Bayesian Hierarchical Changepoint Methods in Modeling the Tumor Growth Profiles in Xenograft Experiments

Lili Zhao1, Meredith A. Morgan2, Leslie A. Parsels3, Jonathan Maybaum3, Theodore S. Lawrence2, and Daniel Normolle4

Abstract

**Purpose:** The standard approach of using tumor doubling time to assess growth delay may not accurately represent tumor response, especially if the growth rates are not constant. Therefore, we developed a method to compare the antitumor activities of different treatments in xenograft experiments that uses the entire growth curve to estimate nonconstant growth rates.

**Experimental Design:** A Bayesian hierarchical changepoint (BHC) method was used to model logarithmically transformed tumor volumes (TV). Each tumor was assumed to have a growth profile, represented by a prenadir regression rate, a regression period, a nadir volume, and a postnadir regrowth rate. Confidence intervals were calculated to compare these features between different treatments. We used data from a study assessing the effects of radiation, gemcitabine, and a Chk1/2 inhibitor on MiaPaCa-2 xenografts.

**Results:** We found that the BHC model provided a good fit to the data and more descriptive features than the tumor doubling approach. This model detected significant tumor regression in the AZD7762 + 1 Gy and GEM + 1 Gy that was not detected when comparing the tumor doubling times. The BHC model also provided evidence that the growth inhibition resulted from a direct tumor effect rather than an indirect effect on the tumor bed, as evidenced by dramatic tumor regression in response to effective treatments and similar postnadir regrowth rates across all treatment groups.

**Conclusions:** Compared with the tumor doubling time approach, the BHC model utilizes all data, providing more descriptive features that address mechanisms underlying tumor growth inhibition and maximize the biological information obtained from tumor xenografts studies. Clin Cancer Res; 17(5); 1057–64. ©2010 AACR.

Introduction

Tumor xenograft models play an important role in translational cancer research. In these models, immunocompromised mice are grafted with cancer cells, treated with anticancer therapies, and then monitored for the effects of therapy on tumor growth during treatment as well as the more sustained effects on tumor regrowth after treatment.

Tumor regression and regrowth is complicated and involves several biological processes. Depending on the treatment, tumor growth patterns can be quite different. For example, although untreated tumors may grow during an entire study period, radiation-treated tumors usually regress and then subsequently regrow. The time until TV doubling, defined as the earliest day on which the TV is at least twice as large as on the first day of treatment, is the most commonly used end point in these studies.

There are, however, two major disadvantages with using doubling time. First, by ignoring the measurements taken after time-to-tumor doubling, biologically important aspects of a treatment effect might be missed. Second, the single estimate of doubling time does not address the biological mechanisms underlying different patterns of tumor growth. For instance, in response to an effective therapy, tumors may regress to a level below the limit of quantitation (10 mm³ in this study) for some time, possibly until the end of the observation period. Volumes below the limit of quantitation are not considered missing, but left censored; the exact volume cannot be estimated beyond stating that the tumor is less than 10 mm³. Thus, an approach that can assess the regression period and volume...

Authors’ Affiliations: 1Biostatistics Unit, University of Michigan Comprehensive Cancer Center; Departments of 2Radiation Oncology and 3Pharmacology, University of Michigan Medical School, Ann Arbor, Michigan; and 4Biostatistics Facility, University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania

Corresponding Author: Lili Zhao, Biostatistics Unit, University of Michigan Comprehensive Cancer Center, 1500 East Medical Center Drive, Ann Arbor, MI 48109. Phone: 7347636898; Fax: 7349364540. E-mail: zhaolili@umich.edu
doi: 10.1158/1078-0432.CCR-10-1935
©2010 American Association for Cancer Research.
Zhao et al. colleagues (5) developed a and colleagues (3), Fang and colleagues (4), and Tian and colleagues (5) developed a $t$ test via the EM (expectation maximization) algorithm and, also, Bayesian approaches for testing differences between 2 treatment regimens by analyzing longitudinal data and taking into account the censored data. Although the models proposed by Tan, Fang, and Tian are an improvement over the previously mentioned approaches (Wilcoxon, repeated-measures ANOVA), their test hypotheses are based on comparing either specific time points or a function of selected time points, which cannot be used to study tumor growth patterns. Liang (6) developed a nonparametric approach to estimate the tumor growth profiles by penalized splines, but important characteristics of the tumor growth curve, such as tumor regression and growth rates, cannot be estimated using this approach. Other studies have utilized the time in days for the tumors to reach a predefined TV as an end point. The choice of the end point tumor size at which to assess delay, however, is critical in determining the magnitude of delay (7). Although these end points are useful to estimate the tumor growth delay, they neither capture all of the data nor address the different mechanisms underlying tumor growth. Thus, analyses that take into account the entire data set, as well as the characteristics of the tumor growth curve, are needed to maximize the biological information gained from tumor xenografts studies.

In this article, we propose a novel Bayesian hierarchical changepoint (BHC) approach to model the logarithmically transformed TVs. The 2-piece linear (‘broken stick’) line describes tumor regression, followed by tumor regrowth. The hinge is when the tumor starts to regrow and the volume at the hinge is the nadir. Bayesian changepoint methods have been successfully applied to CD4 counts to predict the timing of an HIV viral rebound (8), to longitudinal PSA (prostate-specific antigen) series to predict cancer onset (9, 10), and to cognitive function over time to predict the decline in memory that precedes diagnosis of dementia (11). In this study, we applied the BHC model to the tumor xenograft volume profile data, and assessed "during-treatment" and "after-treatment" effects, by testing if the features of the tumor growth profiles (e.g., regression rate, regression period, volume nadir, and regrowth rate) differ between treatment groups.

Methods

Transformation of volume measurements

In xenograft experiments, tumor growth rates are typically close to exponential both before and after the nadir. For this reason, TVs were logarithmically transformed before analysis to obtain linear growth profiles before and after the nadir. The base 2 logarithms have scientifically useful interpretations: the postnadir $\log_2$ volume growth rate is equivalent to the number of times the tumor doubles per day and its reciprocal corresponds to the tumor doubling time. Similarly, the $\log_2$ volume regression rate is the number of times the tumor halving per day and its reciprocal corresponds to the tumor halving time.

BHC model

We assume that the $\log_2$ volumes ($y_{ijk}$, the $i$th tumor in the $j$th animal at time $t_j$) are normally distributed, with an expected value $\mu_{ijk}$ and variance $\sigma_i^2$, as described by the tumor-level piecewise linear model:

$$
\mu_{ijk} = \begin{cases} 
\log_2 a_{ik} - \epsilon^{b_{ik}}(r_{ik} - t_j), & t_j \leq r_{ik} \\
\log_2 a_{ik} + \epsilon^{b_{ik}}(t_j - r_{ik}), & t_j \geq r_{ik} 
\end{cases}
$$

The model assesses each tumor growth profile, represented by a prenadir regression rate $[\exp(b_{i1})]$, a regression period $[r_{ik}]$, a volume nadir $[a_{ik}]$, and a postnadir regrowth rate $[\exp(b_{i2})]$. For left-censored observations, the condition that a given $y_{ijk}$ is less than the limit of quantitation [in the example, $\log_2(10) = 3.3$] is incorporated in the estimation process in the same fashion as censored data are considered in parametric survival models.

The random effects $b_{i1}$ and $b_{i2}$ are exponentiated to model both regression-regrowth profile and growth-only profile. A growth profile does not exhibit regression and uses only the regrowth line of the model, with $r_{ik} \leq 0$. In such tumors, the regrowth starts immediately after the initiation of the therapy. To complete the BHC model, it

Translational Relevance

The new statistical model that we propose in this study provides a comprehensive analysis of the tumor growth profiles observed in tumor xenograft experiments including tumor regression rate, regression period, nadir volume, and tumor regrowth rate. Each of these features provided by this model could have clinical implications for treatment. For example, higher regression rate, longer regression time, and lower regrowth rate could predict longer intervals to, and/or less frequent, tumor recurrence. This statistical approach maximizes the biological information obtained from tumor xenografts, which should facilitate the translation of preclinical xenograft studies to more rational clinical trial designs.
is then assumed that the tumor-level random effects $a_{ik}$, $b_{ik}$, and $r_{ik}$ are drawn from probability distributions with treatment-specific parameters; these parameters are estimated with 90% highest probability density (HPD) intervals. Given any treatment group, if the lower bound of the 90% HPD interval on the regression period is smaller than zero, then there is a high probability that there is no tumor regression, and only the postnadir regrowth rate will be presented. If the HPD interval of the difference in any feature between 2 treatments covers 0, the 2 treatments are not significantly different on that feature.

The treatment-specific parameters are assumed to have been drawn from probability distributions with vague priors. The full BHC model setup is described mathematically in the Appendix. Estimation was implemented in WinBUGS (12), a statistical software package that uses Markov Chain Monte Carlo to generate samples from the relevant posterior distributions. The mechanics of estimation in the presence of left-censored data are described in the WinBUGS code in the Appendix as well.

Estimates for treatment groups

The model provided a good fit to the data, demonstrated by a reasonable fit of the estimated growth profiles on the observed profiles (see Fig. 1) as well as justified by the diagnostic test embedded in the statistical package.

The estimates of the BHC model parameters are shown in Table 1 and estimated growth profiles for all treatments are presented in Figure 2. Control, AZD7762, GEM, 1 Gy, and GEM + AZD7762 did not induce tumor regression, and the treatment groups were not statistically different [90% HPD intervals of regrowth rate, expressed as the number of times a tumor doubles per day, were (0.03–0.09), (0.03–0.07), (0.03–0.07), (0.03–0.08), and (0.02–0.07)], respectively; the 90% HPD intervals of all pairwise differences covered zero. With a radiation dose of 2 Gy (such as 2 Gy, AZD7762 + 2 Gy, GEM + 2 Gy, and GEM + AZD7762 + 2 Gy), tumor regression rates were significantly higher, tumor regression periods were significantly longer, and the nadir volumes were significantly lower than the corresponding treatments with 1 Gy. For example, the 90% HPD interval of nadir volume in AZD7762 + 1 Gy is 152 to 221 mm$^3$, which is substantially above the interval of 3 to 13 mm$^3$ in AZD7762 + 2 Gy. The addition of AZD7762 to radiation and gemcitabine resulted in enhanced tumor regression; 90% HPD intervals of the difference in nadir volume were (96 mm$^3$, 173 mm$^3$) for GEM + 1 Gy versus GEM + AZD7762 + 1 Gy and (6 mm$^3$, 21 mm$^3$) for GEM + 2 Gy versus GEM + AZD7762 + 2 Gy. In summary, this model suggests that the growth inhibition results from direct effects on the tumor cells and not from indirect effects on the tumor bed, evidenced by dramatic tumor regression in response to effective treatments and similar postnadir regrowth rates across all treatments.

Comparison with time-to-tumor volume doubling test

This experiment was also analyzed using the time-to-tumor volume doubling end point by log-rank tests. The 2 approaches (BHC vs. doubling) did not yield entirely consistent results. The BHC model found 2 significant comparisons that were not found by the time-to-doubling approach; AZD7762 + 1 Gy versus 1 Gy (log-rank: $P = 0.45$) and GEM + 1 Gy versus 1 Gy (log-rank: $P = 0.73$). Although these treatments are associated with similar doubling times, both AZD7762 + 1 Gy and GEM + 1 Gy demonstrate significant tumor regression by the BHC model, which is confirmed by inspection of the growth curves. These findings support the use of the BHC model, which has more descriptive features than the time-to-tumor volume doubling test.

The time-to-tumor volume doubling test found 3 significant comparisons: control versus AZD7762 ($P = 0.04$), control versus GEM + AZD7762 ($P < 0.0001$), and GEM versus GEM + AZD7762 ($P = 0.0002$), which were not estimated to be significant by the BHC model in terms of tumor regrowth rate. However, consistent with the BHC model estimates, the time-to-tripling test showed nonsignificant results for all 3 of these comparisons ($P = 0.64$ for
AZD7762 vs. control, $P = 0.46$ for GEM + AZD7762 vs. control and $P = 0.65$ for GEM + AZD7762 vs. GEM. If tumor growth were consistently different between the treatment conditions, the time-to-doubling or -tripling test should yield the same results. Given the agreement between the BHC and time-to-tripling models, it seems likely that the differences found in the time-to-doubling test result from the limitations of analyzing a single time point rather than a fundamental difference between the growth curves.

**Discussion**

In this study, we have found that, in comparison with the most commonly used tumor growth model of time to doubling, the BHC approach helped to clarify the underlying biology by providing additional features of the growth profiles, including tumor regression rate, regression period, nadir volume, and regrowth rate.

During anticancer therapy, tumor cells are killed and eventually cleared in the circulation. As dead cells are cleared from the tumor, the blood and nutrient supply to the tumor improves, allowing repopulation of the tumor. When the rate of tumor repopulation overtakes the rate of tumor cell loss, the tumor starts to regrow, and at the same time, its volume achieves the nadir. For some effectively treated tumors, a high rate of regression and/or a longer period of regression could ultimately result in a complete tumor regression in the study period. If tumor repopulation starts very early and the rate of repopulation is faster than the rate of cell loss, the tumors typically grow during the entire study period. The BHC model efficiently assessed whether the treatment induced a significant tumor regression, and if so then the rate and period of the regression. This model also distinguished the initial tumor growth inhibition (i.e., cell killing) from the subsequent inhibition of tumor regrowth (i.e., tumor bed effect; ref. 14), which will be interpreted as tumor growth delay by the time-to-doubling approach. All of these features could have clinical implications for treatment. For example, higher regression rate, longer regression time, and lower regrowth rate could predict longer intervals to, and/or less frequent, tumor recurrence.
In summary, a BHC model was developed to describe clinically meaningful characterization of nonlinear tumor growth profiles observed in many xenograft experiments, using off-the-shelf programming components. It not only provides new insights into the current xenograft study, but also lays ground work for future studies in cancer stem cell biology.
Prior specification

\[
\begin{align*}
(b_1, b_2) | \Omega_b & \sim MVN \left\{ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \Omega_b \right\} \\
(b_1, b_2) & \sim MVN \left\{ \begin{pmatrix} -3 \\ -3 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0.1 \end{pmatrix} \right\} \\
\Omega_b & \sim \text{Wishart} \left\{ \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, 2 \right\}
\end{align*}
\]

\[
\log_2 \, a_{ik} | \tau_i \sim \text{N}(a, \tau_i) \\
a \sim \text{N}(0, 0.01) \\
\tau_i \sim \text{Gamma}(0.1, 0.1) \\
r_{ik} | r, \tau_i \sim \text{N}(r, \tau_i) \\
r \sim \text{N}(20, 0.001) \\
\tau_r \sim \text{Gamma}(0.1, 0.1) \\
\tau \sim \text{Gamma}(0.01, 0.01)
\]

Note that WinBUGS describes the normal distribution in terms of a mean and a precision, where precision is the inverse of the variance ($\tau = 1/\sigma^2$), so the second term in the specification of a normal distribution will refer to precision in this Appendix.

We use hierarchical priors for the parameters $a_{ik}$, $b_1$, $b_2$, and $r_{ik}$, so the individual parameters are viewed as a random sample from a common distribution. For example, the hierarchical model assumes that the regression time, $r_{ik}$, is randomly selected from a normal distribution with treatment-specific mean $r$ and variance $1/\tau$. This hierarchical structure allows borrowing information across tumors, resulting in correlation between tumors in the same treatment group. In other words, the data from each tumor provide information about all other tumors in the same group. Therefore, the BHC model will be less sensitive to the outliers (tumor volumes or tumor profiles that are far from the others). In addition, an example code of borrowing information across similar treatment groups is shown in the WinBUGS code saved as TreatALL.txt.

We performed a sensitivity analysis to investigate the effect of the prior distributions on the estimates of the BHC parameters. We observed the effect of increasing the variance of the prior distribution on the estimates. For example, we increased the variance of the prior distribution of $b_1$ and $b_2$ from 10 to 100 and found...
that increasing the variances in the prior distributions had a negligible effect on these estimates, concluding that the priors we used provide little information in the final estimates.

**WinBUGS Code**

We ran 2 independent parallel chains with different starting values. The chains were run with a burn-in of 5,000 iterations and an additional 50,000 iterations for inference. With a thinning interval of 50, posterior distributions for the parameter estimates were therefore based on 2,000 iterations. Point estimates of parameters were calculated using the mean of the posterior distribution; 90% HPD intervals were obtained as well. The convergence of the Markov chain was assessed visually, and by the Gelman–Rubin diagnostic criterion (16). In addition, we visually checked the predicted growth profiles for each tumor, and they all fit very well to the observed profiles (see Fig. 1).

```winbugs
# WinBUGS code for a given treatment group (saved as Treat1.txt file)
# N: number of observations
# T: total number of time points
# NP: total number of mice
# y is the log2 tumor volume, and it is missing if data is censored
# ID is the mouse id and Loc is the flank index
# k is the index of flank
# ma, mu.r, mb1, and mb2 are treatment-specific volume nadir, regression period, regression rate, and regrowth rate, respectively.
# An indicator function (I) is added to the model when there are tumor volumes below the limit of quantitation, i.e. t.cen is 3.3 (log210) if data is censored, 0 otherwise.

model {
  # likelihood
  for (l in 1:N) {
    y[l] ~ dnorm(mu[l], tau)I(,t.cen[l])
    mu[l] <- a[ID[l],Loc[l]] - exp(b[1][ID[l],1])*(T[l]-(r[1][ID[l],Loc[l]])) * step((r[1][ID[l],Loc[l]])-T[l]) + exp(b[1][ID[l],2])*(T[l]-(r[1][ID[l],Loc[l]])) * step(T[l]-(r[1][ID[l],Loc[l]]))
  }
  # priors
  for (i in 1:NP) {
    b[i,1:2] ~ dnorm(mu.b[i,1:2], OmegaB[i,1:2,1:2])
    for (k in 1:2) {
      a[i,k] ~ dnorm(mu.a[i],tau.a)
      r[i,k] ~ dnorm(mu.r[i],tau.r)
    }
    tau ~ dnorm(0.001)
    tau.a ~ dgamma(0.001,0.001)
    tau.r ~ dnorm(20,0.001)
    OmegaB[i,1:2,1:2] ~ dwish(OmegaTau[,],2)
    mu.b[i,1:2] ~ dnorm(mub[i,1:2],tau[i,1:2])
    mu.r[i] ~ dnorm(20,0.001)
  }
  ma ~ pow(2, mu.a)
  mb1 ~ exp(mu.b[1])
  mb2 ~ exp(mu.b[2])
}

# WinBUGS for all treatments (saved as TreatALL.txt file)
# S is the number of treatment groups
# NP: total number of mice in all treatment arms
# v and m[i] is the index of treatment
# this code can be easily modified to include a hierarchical structure for similar treatments, for example, replacing mu.r[v]~dnorm(20,0.001) by mu.r[v]~dnorm(Omu.r, 0.001) will allow borrowing information across similar treatments on estimating the changepoint

model {
  # likelihood
  for (l in 1:N) {
    y[l] ~ dnorm(mu[l], tau)I(,t.cen[l])
    mu[l] <- a[ID[l],Loc[l]] - exp(b[1][ID[l],1])*(T[l]-(r[1][ID[l],Loc[l]])) * step((r[1][ID[l],Loc[l]])-T[l]) + exp(b[1][ID[l],2])*(T[l]-(r[1][ID[l],Loc[l]])) * step(T[l]-(r[1][ID[l],Loc[l]]))
  }
  # priors
  for (i in 1:NP) {
    b[i,1:2] ~ dnorm(mu.b[i,1:2], OmegaB[i,1:2,1:2])
    for (k in 1:2) {
      a[i,k] ~ dnorm(mu.a[i],tau.a[i])
      r[i,k] ~ dnorm(mu.r[i],tau.r[i])
    }
    for (v in 1:S) {
      mu.a[v] ~ dnorm(0.001)
      tau.a[v] ~ dgamma(0.001,0.001)
      mu.b[v,1:2] ~ dnorm(mub[v,1:2],OmegaB[v,1:2,1:2])
      OmegaB[v,1:2,1:2] ~ dwish(OmegaTau[,],2)
      tau.r[v] ~ dnorm(20,0.001)
      mu.r[v] ~ dnorm(20,0.001)
    }
    ma ~ pow(2, mu.a[v])
    mb1[v] ~ exp(mu.b[v,1])
    mb2[v] ~ exp(mu.b[v,2])
  }
  # Model convergence diagnostics
  Gelman Rubin diagnostic criterion (16).
  # Predicted growth profiles for each tumor
  # A visual check of the predicted growth profiles for each tumor, and they all fit very well to the observed profiles (see Fig. 1).
}
```

```
1. Shusterman S, Grupp SA, Barr R, Carpentieri D, Zhao H, Maris JM. References

2. Sauer LA, Dauchy RT, Blask DE, Krause JA, Davidson LK, Dauchy EM.


Bayesian Hierarchical Changepoint Methods in Modeling the Tumor Growth Profiles in Xenograft Experiments

Lili Zhao, Meredith A. Morgan, Leslie A. Parsels, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-10-1935

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2011/02/28/1078-0432.CCR-10-1935.DC1

Cited articles
This article cites 15 articles, 4 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/17/5/1057.full.html#ref-list-1

Citing articles
This article has been cited by 8 HighWire-hosted articles. Access the articles at:
/content/17/5/1057.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.