L01}$	2.132	**0.019**	**0.013**	2.169	**5.5E-3**	**4.3E-3**
$\hat{b}_{L11}$	-0.017	**4.8E-3**	**3.2E-3**	-7.3E-3	**3.3E-4**	**1.4E-4**
$\hat{b}_{L21}$	0.012	**1.6E-3**	**1.2E-3**	-0.021	**2.2E-3**	**2.0E-3**

† FLME is estimated through two-stage piecewise linear approximation.

Simulation results from the Two-stage model under MNAR (Stage2)

Parameter estimates (Bias) for b_{L0}, b_{L1}, and b_{L2} from TB model						
	FLME †	Bias(PLME)	Bias(PJM)	FLME †	Bias(PLME)	Bias(PJM)
	Model 1 (Null condition – no difference)			**Model 2 (Different shrinkage rates)**		
$\hat{b}_{L02}$	1.502	**0.034**	**0.011**	1.433	**0.037**	**0.012**
$\hat{b}_{L12}$	6.4E-3	**4.1E-3**	**1.4E-3**	6.4E-3	**4.0E-3**	**1.3E-3**
$\hat{b}_{L22}$	-2.6E-4	**9.1E-4**	**3.5E-4**	2.0E-3	**1.3E-3**	**5.4E-4**
	Model 3 (Different progression rates)			**Model 4 (Stability vs. progression)**		
$\hat{b}_{L02}$	1.570	**0.037**	**0.011**	1.701	**0.013**	**4.8E-3**
$\hat{b}_{L12}$	5.7E-3	**4.2E-3**	**1.4E-3**	-5.5E-3	**4.4E-4**	**8.8E-5**
$\hat{b}_{L22}$	2.7E-3	**1.0E-3**	**4.3E-4**	8.0E-3	**2.6E-3**	**1.2E-3**

† FLME is estimated through two-stage piecewise linear approximation.

Conclusion:
More reduction in Bias observed in Stage2 due to more drop-outs.

Comparisons between one-stage and two-stage joint models

- **Goodness-of-fit Comparisons between one-stage and two-stage joint models**

	Model 1		Model 2		Model 3		Model 4	
	one-stage	two-stage	one-stage	two-stage	one-stage	two-stage	one-stage	two-stage
AMSE*	0.048	0.016	0.061	0.024	0.046	0.017	0.024	0.007

***AMSE is the averaged mean squared error between the observed and predicted main outcomes**

Conclusions

- Under MAR and linear TB, JM produces unbiased estimates and shows minimal loss of efficiency compared to the correctly specified LME
- Under MNAR, linear TB, and large inter-subject variability, JM reduces estimation bias compared with the misspecified LME
- Under MNAR and non-linear TB distributions, JM appears to have more favorable operating characteristics compared to LME
- Further exploration is required to incorporate non-linear TB modeling and corresponding estimation algorithms

References

Guo, X. and Carlin, B.P. (2004). "Separate and Joint Modeling of Longitudinal and Event Time Data Using Standard Computer Packages," *The American Statistician*, 58, 16-24

Henderson, R., Diggle, P. J., and Dobson, A. (2000). "Joint Modeling of Longitudinal Measurements and Event Time Data," *Biostatistics*, 1, 465-480

Schafer, J. L. and Graham, J. W. (2002). "Missing data: Our view of the state of the art," *Psychological Methods*, 7, 147-177

Wang, Y. et al. (2009). "Elucidation of relationship between tumor size and survival in non-small cell lung cancer patients can aid early decision making in clinical drug development," *Clin. Pharmacol. Ther*, 86, 167-174

Thank you!

Appendix: Wang et. al. (2009)'s non-linear TB model

- Wang et. al. (2009) validated this model in four Phase III studies and provided estimates of the parameters. This relatively flexible model offers a wide range of plausible trajectories for TB data.

- The following notation is used to explain the model.

  $TSO_i(t)$ = Observed tumor burden of subject i at time t

  $TS_i(t)$ = Tumor burden of subject i at time t

  $Base_i$ = Baseline tumor burden of subject i

  SR_i = Tumor Shrinkage Rate of subject i

  PR_i = Progression Rate of subject i

  ϵ = Random measurement error

- Then the no-linear longitudinal TB model is defined by the equations:

  $$TS_i(t) = Base_i \times \exp(-SR_i \times t) + PR_i \times t$$

  $$TSO_i(t) = TS_i(t) \times \exp(\epsilon)$$