Bayesian Modeling and Prediction of Patient Accrual in Multi-Regional Clinical Trials

Xiaoxi Zhang[1] and Qi Long[2*]

[1] *Pfizer Inc., New York, NY 10017.* [2] *Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322.*

[*] *Corresponding author: qlong@emory.edu*

SUMMARY

With advances in medical research, effective treatments are becoming standard of care for many diseases. Consequently, when testing a new treatment, the sample size in a clinical trial is on the rise in order to demonstrate a moderate, yet clinically meaningful, improvement in therapeutic effect (compared to an active control), which often makes it impossible to enroll all patients from one region. In addition, regulatory requirements as well as other considerations may require trials to be conducted in multiple regions across the world. In multi-regional trials, the underlying overall and region-specific accrual rates often do not hold constant over time and different regions could have different start-up times, which combined with initial jump in accrual within each region often leads to a discontinuous overall accrual rate, and these issues associated with multi-regional trials have not been adequately investigated. In this paper, we clarify the implication of the multi-regional nature on modeling and prediction of accrual in clinical trials and investigate a Bayesian approach for accrual modeling and prediction, which models region-specific accrual using a nonhomogeneous Poisson process (NHPP) and allows the underlying Poisson rate in each region to vary over time. The proposed approach can accommodate staggered start-up times and different start-up accrual rates across regions. Our numerical studies show that the proposed method improves precision of accrual prediction (i.e., tighter posterior predictive credible intervals) compared to an existing NHPP model that ignores region-specific data.

KEY WORDS: Bayesian modeling, Clinical Trial, multi-regional Trial, Nonhomogeneous Poisson Process, Patient Accrual

1. Introduction

With advances in medical research, effective treatments are becoming standard of care for many diseases. Consequently, when testing a new treatment, the sample size in a clinical trial is on the rise in order to demonstrate a moderate, yet clinically meaningful, improvement in therapeutic effect (compared to an active control), which often makes it impossible to enroll all patients from one region. In addition, differences in regulatory requirements among different countries and concern on differential treatment effects across different racial designations often lead to a mandate on recruiting patients from different regions and racial groups in clinical trials. Of note, we use "region" loosely throughout, which could represent a center or a geographic region.

Anisimov and Fedorov [1] proposed a nice statistical approach to model and predict patient accrual in multi-center trials, in which patient accrual at different centers was assumed to follow homogeneous Poisson processes (HPP) with constant rates and maximum likelihood or method of moments were used for estimation. Since accrual in real trials often does not follow an HPP, Zhang and Long [2] proposed to use non-homogeneous Poisson processes (NHPP) with time-varying rates to model accrual and showed that the NHPP approach outperformed the HPP approach when the assumption of constant rates was not met; their subsequent investigation also showed that the NHPP and HPP approaches were comparable when this assumption held. However, the NHPP approach in [2] ignores region-specific data and thus has some limitations when used in multi-regional trials. When a region first clears regulatory hurdles and is ready for recruitment, this usually brings a "boost" or initial jump in accrual, which combined with staggered region start-up time translates into discontinuity in the underlying overall accrual rate. In other words, the overall time-varying rate for NHPP is not only non-constant but also discontinuous. The original NHPP approach cannot accommodate such features since it assumes that the time-varying rate is continuous and smooth over time, and hence it is inadequate in such settings.

In this article, we clarify the implication of the multi-regional nature on modeling and prediction of patient accrual in clinical trials and investigate Bayesian modeling and prediction of accrual in multi-regional trials. The proposed approach borrows strength of the accrual profiles across regions and accommodates staggered regional start-up times and initial jump in accrual within each region. The underlying accrual rate within each region is allowed to vary over time and is assumed to be continuous over time, but discontinuity is allowed for

the overall accrual rate over time, a more realistic assumption compared to what underlies the original NHPP approach [2]. In Section 2, we describe the proposed NHPP approach for multi-regional trials and contrast it with the original NHPP approach. In Section 3, we conduct simulation studies to investigate their performance in finite samples. In Section 4, we illustrate the proposed approach using a real clinical trial for colon cancer. We conclude this paper with discussion in Section 5.

2. Methodology

Suppose a multi-regional clinical trial plans to enroll a total of n patients in J participating regions. Let j ($1 \leq j \leq J$) index the regions; without loss of generality, let t ($t = 1, 2, \cdots$) index the time in days with the date of the study start as the reference point (i.e., $t = 1$). We write the number of patients enrolled in region j on day t as N_{jt}, which is subject to the randomness of the enrollment process; by definition, N_{jt} is a random variable. We denote its realization in the current trial, or the observed value, by n_{jt}. Suppose an interim look of accrual occurs at time T, we write the observed enrollment from all J regions by that time as $\mathbf{n} = \{n_{jt} : j = 1, \cdots, J, t = 1, \cdots, T\}$. The total number of patients enrolled across all regions on day t can be written as a random variable $N_{\cdot t} = \sum_{j=1}^{J} N_{jt}$, with its realization $n_{\cdot t} = \sum_{j=1}^{J} n_{jt}$ ($n_{jt} = 0$ for regions that have not started enrollment by time t). We are interested in predicting the time when at least a total of n patients are enrolled across all regions, namely, $\tau = \arg\min_r \left(\sum_{t=1}^{r} n_{\cdot t} \geq n \right)$.

Since the patients are independent individuals, the number of patients enrolled each day in any given region can be modeled as a random variable following a Poisson distribution. We denote the underlying Poisson rate of region j at time t by λ_{jt}, which implies that the rate is time-varying, and we further write the vector of the underlying accrual rates for all regions by time T as $\boldsymbol{\lambda} = \{\lambda_{jt} : j = 1, \cdots, J, t = 1 \cdots, T\}$. The distribution of the observed data then follows, i.e.,

$$\Pr(\mathbf{N} = \mathbf{n} \mid \boldsymbol{\lambda}) = \prod_{j=1}^{J} \prod_{t=1}^{T} \frac{e^{-\lambda_{jt}} \lambda_{jt}^{n_{jt}}}{n_{jt}!}. \tag{1}$$

2.1. Modeling Time-Varying Region-Specific Rate λ_{jt}

To model region-specific accrual, we assume that the underlying accrual rate within each region is continuous over time, or close to continuous, which can be approximated with a cubic B-spline [3]. We define the B-spline basis at time t as $\boldsymbol{\phi}(t) = \big(\phi_1(t), \cdots, \phi_q(t)\big)^T$. Note that the dimension of the B-spline (i.e., the number of basis

functions), q, is determined by the number of knots, p, i.e., $q = p + 4$. We denote the B-spline coefficients for region j by $\boldsymbol{\beta}_j$ and write the coefficients for all regions as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_J)^T$. It follows that the true accrual rate from region j at time t is $\lambda_{jt} = \boldsymbol{\beta}_j^T \boldsymbol{\phi}(t)$.

In real life multi-regional trials, it is rare for enrollment to initialize in all regions at the same time. The kick-off of the enrollment usually depends on the regional IRB/regulatory schedules and processes. As a consequence, the region start-up time usually staggers. It is not uncommon that some regions may lag behind for months, or even longer, in the trial initialization activities compared to other regions. To address this, we introduce an offset time for each region, i.e., t_j^0 for region j, which represents the time when the specific region (j) is ready for enrollment (relative to the study start) and can be considered the delay of the enrollment initiation in region j. This conceptually divides the underlying accrual rate of region j into two periods, one before t_j^0 and the other one afterwards. Clearly, patients can only be enrolled into the study during the latter period, i.e., $t \geq t_j^0$. The underlying accrual rate of region j is then expressed as

$$\lambda_{jt} = \boldsymbol{\beta}_j{}^T \boldsymbol{\phi}\big((t - t_j^0)_+\big), \tag{2}$$

where $t_+ = t$ when $t \geq 0$ and $t_+ = 0$ when $t < 0$. We note that the proposed model encompasses the original NHPP model as a special case when $J = 1$. The comparison of the two models will be elaborated in Section 2.4.

We assign a common prior distribution to govern the region-specific B-spline coefficients, $\boldsymbol{\beta}_j$. More specifically, we assume $\boldsymbol{\beta}_j \sim \text{MVN}_q(\boldsymbol{\nu}, \boldsymbol{\Gamma})$, for all $j = 1, \cdots, J$, that is,

$$f(\boldsymbol{\beta}_j \mid \boldsymbol{\nu}, \boldsymbol{\Gamma}) \quad \propto \quad |\boldsymbol{\Gamma}|^{-1/2} \exp\Big\{ -.5(\boldsymbol{\beta}_j - \boldsymbol{\nu})^T \boldsymbol{\Gamma}^{-1} (\boldsymbol{\beta}_j - \boldsymbol{\nu}) \Big\},$$

where $\boldsymbol{\nu}$ and $\boldsymbol{\Gamma}$ are pre-specified values. It follows that

$$f(\boldsymbol{\beta} \mid \boldsymbol{\nu}, \boldsymbol{\Gamma}) = \prod_{j=1}^{J} f(\boldsymbol{\beta}_j \mid \boldsymbol{\nu}, \boldsymbol{\Gamma}) \propto |\boldsymbol{\Gamma}|^{-J/2} \exp\Big\{ -.5 \sum_{j=1}^{J} (\boldsymbol{\beta}_j - \boldsymbol{\nu})^T \boldsymbol{\Gamma}^{-1} (\boldsymbol{\beta}_j - \boldsymbol{\nu}) \Big\}. \tag{3}$$

2.2. Specification of Prior Distributions

We now discuss how to determine the values of $\boldsymbol{\nu}$ and $\boldsymbol{\Gamma}$ in the prior distribution of $\boldsymbol{\beta}_j$. In a general setting of Bayesian analysis, there may lack proper prior knowledge on the model parameters, in which circumstance a noninformative prior distribution may be reasonable. We argue that this is usually not the case with accrual modeling. At the design stage, some information regarding the regional recruitment capacity should exist, for instance based on the recruitment history from earlier trials for a similar indication, from the investigators'

judgement, or a combination of the two. Those can then serve as a reasonable basis for the values of $\boldsymbol{\nu}$ and $\boldsymbol{\Gamma}$. Additionally, sensitivity analyses may be conducted to evaluate the choice of those parameter values.

Following Zhang and Long [2], we use $A_{\max}$ to denote the anticipated maximum accrual rate across all regions. Assuming regions of equal enrollment capacity *apriori*, this suggests a maximum accrual rate of $A_{avg} = A_{\max}/J$ from each region. We thereby let $\boldsymbol{\nu} = (A_{avg}, \cdots, A_{avg})^T$. Next, we introduce a coefficient of variation (cv), ρ ($0 \leq \rho \leq 1$), to quantify the variability across the regions in their deviation from the average accrual rate. That is, $\boldsymbol{\Gamma} = diag(\rho^2 \boldsymbol{\nu}^2)$, with larger value of ρ indicating larger variability in the enrollment profiles across the regions. On an empirical note, many multi-regional trials consist of a few high-enrollers (i.e., regions that each contribute a large number of patients) and many poor-enrollers (i.e., regions that each contribute a small number of patients). If this is expected to be the case, it is desirable to use a relatively large value of ρ or pool multiple poor-enrollers to form a more robust entity.

2.3. Estimation and Prediction

Given Equations (1), (2), and (3), the joint distribution of the model parameters and the data is

$$f(\boldsymbol{\beta}, \mathbf{n} \mid \boldsymbol{\nu}, \boldsymbol{\Gamma}) \propto \exp\Big\{ -.5 \sum_{j=1}^{J} (\boldsymbol{\beta}_j - \boldsymbol{\nu})^T \boldsymbol{\Gamma} (\boldsymbol{\beta}_j - \boldsymbol{\nu}) - \sum_{t=1}^{T} \sum_{j=1}^{J} \mathbf{b}_j^T \boldsymbol{\phi}\big((t - t_{0j})_+\big) \Big\} \times \prod_{t=1}^{T} \prod_{j=1}^{J} \Big[\mathbf{b}_j^T \boldsymbol{\phi}\big((t - t_{0j})_+\big) \Big]^{n_{jt}}.$$

The value of $\boldsymbol{\beta}_j$ for each region can then be updated from the following conditional posterior distribution using the *arms* function in the R package *HI*,

$$f(\boldsymbol{\beta}_j \mid \cdot) \propto \exp\Big\{ -.5 (\boldsymbol{\beta}_j - \boldsymbol{\nu})^T \boldsymbol{\Gamma}^{-1} (\boldsymbol{\beta}_j - \boldsymbol{\nu}) - \sum_{t=1}^{T} \mathbf{b}_j^T \boldsymbol{\phi}\big((t - t_{0j})_+\big) \Big\} \prod_{t=1}^{T} [\mathbf{b}_j^T \boldsymbol{\phi}\big((t - t_{0j})_+\big)]^{n_{jt}},$$

where $f(\boldsymbol{\beta}_j \mid \cdot)$ denotes the conditional posterior distribution of $\boldsymbol{\beta}_j$ given all the other parameters and the observed enrollment by time T.

We are interested in the posterior predictive distribution of τ, $f(\tau \mid \mathbf{n})$, which follows from the posterior predictive distribution of the future region-specific enrollment, $f(\tilde{\mathbf{n}} \mid \mathbf{n})$ where $\tilde{\mathbf{n}} = \{\tilde{n}_{jt}, j = 1 \cdots, J, t > T\}$. By definition, $f(\tilde{\mathbf{n}} \mid \mathbf{n}) = \int f(\tilde{\mathbf{n}} \mid \tilde{\boldsymbol{\lambda}}) f(\tilde{\boldsymbol{\lambda}} \mid \mathbf{n}) d\tilde{\boldsymbol{\lambda}}$. The future accrual can then be generated from $\tilde{n}_{jt} \sim \text{Poisson}(\tilde{\lambda}_{jt})$ for $t > T$. We write $\tilde{\boldsymbol{\lambda}} = \{\tilde{\lambda}_{jt}, j = 1, \cdots, J, t > T\}$ as the future region-specific accrual rate. It is known that the spline model can be unstable when used to make extrapolations, i.e., $\tilde{\lambda}_{jt}$ for $t > T$. Therefore, we propose to set $\tilde{\lambda}_{jt} = \lambda_{jT}$ for $t > T$, where λ_{jT} comes from the posterior distribution of $f(\boldsymbol{\lambda} \mid \mathbf{n})$, i.e.,

assuming the underlying accrual rate within each region remains constant beyond the day of the interim look, T. This approach is generally conservative, i.e., it tends to underestimate the future accrual rate, as the patient enrollment within a region usually increases before reaching a plateau (i.e., the maximum enrollment capacity), and consequently this approach tends to overestimate the time to reach the overall enrollment goal. Compared to a naive approach of projecting the observed accrual at time T as the future accrual rate, i.e., $\tilde{n}_{jt} = n_{jT}$ for all j and $t > T$, the proposed approach properly models the randomness in the observed daily enrollment and avoids projecting random highs or lows as the future accrual rate for $t > T$.

We comment that, in the above, we illustrate as an example how to make prediction for τ using the proposed model. If quantities other than τ are of interest, such as the number of subjects enrolled by a given time in the future, they can be derived from the posterior predictive distribution of $f(\tilde{\boldsymbol{\lambda}} \mid \mathbf{n})$ following the similar lines. Furthermore, as we noted in Section 2.2, sensitivity analyses can be performed based on different prior assumptions on the maximum accrual capacity, $A_{\max}$, and/or the variability across regions, ρ.

2.4. Comparison with the Original NHPP Model

Following the notation in Section 2, it can be readily shown from the properties of Poisson distributions that the overall enrollment on day t, $N_{\cdot t}$, also follows a Poisson distribution, i.e., $N_{\cdot t} \sim \text{Poisson}(\lambda_{\cdot t})$ with $\lambda_{\cdot t} = \sum_{j=1}^{J} \lambda_{jt}$, where λ_{jt}'s are defined in Equation (2). It follows that $\{N_{\cdot t}, t = 1, 2, \cdots\}$ can be modeled using a NHPP model with a time-dependent accrual rate, $\lambda_{\cdot t}$. Again, we emphasize that the original NHPP model in Zhang and Long [2] has some limitations in the case of multi-regional trials. Specifically, if there is delay in the enrollment initiation (namely, t_j^0) in some regions, then the overall accrual rate across all regions, $\lambda_{\cdot t}$, may not be smooth; if, in addition, the accrual rate at the enrollment initiation (namely, $t_{jt_j^0}$) is greater than 0, then $\lambda_{\cdot t}$ may not even be continuous. As a result, a key assumption underlying the NHPP model, i.e., the overall accrual rate $\lambda_{\cdot t}$ is smooth and can be modeled using splines, may not hold in multi-regional clinical trials. While one could still use the NHPP model, the performance may suffer as a result of these complications. The comparisons between these two types of models will be further evaluated in our simulation studies.

3. Simulation Studies

3.1. Simulation Settings

We conduct simulation studies to investigate the properties of the proposed method. Suppose that we are interested in a multi-regional trial with a sample size of 3000 and a targeted maximum daily accrual of 12 patients across all regions, i.e., $n = 3000$ and $A_{\max} = 12$. We simulate the trial accrual with either two or five regions ($J = 2$ or 5), where the true accrual rate over time within each region follows the shape of the cumulative distribution function (c.d.f.) of a Gamma random variable. This choice of the true accrual function over time is flexible with five parameters, including two parameters of the Gamma c.d.f. ($a_j > 0$ and $b_j > 0$), two parameters to scale the c.d.f. ($c_j > 0$ and $0 \leq d_j \leq 1$), and one parameter to denote the region start-up time ($t_j^0 \geq 0$), i.e.,

$$\lambda_{jt} = d_j c_j + (1 - d_j) c_j \frac{b_j^{a_j} \left[(t - t_j^0)_+\right]^{a_j - 1} \exp\left[- b_j (t - t_j^0)_+ \right]}{\Gamma(a_j)}, \text{ for } t = 1, 2, \cdots,$$

where t_+ is defined as in Section 2.1. In other words, c_j is the maximum true accrual capacity of region j and d_j represents the ratio of the accrual rate at the start of the enrollment relative to the maximum enrollment capacity. Within each region, a_j, b_j, and c_j are drawn from the following distributions, $a_j \sim \text{Unif}(0.04, 0.12)$, $b_j \sim \text{Unif}(2, 6)$, and $c_j \sim \text{Unif}(8/J, 16/J)$. These parameters allow for a considerable heterogeneity in the simulated true accrual rate across the regions. When all J regions have reached the maximum capacity (i.e., the respective c_j), the overall accrual capacity is in the neighborhood of $A_{\max} = 12$. We consider two types of scenarios for the value of d_j, one with $d_j = 0$, i.e., the overall accrual rate ($\lambda_{\cdot t}$) is continuous over time throughout, and the other one with $d_j = 0.5$, which suggests discontinuity in the overall accrual rate whenever a region first starts enrollment. We first allow the regions to stagger in their start-up time. Without loss of generality, we set the start-up time of one region as the study start, or Day 1 ($t_{j'}^0 = 0$ for region j'), and draw the start-up time for the rest regions from $t_j^0 \sim \text{Unif}(1, 150)$ for $j \neq j'$. We also investigate a simplified scenario when all regions start accrual at the same time, i.e., $t_j^0 = 1$ for all j. We denote the observed time to reach a total of n patients in the simulated dataset by τ_{true}.

For illustration purposes, we conduct two interim analyses of the accrual, one at 40% of the total targeted enrollment and the other one at 70%. We set the prior parameters at $\boldsymbol{\nu} = (A_{\max}/J, \cdots, A_{\max}/J)^T$ and $\rho = 0.1$ or $\rho = 0.3$. Both the proposed model and the original NHPP model are fitted using a cubic B-spline with three

equally spaced internal knots, i.e., $T/4$, $T/2$, and $3T/4$. For each simulated dataset, we run 5000 iterations of both models after discarding 1000 iterations of burn-in. We summarize the median of the posterior draws of τ as the predicted time of full accrual (sub-sampled to every fifth draw for a total of 1000 posterior draws), T_{exp}. The posterior credible interval of τ is then computed as the 2.5^{th} and 97.5^{th} percentile of the posterior draws of τ within each dataset, denoted by T_L and T_U. We denote the width of the 95% posterior CI by $w = T_U - T_L$.

We run a total of 1000 simulated trials under each setting and compare the proposed approach with the NHPP approach [2] using the following summary statistics, mean prediction errors, calculated as $\text{PE} = \text{E}(T_{\text{exp}} - \tau_{\text{true}})$, root mean square errors, calculated as $\text{rMSE} = \left\{\text{E}(T_{\text{exp}} - \tau_{\text{true}})^2\right\}^{1/2}$, mean coverage rate of the 95% posterior CI, calculated as $\text{CR} = \text{E}\big(\text{I}(T_L \leq \tau_{\text{true}} \leq T_U)\big)$, where $\text{I}(\mathcal{A})$ is the index function with value 1 when $\mathcal{A}$ is true. In addition, we summarize the mean width of the posterior CI of each method across the 1000 simulated datasets, i.e., $\bar{w}_1 = \text{E}w_1$ and $\bar{w}_2 = \text{E}w_2$, where the subscript $_1$ denotes the NHPP method and $_2$ denotes the proposed method. The percentage of simulated trials in which the proposed method produces tighter posterior CI, i.e., $\Pr(w_1 > w_2)$, is also provided for comparison. Of note, since Zhang and Long [2] showed that the performance of the HPP model could be very poor when the assumption of constant underlying accrual rate was not met, models using HPP are not included in the simulation studies.

3.2. Simulation Results

Table I presents the results from both approaches when the initiation of accrual is staggered among regions. We first focus on the settings where there are two regions ($J = 2$). One can observe that the proposed method always provides smaller rMSE, e.g., $\text{rMSE}_2 = 7.72$ for the proposed method when $d_j = 0$ and $\rho = 0.1$ at the first interim look, compared to $\text{rMSE}_1 = 8.71$ for the NHPP model. In the meantime, the proposed method also produces tighter posterior CI, on average, than the NHPP method, e.g., $\bar{w}_2 = 29.15$ compared to $\bar{w}_1 = 49.44$ when $d_j = 0$ and $\rho = 0.1$ at the first interim look. As a matter of fact, the proposed method almost always yields tighter CI, i.e., $\Pr(w_1 > w_2) = 1.00$, under the same $d_j = 0$ and $\rho = 0.1$ at both the first and the second looks. In the meantime, the proposed method provides comparable, if not higher, CR as the NHPP method despite the tighter posterior CI, e.g., 98.9% versus 99.3% at the second look when $d_j = 0$ and $\rho = 0.1$. When a more diffused prior distribution is utilized, i.e., $\rho = 0.3$ instead of $\rho = 0.1$, the results from both methods become more variable with higher rMSE and $\bar{w}$, nevertheless, the impact on the proposed method

is relatively smaller compared to the NHPP method. For example, the rMSE of the NHPP method increases from 8.71 to 17.97, which is more than doubled, when ρ increases from 0.1 to 0.3 with $d_j = 0$ at the first look, compared to a moderate increase from 7.72 to 11.69 with the proposed method. We also comment that the performance of both methods improve as information accumulates for the prediction, i.e., from the first look with 40% observed enrollment to the second one with 70%. The performance of both methods are generally similar when $d_j = 0.5$ compared to the respective ones under $d_j = 0$.

In the simulated settings with five regions ($J = 5$), the above observations still hold. Moreover, the proposed method provides even smaller rMSE and tighter posterior CI compared to the respective results with two regions. For instance, the rMSE reduces to 5.02 ($J = 5$) from 7.72 ($J = 2$) when $d_j = 0$ and $\rho = 0.1$ at the first look, whereas the average width of the posterior CI also reduces to 22.27 ($J = 5$) from 29.15 ($J = 2$). Nevertheless, the performance of the original NHPP method remains similar. This further demonstrates the improvement in precision of the proposed method by properly utilizing the regional accrual information.

For the simplified scenario when all regions start accrual at the same time, i.e., $t_j^0 = 1$ for all j, the results are summarized in Table II and are similar to what are observed in Table I.

In summary, the proposed method not only provides region-specific accrual prediction, but also produces tighter posterior CI of τ_{true} overall. The advantage of the proposed approach compared to the NHPP method can impact multiple aspects of trial operations, such as regional drug supply and distribution of laboratory kits, as well as proper allocation of clinical staff within each region, all attributable to the granularity of regional accrual adjusted for in the proposed model. More importantly, the tighter CIs produced by the proposed method translate into less uncertainty regarding the projected accrual, and hence better confidence when addressing issues such as slow accrual.

4. Data Example: a Real Cancer Trial

In this section we retrospectively apply the proposed method to a real oncology trial [4] and compare the results with the original NHPP method. In this randomized Phase III study of adjuvant treatments of colorectal cancer, a total of 1794 Stage III patients were planned to be enrolled from 32 countries. Although Stage II patients satisfying certain criteria were later on also enrolled per a protocol amendment, they were not included in the primary analysis, and the enrollment goal was only tracked for Stage III patients. Therefore, for the purpose

of illustration, we only focus on the Stage III patients in this analysis. Under a reasonable expectation of an enrollment duration of 24 months (523 workdays, excluding weekends), a daily enrollment of 3.45 patients is assumed. The accrual goal was met on Day 570 with a total of 1801 patients, i.e., $\tau_{\text{obs}} = 570$.

In the following, we illustrate a retro-perspective enrollment monitoring/prediction using the proposed method at two interim looks with 40% and 70% of patients enrolled, which should occur on Days 386 and 474, respectively. Among all countries participated in this study, one country contributed 30% of the patients. Therefore, we first consider grouping all countries into two regions, i.e., this one country contributing the most enrollment vs. the rest of the world. The results are shown in Table III. The predicted T_{exp} generally excludes the originally planned 523 days, indicating that it is highly unlikely to meet the original accrual goal given the observed data. Nevertheless, as the maximum overall daily accrual rate across all regions is much higher than the anticipated $A_{\max} = 3.45$, one can observe that the prediction, T_{exp}, does not include the observed truth ($\tau_{\text{obs}} = 570$) when $\rho = 0.1$, denoted by the $*$ next to the posterior CI in Table III. When the prior distribution is more diffused, e.g., $\rho = 0.3$ or 0.5, both the NHPP and the proposed method cover τ_{obs}. In other words, one can still make a reasonable accrual projection based on the data despite a dubious yet diffused prior distribution. As in the simulation studies, we generally observe tighter posterior CIs with the proposed method, e.g., $w_2 = 52$ compared to $w_1 = 58$ when $\rho = 0.5$ at the first look and $w_2 = 33$ compared to $w_1 = 37$ at the second look. We also compare an *ad hoc* determination of $A_{\max}$ which is based on the observed accrual in the past 20 workdays at the time of the interim look. As shown in Table III, the prediction improves. Notably, all the posterior CIs of the proposed method cover τ_{obs}, even when $\rho = 0.1$.

Following the highest enrolling country, there were five countries each contributing between 5% and 11% of the patients. The rest of the countries each contributed a maximum of 4% patients. Hence, based on the enrollment capacity, one may group all countries into three categories, i.e., high enrolling region (the one country with the most enrollment), medium enrolling region (the five countries with moderate enrollment), and low enrolling region (the rest of the countries). As shown in Table III, the proposed method produces similar results as those with two regions, with the exception at the first look when $\rho = 0.3$, which narrowly misses τ_{obs} with a 95% posterior CI of $[574, 632]$.

Alternatively, one can set the highest enrolling country as a region and group the rest of the countries based on their respective enrollment start-up time, for instance those countries that initialize enrollment within 50 days from the trial start, between Days 51 and 100, between Days 101 and 200, and beyond Day 201. These

four regions consist of four, seven, five, and fifteen countries respectively. For a total of 5 regions ($J = 5$), the results remain similar as shown in Table III, which suggests the robustness of the proposed method with respect to different ways to define the regions.

Overall, the proposed method performs reasonably well and further improves on the NHPP method with tighter posterior CIs. The sensitivity analyses on ρ ($\rho = 0.1$, 0.3 and 0.5) as well as the number of regions ($J = 2$, 3, and 5) both suggest the robustness of the proposed method. Had the accrual monitoring tool been available during the trial, it could have detected issues with accrual relatively early and alerted the study team, or the coordinating data center, for proper actions.

5. Discussion

In this paper, we propose a Bayesian modeling approach for patient accrual in multi-regional clinical trials, which models region-specific accrual using NHPP and allows the underlying overall accrual rate to be discontinuous and change over time. Our numerical studies show that the proposed method improves precision of accrual prediction (i.e., tighter posterior predictive credible intervals) compared to the original NHPP approach that ignores region-specific data [2]. In practice, improved precision of prediction leads to improved decision making on resource allocation. Generally speaking, the proposed method also allows the research team to identify potential enrollment problems with certain regions and hence enable the team to address the problem with a more targeted approach, such as stressing the detected deficiency in recruitment for certain regions or adding satellite regions to increase enrollment.

Along the lines of [2], Deviance Information Criterion (DIC) can be used to perform model selection to select optimal spline models for region specific accrual rates and evaluate goodness of fit. In particular, a model assuming that the region specific accrual rate is constant, i.e., $\lambda_{jt} = \lambda_j$, is a special case of the proposed model, which could be selected by DIC in cases where the constant accrual rate assumption is indeed met. In the numerical studies, we observe that the choice of the prior parameters ($A_{\max}$ and ρ) has an effect on the results to certain degree. Therefore, we recommend a possibly iterative approach to adjust the prior parameters. For instance, if the fitted profile of the underlying accrual rate deviates substantially from the initial guess, as in the real data example, one may want to adjust the values of $A_{\max}$ and/or ρ accordingly and refit the model. Alternatively, one may resort to an *ad hoc* determination of $A_{\max}$ as discussed in Section

4, e.g., taking $A_{\max}$ as the average of the recent observed accrual.

In clinical trials with time-to-event endpoints, the statistical power is primarily driven by the number of events observed. In such cases, monitoring and prediction of event times is very important and has been investigated in the literature [5, 6]. It is of interest to extend the proposed approach to monitor and predict both patient accrual and event times in multi-regional clinical trials.

REFERENCES

1. Anisimov VV and Fedorov VV. Modelling, Predictions and Adaptive Adjustment of Recruitment in Multicentre Trials. *Statistics in Medicine* 2007; **26**:4958-4975. DOI: 10.1002/sim.2956
2. Zhang X and Long Q. Stochastic modeling and prediction for accrual in clinical trials. *Statistics in Medicine* 2010; **29(6)**:649-658. DOI: 10.1002/sim.3847
3. de Boor C. *A Practical Guide to Splines*. Springer-Verlag: New York, 1978; 113-114.
4. Van Cutsem E, Labianca R, Bodoky G, Barone C, Aranda E, Nordlinger B, Topham C, Tabernero J, Andre T, Sovrero A, Mini E, Greil R, Di Costanzo F, Colette L, Cisar L, Zhang X, Khayat D, Bokemeyer C, Roth A, Cunningham D. Randomized Phase III Trial Comparing Biweekly Infusional 5-Fluorouracil/Leucovorin Alone or with Irinotecan in the Adjuvant Treatment of Stage III Colon Cancer: PETACC-3. *Journal of Clinical Oncology* 2009; **27**:2117-2125.
5. Ying GS, Heitjan DF, Chen TT. Nonparametric prediction of event times in randomized clinical trials. *Clinical Trials* 2004; **1**:352361.
6. Ying GS, Heitjan DF. Weibull prediction of event times in clinical trials. *Pharmaceutical Statistics* 2008; **7**:107-120.

Table I. Comparison of mean prediction errors (PE), root mean squared errors (rMSE), mean coverage rates (CR), and mean width of the 95% posterior CI ($\bar{w}$) of the width from the proposed method (prop.) and the NHPP method, as well as the probability of the proposed method having tighter 95% CI than the NHPP method (Prob.), i.e., $\Pr(w_1 > w_2)$, based on 1000 simulated trials with two/five regions, when the region start-up staggers.

Settings				PE		rMSE		CR		w		Prob.
J	d_j	ρ	pct.	NHPP	prop.	NHPP	prop.	NHPP	prop.	NHPP	prop.	$w_1 > w_2$
2	0	0.1	40%	0.54	3.11	8.71	7.72	0.99	0.94	49.44	29.15	1.00
			70%	-0.68	-0.56	4.74	3.28	0.99	0.99	23.68	15.54	1.00
		0.3	40%	1.77	2.14	17.97	11.69	0.96	0.96	74.88	49.91	0.99
			70%	-1.28	-0.72	8.12	5.27	0.95	0.98	32.34	24.00	0.98
	.5	0.1	40%	0.81	2.66	8.53	6.80	1.00	0.98	50.80	34.11	1.00
			70%	-0.64	-0.46	4.45	3.28	1.00	1.00	24.19	17.41	1.00
		0.3	40%	2.01	2.54	18.73	13.37	0.96	0.97	80.27	59.99	0.98
			70%	-0.99	-0.69	7.92	5.78	0.97	0.98	33.99	26.76	0.97
5	0	0.1	40%	1.85	2.46	8.60	5.02	0.99	0.98	49.15	22.27	1.00
			70%	-0.69	-0.27	4.50	2.78	0.99	0.98	23.14	12.68	1.00
		0.3	40%	4.72	5.00	18.67	9.41	0.95	0.96	73.62	38.24	1.00
			70%	-1.04	-0.57	7.61	3.84	0.96	0.99	30.85	18.34	1.00
	.5	0.1	40%	1.83	1.62	8.67	4.74	1.00	0.99	50.48	23.39	1.00
			70%	-0.63	-0.22	4.59	2.84	0.99	0.99	23.56	13.21	1.00
		0.3	40%	4.54	4.64	18.30	9.61	0.97	0.98	77.95	42.67	1.00
			70%	-0.96	-0.50	7.86	4.23	0.96	0.98	32.31	19.86	1.00

Table II. Comparison of mean prediction errors (PE), root mean squared errors (rMSE), mean coverage rates (CR), and mean width of the 95% posterior CI ($\bar{w}$) of the width from the proposed method (prop.) and the NHPP method, as well as the probability of the proposed method having tighter 95% CI than the NHPP method (Prob.), i.e., $\Pr(w_1 > w_2)$, based on 1000 simulated trials with two/five regions, when all regions initialize accrual at the same time.

Settings				PE		rMSE		CR		w		Prob.
J	d_j	ρ	pct.	NHPP	prop.	NHPP	prop.	NHPP	prop.	NHPP	prop.	$w_1 > w_2$
2	0	0.1	40%	-1.45	-0.57	8.24	4.78	1.00	1.00	50.16	28.44	1.00
			70%	-0.90	-0.57	4.63	3.18	0.99	0.99	24.45	15.97	1.00
		0.3	40%	-1.87	-1.49	18.21	10.00	0.96	0.99	79.32	50.58	1.00
			70%	-1.09	-0.96	8.38	5.37	0.97	0.99	34.88	25.81	0.98
	.5	0.1	40%	-0.80	-0.51	8.09	5.39	1.00	1.00	51.65	37.91	1.00
			70%	-0.88	-0.63	4.63	3.58	0.99	1.00	24.92	19.87	1.00
		0.3	40%	-1.20	-1.38	18.38	14.13	0.97	0.98	85.21	72.22	0.96
			70%	-0.99	-1.14	8.45	7.06	0.98	0.98	36.24	32.54	0.88
5	0	0.1	40%	0.24	0.44	8.33	4.09	1.00	0.99	50.55	21.77	1.00
			70%	-0.75	-0.47	4.44	2.68	1.00	1.00	24.20	12.90	1.00
		0.3	40%	1.77	1.04	18.78	7.40	0.97	0.99	80.69	39.69	1.00
			70%	-0.82	-0.75	7.95	3.90	0.97	0.99	34.21	20.30	1.00
	.5	0.1	40%	-0.52	-0.67	7.95	4.11	1.00	1.00	51.41	26.68	1.00
			70%	-0.94	-0.60	4.33	2.83	1.00	0.99	24.50	15.16	1.00
		0.3	40%	-0.67	-0.65	18.29	9.33	0.98	1.00	84.03	55.84	1.00
			70%	-1.19	-1.21	7.90	5.00	0.97	1.00	35.52	26.63	1.00

Table III. Data Example: 95% posterior credible intervals (CI) of the predicted accrual duration (τ) and $w = T_U - T_L$ which is the width of the 95% posterior CI, using the proposed method versus the original NHPP method.

Real data			Original		Proposed					
					Two regions		Three regions		Five regions	
ad hoc	cv	pct.	CI	w_1	CI	w_2	CI	w_2	CI	w_2
No	0.1	40%	[596, 657]*	61	[632, 686]*	54	[643, 691]*	48	[656, 703]*	47
		70%	[579, 611]*	32	[592, 622]*	30	[595, 622]*	27	[598, 625]*	27
	0.3	40%	[533, 595]	62	[561, 623]	62	[574, 632]*	58	[598, 654]*	56
		70%	[555, 593]	38	[564, 598]	34	[565, 593]	28	[570, 598]	28
	0.5	40%	[526, 584]	58	[538, 590]	52	[543, 599]	56	[567, 622]	55
		70%	[552, 589]	37	[558, 591]	33	[556, 593]	37	[560, 588]	28
Yes	0.1	40%	[594, 660]*	66	[547, 587]	40	[550, 584]	34	[556, 588]	32
		70%	[578, 611]*	33	[552, 574]	22	[552, 570]	18	[551, 568]	17
	0.3	40%	[533, 593]	60	[528, 585]	57	[526, 583]	57	[540, 586]	46
		70%	[553, 589]	36	[549, 581]	32	[546, 576]	30	[549, 570]	21
	0.5	40%	[525, 594]	69	[522, 582]	60	[518, 567]	49	[533, 586]	53
		70%	[548, 591]	43	[552, 587]	35	[549, 579]	30	[545, 572]	27

$*$: denotes the 95% posterior CI does not cover the true value ($\tau = 570$).