Model Tumor Pattern and Compare Treatment Effects Using Semiparametric Linear Mixed-Effects Models

Changming Xia[1], Jianrong Wu[2] and Hua Liang[1]*

[1]Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, NY 14642, USA
[2]Department of Biostatistics, St Jude Children's Research Hospital, Memphis, TN 38105, USA

Abstract

To analyze responses of solid tumor to treatments and to compare treatment effects with antitumor therapies, we applied semiparametric mixed-effects models to fit tumor volumes measured over a period. The population and individual nonparametric functions were approximated by smoothing spline. We also proposed an intuitive method for a comparison of the antitumor effects of two different treatments. Biological interpretation was also discussed.

Introduction

In anticancer drug development, demonstrating the antitumor activity of anticancer agents in preclinical animal model is important. Tumor volume is a commonly used endpoint of treatment efficacy in the evaluation of antitumor agents in such a preclinical animal tumor model. Intuitively, tumor volumes of animals treated on different antitumor agents may be used to compare the antitumor activity of the treatments. Appropriate analysis of tumor volume is therefore important in anticancer drug development. Survival analysis based on the tumor growth delay [1-3] is often conducted but it sometimes provides insufficient information or even invalid comparison of two treatments when both survival times are the same but tumor volumes are different. Another endpoint is tumor growth inhibition [2-5] that is generally assessed at a pre-specified time point. These approaches give valid results but can be inefficient because the information at other time points is discarded. An alternative approach is to fit tumor growth curves, such as multivariate analysis, regression modeling [6-8]. More recently, Liang and Sha [9] applied a parametric nonlinear mixed-effects model [10,11] to analyze changes in tumor volume. To reduce the model's assumptions and make the methods more general and robust, Liang [12] proposed a nonparametric method to model tumor volume. Although these approaches use the entire dataset, each has their own limitations. For example, parametric mixed-effects models [13-15] impose strong assumptions on underlying biology mechanisms and might produce coefficients with limited biological relevance [16], whereas nonparametric mixed-effects models impose no assumptions and may lose useful information when some information is available. In this paper, we suggest a compromised strategy and propose a semiparametric linear mixed-effects (SLM) model to fit tumor volumes.

This research is motivated by the data from a drug combination tumor xenograft study generated in the Pediatric Preclinical Testing Program (PPTP) [17]. In this study, the human rhabdomyosarcoma cell line Rh30 was used to evaluate the therapeutic enhancement for the combination of rapamycin with cytotoxic agents. A total of 140 SCID female mice were used to propagate subcutaneously implanted Rh30 tumors. After tumors grew to a certain size, tumor-bearing mice were randomized into 14different treatment groups with 10 mice per group. Cytotoxic agents were administered at their maximum tolerated dose (MTD), 0.5MTD or 0.63MTD or without concomitant rapamycin treatment. All mice were treated for 6 weeks and followed for another 6 weeks without any treatment. The volume of each tumor is measured

at the initiation of the study and weekly up to 12 weeks. Mice were euthanized usually when the tumor volume reaches four times its initial volume, thus resulting in incomplete longitudinal tumor volume data, as shown in Figure 1.

We would like to establish statistical significance of between-group differences in growth profiles and investigate the underlying biology. It is desirable to have interpretable parameters that represent characteristics of the growth curves, such as slope interpreted as the tumor growth rate. We also want the model to be flexible enough to allow different shapes of the curves. As in Figure 1, there is generally an upward/downward trend for a given group, but the growth patterns seem to be non-linear with different patterns among the groups. Straight-line regression models are likely to underfit the data. Polynomial regression or nonlinear models fit the observations better, but coefficient estimates can be sensitive to nonlinearity assumptions that one cannot evaluate robustly from the dataset in hand. In this case, it is reasonable to use a class of semiparametric models that keep the trend modeled parametrically, while letting the rest of the model be driven by data nonparametrically. Thus we may take advantage of both the flexibility of nonparametric models and interpretability and parsimony of parametric models.

We use smoothing spline to fit the nonparametric component, which was initially developed for smooth interpolation [18]. Under statistical context it is more appealing to fit curves that pass near the noisy data, but are not restricted to interpolating exactly. The optimum curve under certain criteria can be found by solving a penalized least squares problem [19], which is discussed in the later Section. The idea is to find a curve that is a good compromise between fidelity and smoothness. A semiparametric model combining linear predictors and smoothing splines can be written in the form of Linear Mixed Effect models (LME) [20], which enables utilization of theory and computation power in the

---

*Corresponding author: Hua Liang, Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, New York 14642, USA, E-mail: hliang@bst.rochester.edu
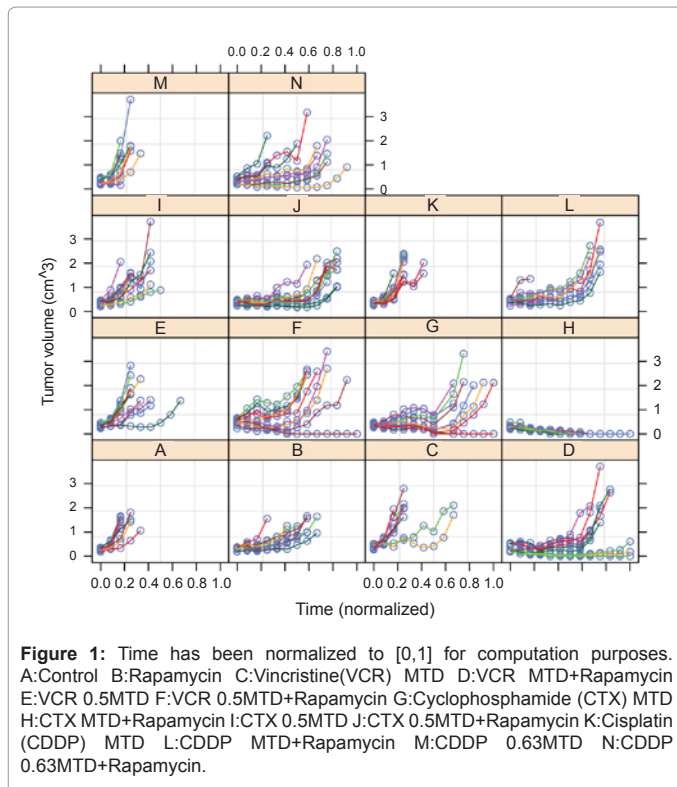
Figure 1: Time has been normalized to [0,1] for computation purposes. A:Control B:Rapamycin C:Vincristine(VCR) MTD D:VCR MTD+Rapamycin E:VCR 0.5MTD F:VCR 0.5MTD+Rapamycin G:Cyclophosphamide (CTX) MTD H:CTX MTD+Rapamycin I:CTX 0.5MTD J:CTX 0.5MTD+Rapamycin K:Cisplatin (CDDP) MTD L:CDDP MTD+Rapamycin M:CDDP 0.63MTD N:CDDP 0.63MTD+Rapamycin.

field. Subject effects and within-subject variations can also be naturally accounted for under this LME framework.

The paper is organized as follows. In Section 2, the SLM model is reviewed and the detail about constructing corresponding models is given. In Section 3, we apply this method to data from a cancer treatment trial and analyze the results, followed by a conclusion section.

Model and Method

The general semi-parametric linear mixed-effects model assumes that [21]

$$y = f + X\beta + Zb + \varepsilon \tag{1}$$

where $y = (y_1,...,y_n)^T$ are responses; $f = (f(t_1),...,f(t_n))^T$ is an unknown function of an independent variable t with $t \in T$ and $f \in H$, a given Reproducing Kernel Hilbert Space (RKHS), X is the design matrix for some fixed effects with parameters β, Z is the design matrix for some random effects b, $b \sim N(0, \sigma^2 D)$; and random errors $\in \sim N(0, \sigma^2 G)$.

H can be decomposed orthogonally as

$$H = H_{null} + H_{penalized}, \tag{2}$$

where H_{null} is a finite dimensional RKHS spanned by basis functions $\phi_1(t),...,\phi_M(t)$, and $H_{penalized}$ is also a RKHS with reproducing kernel $R_{penalized}(s,t)$. Functions in H_{null} are not penalized and "preferred" over functions in $H_{penalized}$. For more details on the topic see Aronszajn [22] and Wahba [23]. The estimate of f can be found by minimizing the following penalized sum of squared errors:

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - f(t_i) - x^T\beta)^2 + \lambda \left\| P_{penalized}f \right\|^2, \tag{3}$$

where $P_{penalized}$ is the orthogonal projection operator of f from $\underline{H}$ onto the penalizing space $H_{penalized}$. As a special case, cubic spline penalizes roughness by letting $P_{penalized}f = f''$, so linear functions are not penalized. λ is the smoothing parameter controlling the balance between fidelity to original data and departure from non-penalizing space H_{null}. λ can be estimated by automatic methods such as Generalized Cross-Validation (GCV), Unbiased Risks (UBR) or Generalized Maximum Likelihood (GML), and it is treated as constant once estimated. The three estimates behave similarly for large sample size [21]. GML is used to estimate the smoothing parameter in this paper.

The function that minimizes (3) has the form

$$\hat{f}_\lambda(t) = \sum_{i=1}^{M} d_i\phi_i(t) + \sum_{j=1}^{n} c_j\xi_j(t) \tag{4}$$

where the coefficients $d = (d_1,...,d_M)^T$ and $c = (c_1,...,c_n)^T$ are solutions to

$$(\Sigma + n\lambda I)c + Td = y, \tag{5}$$

$$T^T c = 0,$$

where $\xi_j(t) = R_{penalized}(t,t_j)$, $T_{n\times M} = \{\phi_i(t_j)\}_{i=1\,j=1}^{M\,n}$ and $\Sigma = \{<\xi_i,\xi_j>\}_{i,j=1}^{n}$.

A more general form of (3) enables modeling of the covariance matrix using a weight matrix W, and different smoothing components with multiple RKHS decompositions $H_k = H_{k,null} + H_{k,penalized}$:

$$\frac{1}{n}(y - f - X\beta)^T W(y - f - X\beta) + \lambda\sum_{k=1}^{p}\theta_k^{-1}\left\| P_{k,penalized}f_k \right\|^2. \tag{6}$$

Minimizer to (6) is a simple extension to (4) and (5), and can be directly related [24,25] to the Restricted Maximum Likelihood (REML) solution of the following LME model

$$y = Td + \sum_{k=1}^{p} z_{sk}b_{sk} + X\beta + Zb + \in$$

$$= (T,X)(d^T,\beta^T)^T + (Z_1,...,Z_p,Z)(b_1^T,...,b_p^T,b^T)^T + \in, \tag{7}$$

where $\xi_{kj}(t) = R_{k,penalized}(t,t_j)$, $\Sigma_k = \{<\xi_{ki},\xi_{kj}>\}_{i,j=1}^{n}$, $\Sigma_k = Z_{sk}Z_{sk}^T$; $T = (T_1,...,T_p)$, $T_k = \{\phi_{kl}(t_j)\}_{l=1\,j=1}^{M_k\,n}$; $b_{sk} \sim N(0,\sigma^2\theta_k I_{M_k}/n\lambda)$, $b \sim N(0,\sigma^2 D)$, $\in \sim N(0,\sigma^2 G)$.

Bayesian credible bands are commonly used to evaluate smoothing spline fitted values [26] by assuming the following prior for f,

$$F(t) = \sum_{i=1}^{M}\zeta_i\phi_i(t) + \delta^{\frac{1}{2}}U(t), \tag{8}$$

where $\zeta_1,...,\zeta_M \sim^{i.i.d} N(0,\kappa)$, $U(t)$ is a zero-mean Gaussian process with covariance function $R_{penalized}(t,t_j)$, ζ_i and $U(t)$ are independent and κ and δ are positive constants. p–values cannot be directly calculated because the distribution (of a function) under the null hypothesis is unknown. Such credible bands must be interpreted in an across-the-function fashion, and they cannot be used for describing features of the curve. Since we are estimating a function instead of a real-value parameter, the credible region should represent realizations $f(\cdot)$ of a stochastic process for a fixed time τ. The Bayesian credible bands, however, is based on the average coverage probability (ACP).

$$ACP = \frac{1}{n}\sum_{k=1}^{n} Pr\{f(t_k) \in C(\alpha,t_k)\}, \tag{9}$$

where $f(\cdot)$ is fixed, t s are randomly selected and $C(\alpha, t_k)$ is the corresponding credible band at significance level α. ACP has been shown to be close to nominal coverage rate $1-\alpha$ in simulations [27] and theoretical justification [28]. Note that this average coverage is weaker than pointwise coverage.

When comparing two groups with estimated functions $\hat{f}_1(t)$ and $\hat{f}_2(t)$, the difference of these two groups can be derived

$$\hat{f}_{diff}(t) = \hat{f}_1(t) - \hat{f}_2(t)$$

with standard error

$$se_{diff}(t) = \sqrt{se_1^2(t) + se_2^2(t)},$$

where $se_1(t)$ and $se_2(t)$ are posterior standard errors of the fitted curves $\hat{f}_1(t)$ and $\hat{f}_2(t)$ respectively. A check for group difference can be performed by examining whether the credible band covers the horizontal zero line. Similar idea was once used in Bowman and Young [29] and Liang [30].

As can be seen from (6), the RKHS formulation makes it easy to model parametric and nonparametric components by manipulating the RKHS decomposition, and it combines various spline models under a unified framework. Smoothing spline ANOVA (SS ANOVA) is available with similar interpretation as ordinary ANOVA. Suppose the model space can be decomposed as in (2) and the estimated function is (4) with orthogonal basis $\phi_1(t), ..., \phi_M(t)$, and let   be the projection operator onto unpenalized subspace H_{null} spanned by $\phi_v(t)$ for $v = 1, ..., M$, and $P_{penalized}$ be the projection onto $H_{penalized}$, then the function can be written as

$$f = (P_{10} + \cdots + P_{M0} + P_{penalized})f \cdot \qquad (10)$$

ANOVA and linear regression can be seen as special cases of SS ANOVA by specifying their corresponding decompositions. For an SLM model with linear parametric component and cubic spline nonparametric component, assume the following model

$$y_{kwj} = f(k, w, t_j) + \in_{kwj}; k = 1, \cdots, K; w \in B_k; t_j \in [0,1],$$

where y_{kwj} is the observed tumor volume at time t_j of mouse w in group k. For group k, denote B_k as the population from which the mice in group k were drawn and P_k as the sampling distribution. $f(k, w, t_j)$ is the "true" tumor volume at time t_j of mouse w in the population B_k, and $\in_k$'s are random errors. $f(k, w, t_j)$ is a function defined on $\{\{1\} \otimes B_1, \{2\} \otimes B_2, ..., \{K\} \otimes B_K\} \otimes [0,1]$. Note that $f(k, w, t_j)$ is a random variable since w is a random sample from B_k. What we observe are realizations of this "true" mean function plus random errors. We use label w to denote mice we actually observe.

We define four averaging operators that project the function f onto modular structures constituting this SLM model:

$$A_2 f = \int_{B_k} f(k, w, t) dP_k(w),$$

$$A_1 f = \sum_{k=1}^{K} A_2 f(k, t) / K,$$

$$A_3 f = \int_0^1 f(k, w, t) dt,$$

$$A_4 f = \left\{ \int_0^1 \frac{df(k, w, t)}{dt} \right\} (t - 0.5).$$

Then we have the following SS ANOVA decomposition

$$\begin{aligned} f &= \{A_1 + (A_2 - A_1) + (I - A_2)\}\{A_3 + A_4 + (I - A_3 - A_4)\}f \\ &= A_1 A_3 f + A_1 A_4 f + A_1 (I - A_3 - A_4)f \\ &\quad + (A_2 - A_1)A_3 f + (A_2 - A_1)A_4 f + (A_2 - A_1)(I - A_3 - A_4)f \\ &\quad (I - A_2)A_3 f + (I - A_2)A_4 f + (I - A_2)(I - A_3 - A_4)f \\ &= \mu_0 + \delta_0(t - 0.5) + s_0(t) + \mu_k + \delta_k(t - 0.5) + s_k(t) \end{aligned}$$

$$+ \alpha_{w(k)} + \gamma_{w(k)}(t - 0.5) + s_{w(k)}(t), \qquad (11)$$

which can be interpreted in parallel with the classical mixed models as follows: μ_0 is a constant, $\delta_0(t-0.5)$ is the linear main effect of time, $s_0(t)$ is the smooth main effect of time, μ_k is the main effect of group, $\delta_k(t-0.5)$ is the linear interaction between time and group, $s_k(t)$ is the smooth interaction between time and group, $\alpha_{w(k)}$ is the main effect of mouse, $\gamma_{w(k)}(t-0.5)$ is the linear interaction between time and mouse, and $s_{w(k)}(t)$ is the smooth interaction between time and mouse. We can calculate the main effect of time as $\delta_0(t-0.5) + s_0(t)$, the interaction between time and group as $\delta_k(t-0.5) + s_k(t)$, and the interaction between time and mouse as $\gamma_{w(k)}(t-0.5) + s_{w(k)}(t)$. The first six terms are fixed effects. The last three terms are random effects since they depend on the random variable w. Depending on time only, the first three terms represent the mean curve for all mice. The middle three terms measure the departure of a particular group from the population mean curve. The last three terms measure the departure of a particular mouse from the mean curve of a population from which the mouse was chosen.

For categorical variables, we can either estimate each level by shrinkage estimates, which reduce overall means squared error by penalizing departure from overall mean; or we can fit each level separately as fixed categorical covariates. In this study, it is beneficial to not penalize group difference because we are interested in comparing group differences. Both shrinkage and fixed-effects estimates can be fitted under the RKHS framework with different RKHS decompositions. Suppose we want to model time using cubic spline, the categorical variable group as fixed effects and shrink mouse factors toward constants (modeled as random effects), the fixed effects in Equation (11) can be re-written as

$$\begin{aligned} f_k(t) &= \mu_0 + \delta_0(t - 0.5) + s_0(t) + \mu_k + \delta_k(t - 0.5) + s_k(t) \\ &= \{\mu_0 + \mu_k\} + \{\delta_0(t - 0.5) + \delta_k(t - 0.5)\} + \{s_0(t) + s_k(t)\} \end{aligned} \qquad (12)$$

Analysis of Xenograft Tumor Data

In this section, we use the proposed method to analyze data from the study described in introduction. Intuitively, the longer mice live, the more favorable the treatment combination is. But the mice with same survival times may have different tumor volumes. The differences in tumor volumes might represent quality of life and reveal potential intervention mechanisms of treatments. For instance, scatterplots from Treatment F and G look quite similar. As we investigate measurements from these two groups, mice with Treatment F tend to die earlier if the tumor volumes are not well controlled at early times, while mice in Group G are more likely to survive after having high initial tumor volumes.

Detail of the data set is given as follows. $W=140$ mice were assigned to $K=14$ treatments, and tumor volumes were measured on each mouse weekly for maximum of 12 weeks. Time has been normalized to $[0,1]$. There are two categorical covariates group and mouse, and a continuous covariate time. We treat group and time as fixed. From the design, the

mouse is nested within group and treated as random.

The SLM model discussed in model and methods enables us to (i) estimate the group (treatment) effects; (ii) estimate the population mean volume as functions of time; and (iii) predict response over time for each mouse. For the purpose of this study we are most interested in (i) and (ii).

Based on the SS ANOVA decomposition in (11) and (12), we may fit the following three models. Note that Term 7 in (11), the random effects in intercepts, is not included. From the scatterplots in Figure 1, all mice start out at about the same level (as designed by the study), indicating there is not much of mouse random effects in intercept. Unsurprisingly, the random intercept term causes convergence problems as the near-zero variance is being estimated. Likelihood ratio tests of nested LME objects suggest insignificance of the random intercept beyond the random slope. Therefore, we carry out the analysis without this term.

- **Model 1** includes the first six terms and the eighth term in (11). It fits different population mean curves for each group plus a random slope for each mouse. We assume that $\gamma_{w(k)} \sim^{iid} N(0, \sigma_{\gamma}^2)$. $\in_{kwj} \sim^{iid} N(0, \sigma^2)$, and they are mutually independent.

- **Model 2** includes all terms except 7th term in (11). It fits different population mean curves for each group plus a random slope and a random smooth effect for each mouse. We assume that $\gamma_{w(k)} \sim^{iid} N(0, \sigma_{\gamma}^2)$. $s_{w(k)}(t)$'s are stochastic processes which are independent between mice with mean zero and covariance function $\sigma_s^2 R_{penalized}(s,t)$, with $R_{penalized}(s,t)$ being reproducing kernel of the penalized space for cubic splines (specific form can be constructed from scaled recursive Bernoulli polynomials). We further assume within-group error $\in_{kwj} \sim^{iid} N(0, \sigma^2)$ and independence between the random effects and random errors. These are similar to usual assumptions in LME models.

- **Model 3** fits first-order autoregressive correlations structure AR(1) to Model 2, i.e. $\in_{kwj}$'s are no longer assumed to be independent.

- **Model 4** uses an extra parameter beyond Model 3 to account for unequal variances of within-group error terms $\in_{kwj}$'s by modeling variance as an exponential function of time.

Among all 14 treatments shown in Figure 1, some can be instantly eliminated, such as A and K, because of poor survival times. The analysis is then narrowed down to the 6 groups N, J, F, L, D, G with similar survival times, but may have different tumor volume growth profiles (Figure 2).

Random effects in the smoothing components $s_{w(k)}(t)$ can be fitted manually by constructing a block-diagonal symmetric covariance matrix of dimensions $n \times n$ for all observations. Each block corresponds to the reproducing kernel $R_{penalized}(s,t)$ evaluated at observed design points for each subject. Let $w = 1, ..., 140$, $k = 1, ..., 6$, observed time points for wth mouse $t_w = (t_1, ..., t_{nw})^T$ random smoothing vector for wth mouse $u_w = (s_{w(k)}(t_1), ..., s_{w(k)}(t_{n_w}))^T$, where $Q_w = \{R_{penalized}(s,t)\}_{n_w \times n_w}, s,t = 1, ..., t_{n_w}$; total random smoothing vector $u = (u_1^T, ..., u_{140}^T)^T$. Then random effects in smoothing components can be modeled by specifying $u \sim N(0, \sigma_u^2 Q)$, where $Q = \text{diag}(Q_1, ..., Q_{140})$.

Such SLM models are solved by finding solutions of their LME counterparts, as shown in (7). This connection can also be utilized to
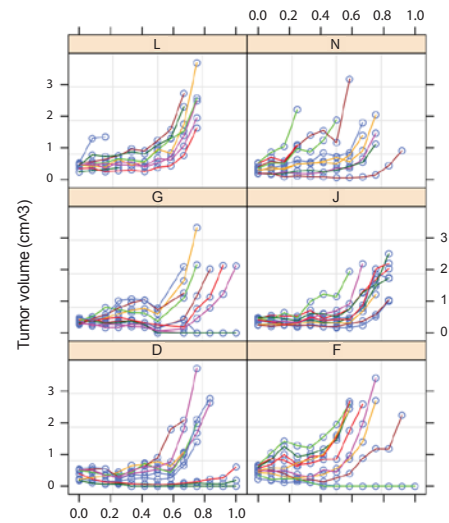

Figure 2: Scatterplots of groups of interest.

Model	AIC	BIC	logLik	Test	L.Ratio	p-value
1	668.0160	754.6460	-314.0080			
2	536.0836	627.0451	-247.0418	1 vs 2	133.9324	<0.0001
3	430.4238	525.7168	-193.2119	2 vs 3	107.6598	<0.0001
4	53.2419	152.8665	-3.62095	3 vs 4	379.1819	<0.0001

Table 1: Model selection criteria for corresponding LME models.
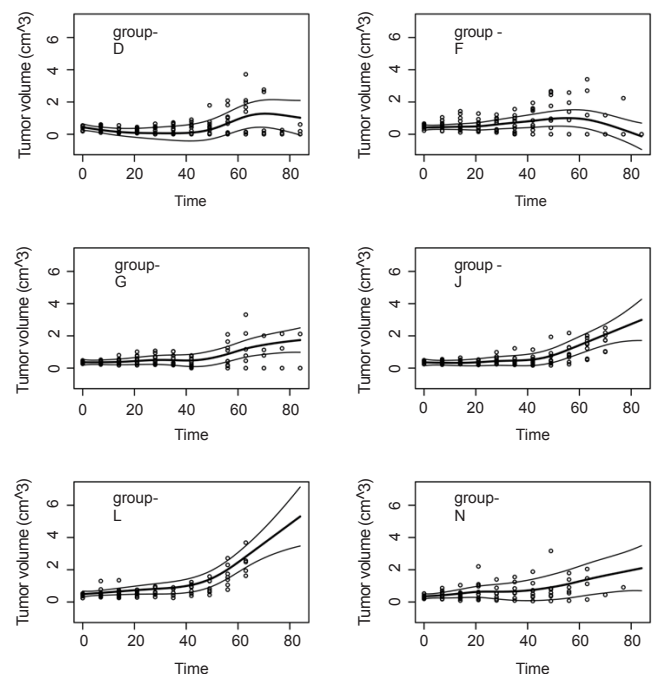

Figure 3: Predicted group means and 95% Bayesian credible bands.

calculate AIC, BIC and LRT in the sense of conventional parametric models for model selection and comparison, as shown in Table 1. LRTs of these four nested models suggest Model 4 to be most favorable. Estimated serial correlation coefficient for AR(1) in Model 3 is large
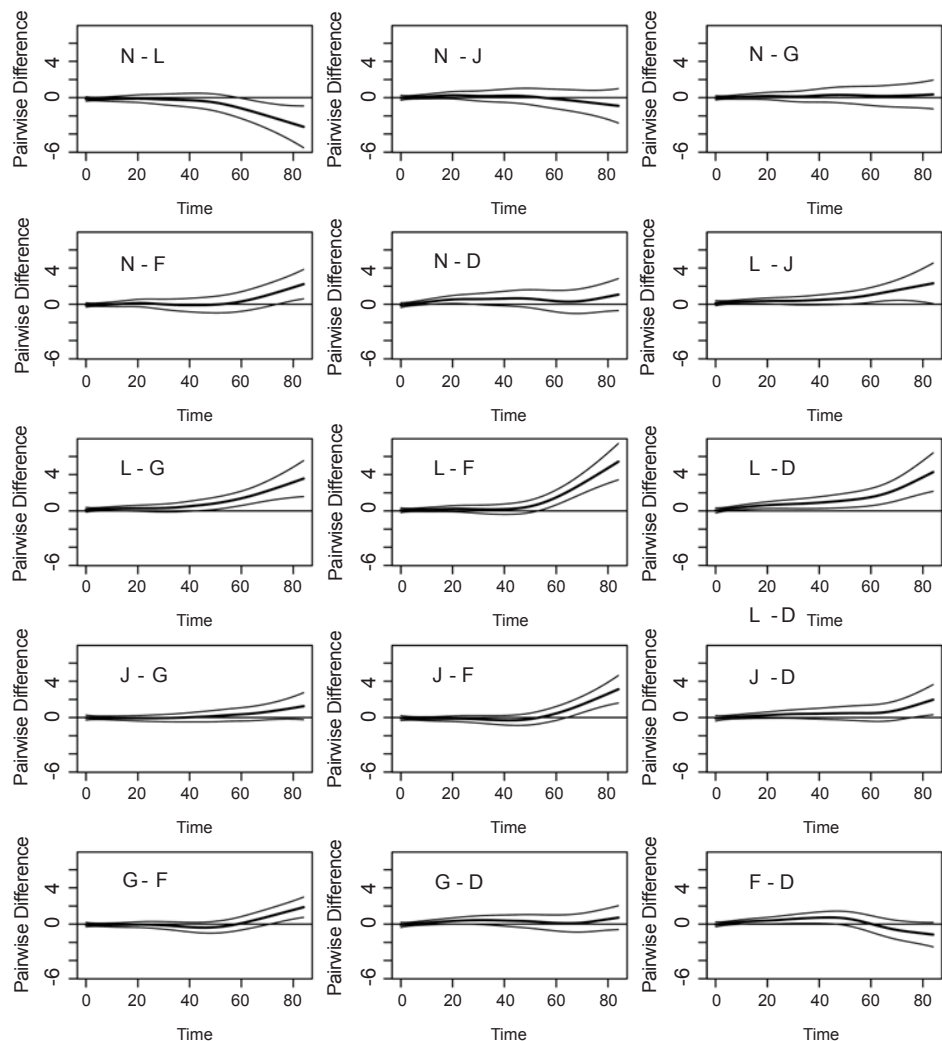
Figure 4: Pairwise difference estimates with Bayesian credible bands ("-" represents an algebraic minus sign in the top left of each plot).

Pair	ANOVA	SLM
N-G	no	no
N-J	no	no
N-L	yes (N)	yes (N)
L-J	yes (J)	yes (J)
N-D	yes (D)	yes (D)
N-F	no	yes (F)
L-D	yes (D)	yes (D)
L-F	yes (F)	yes (F)
L-G	yes (G)	yes (G)
J-D	yes (D)	yes (D)
J-F	no	yes (F)
J-G	no	no
F-D	yes (D)	yes (D)
G-D	no	yes (D)
G-F	yes (G)	yes (F)

Table 2: Comparison of the results based on ANOVA and the proposed method–whether difference is detected (the one with lower tumor volumes if detected).

(0.72), indicating strong within-subject correlation and supporting adequacy of fitting within-subject correlation. Increasing variance over

time is quite obvious from scatterplots, which also coincides with the LRT pick of Model 4.

Figure 3 shows predicted curves along with 95% Bayesian credible bands calculated from posterior distributions of fitted values with a diffuse prior by letting $\kappa \to \infty$ in (8). Note that these predicted curves and credible bands are supposed to represent the mean curves for sub-population B_k (Group k), not individuals. This explains why some estimated tumor growth patterns actually start to go down at the end of the study, such as D and F, reflecting the fact that only mice with lower tumor volumes survived among the population of these groups. Figure 4 shows pairwise differences with Bayesian credible bands. If the credible band runs above or below zero, the two treatments are deemed different. Survival times are not under consideration here, so treatment with lower tumor volumes is better (given a mouse survives till that time). For example, treatments N and J are not different from each other since the zero line is fully contained in the credible band. D is better than L because the credible band is mostly above zero. The result is summarized in Table 2.

As comparison, result from a simple ANOVA analysis for the log-transformed data is also listed in Table 2. SLM model is not only able to

detect more significantly different pairs, but also provides more insight into how the pairs are different from each other. It is worth pointing out that for Treatment G versus Treatment F, although ANOVA and SLM both detect the difference, ANOVA concludes G has lower tumor volumes, while SLM picks F. Simply taking averages in ANOVA ignores information from time and correlation of measurements within each mouse that reveal different behaviors of the two treatment groups towards the end. By looking at the SLM predicted population patterns in Figure 3, while the two tumor growth profiles look similar during early period, tumor volumes in Group F are lower than those in Group G for mice that survived to the final period of the study. For example, if we only look at tumor volumes past normalized time 0.8, the means for Group F and G are 0.4270 and 0.6264 respectively.

Remark

The dimension of the random spline covariance matrix Q increases with the number of observations. As a result, direct computation might take a long time. For this specific dataset with all groups, it took about one day to compute the model with random smoothing effects on a personal computer. To speed up the program, a low-rank approximation algorithm was used to reduce the dimension of Q by eliminating eigenvectors corresponding to small eigenvalues [31]. In our example, with a cutoff value 0.001, the approximation reduced the computation time to about an hour and gave almost identical results.

Conclusion

SLM model only specifies part of the mixed-effects model parametrically and leaves the rest to be modeled non-parametrically by data itself. It incorporates interpretability and parsimony of a parametric model and flexibility of a nonparametric model, and also avoids estimating parameters with little biological relevance. The close connection with LME models enables us to utilize existing LME fitting procedures for computation and model selection. Even when the final goal is to build a fully parametric model, SLM model can be useful for initial data exploration and shed light on parametric models to follow.

The cost of SLM model is relatively heavy computation (but reasonable with optimization) and difficulty in deriving closed-form inferences in the functional space. More research on model selection methods is needed as well.

Further effort can be invested for some special post-hoc methods such as multiple comparison correction for the SLM model. The major difficulty for such extensions would be establishing an expression of the correction, giving a theoretical justification and calculating p-value.

References

1. Wu J (2011) Confidence intervals for the difference of median failure times applied to censored tumor growth delay data. Stat Biopharm Res 3: 488-496.

2. Houghton PJ, Morton CL, Tucker C, Payne D, Favours E, et al. (2007) The pediatric preclinical testing program: description of models and early testing results. Pediatr Blood Cancer 49: 928-940.

3. Corbett TH, White K, Polin L, Kushner J, Paluch J, et al. (2003) Discoveryand preclinical antitumor efficacy evaluations of ly32262 and ly33169. Invest New Drugs 21: 33-45.

4. Wu J (2010) Statistical inference for tumor growth inhibition T/C ratio. J Biopharm Stat 20: 954-964.

5. Hothorn L (2006) Statistical analysis of in vivo anticancer experiments: Tumor growth inhibition. Drug Information Journal 40: 229-238.

6. Kong M, Yan J (2011) Modeling and testing treated tumor growth using cubic smoothing splines. Biom J 53: 595-613.

7. Tan M, Fang HB, Tian GL, Houghton PJ (2002) Small-sample inference for incomplete longitudinal data with truncation and censoring in tumor xenograft models. Biometrics 58: 612-620.

8. Heitjan DF, Manni A, Santen RJ (1993) Statistical analysis of in vivo tumor growth experiments. Cancer Res 53: 6042-6050.

9. Liang H, Sha N (2004) Modeling antitumor activity by using a non-linear mixed-effects model. Math Biosci 189: 61-73.

10. Vonesh EF, Chinchilli VM (1997) Linear and Nonlinear Models for the Analysis of Repeated Measurements. CRC Press, New York, USA.

11. Davidian M, Giltinan DM (1995) Nonlinear Models for Repeated Measurement Data. Chapman and Hall.

12. Liang H (2005) Modeling antitumor activity in xenograft tumor treatment. Biom J 47: 358-368.

13. Roy Choudhury K, Kasman I, Plowman GD (2010) Analysis of multi-arm tumor growth trials in xenograft animals using phase change adaptive piecewise quadratic models. Stat Med 29: 2399-2409.

14. Laajala TD, Corander J, Saarinen NM, Mäkelä K, Savolainen S, et al. (2012) Improved statistical modeling of tumor growth and treatment effect in pre-clinical animal studies with highly heterogeneous responses in vivo. Clin Cancer Res 18: 4385-4396.

15. Zhao L, Morgan MA, Parsels LA, Maybaum J, Lawrence TS, et al. (2011) Bayesian hierarchical changepoint methods in modeling the tumor growth profiles in xenograft experiments. Clin Cancer Res 17: 1057-1064.

16. Heitjan DF (2011) Biology, models, and the analysis of tumor xenograft experiments. Clin Cancer Res 17: 949-951.

17. Houghton PJ, Morton CL, Gorlick R, Kolb EA, Keir ST, et al. (2010) Initial testing of a monoclonal antibody (imc-a12) against igf-1r by the pediatric preclinical testing program. Pediatr Blood Cancer 54: 921-926.

18. Schoenberg IJ (1964) Spline functions and the problem of graduation. Proc Natl Acad Sci U S A 52: 947-950.

19. Kimeldorf G, Wahba G (1971) Some results on tchebycheffian spline functions. J Math Anal Appl 33: 82-95.

20. Laird NM, Ware JH (1982) Random-effects models for longitudinal data. Biometrics 38: 963-974.

21. Wang Y (2011) Smoothing Splines: Methods and Applications. CRC Press, Boca Raton, USA.

22. Aronszajn N (1950) Theory of reproducing kernels. Trans Am Math Soc 68: 337-404.

23. Wahba G (1990) Spline models for observational data. Society for Industrial Mathematics.

24. Wang Y (1998) Mixed effects smoothing spline analysis of variance. J R Stat Soc Series B Stat Methodol 60: 159-174.

25. Opsomer J, Wang Y, Yang Y (2001) Nonparametric regressin with correlated errors. Statistical Science 16: 134-153.

26. Kimeldorf GS, Wahba G (1970) A correspondence between bayesian estimation on stochastic processes and smoothing by splines. The Annals of Mathematical Statistics 41: 495-502.

27. Wahba G (1983) Bayesian "confidence intervals" for the cross-validated smoothing spline. Journal of the Royal Statistical Society. Series B (Methodological) 45: 133-150.

28. Nychka D (1988) Bayesian confidence intervals for smoothing splines. J Am Stat Assoc 83: 1134-1143.

29. Bowman A, Young S (1996) Graphical comparison of nonparametric curves. J R Stat Soc Ser C Appl Stat 45: 83-98.

30. Liang H (2006) A new method of evaluating antitumor activity from measured tumor volumes. Contemp Clin Trials 27: 269-273.

31. Wood S (2003) Thin plate regression splines. J R Stat Soc Series B Stat Methodol 65: 95-114.