A Simple and Robust Model for Accrual Modeling

Xiaoxi Zhang and Bo Huang
Global Product Development, Pfizer Inc.

ABSTRACT

Patient accrual projection is a topic gaining attention in the statistics literature in recent years. A number of methods have been proposed in this area. Some approaches are sophisticated but complicated to implement. We aim to implement a simple and robust empiric Bayes model that is suitable for practical use. We assume the underlying enrollment rate constant over time, which is site-specific and comes from a common Gamma distribution. Choice of prior parameters can be data driven. We tested the model in a number of internal oncology trials with various enrollment patterns, which yields satisfactory results. Compared to a flexible nonparametric model (Zhang and Long, 2010), the new model was associated with narrower credible intervals as a result of parametric assumptions. With the caveat that the model prediction can be off when the model assumption was substantially violated. R codes were available upon request.

INTRODUCTION

By nature, patient enrollment in a clinical trial is a counting process, in which the number of patients enrolled, $\{N(t), t \geq 0\}$, is non-negative, integer, and non-decreasing. When modeling clinical trial enrollment using a counting process, a few dimensions of the model can be considered:

- Overall or center-specific enrollment
- Parametric or non-parametric models
- Constant or variable enrollment rate over time

In the literature, a number of authors proposed various parametric approaches for enrollment modeling. Tang et al. (2012) assumed three phases of underlying overall enrollment. Lan, Tang, and Heitjan (2019), Urbas, Sherlock, and Metcalfe (2020) assumed declining underlying enrollment at each site due to competing accrual. Gajewski et al. (2008) assumed the overall underlying enrollment rate as constant. Additional prior choices were investigated Jiang et al. (2015) (R package *accrual* available). Anisimov and Federov (2007) assumed constant underlying enrollment rate at each center and focused on the special (and arguably unrealistic) case of simultaneous site start-up. Zhang and Long (2010) proposed a non-parametric underlying enrollment model, using cubic spline as underlying enrollment rate.

In this poster, we aim to implement a simple and robust empiric Bayes model that is suitable for practical use and will compare with the method by Zhang and Long (2010).

METHODS

Review of Nonhomogeneous Poisson Process Model (NHPP)

Consider a discrete-time Poisson process following Zhang and Long (2010). Let Nt denote the number of enrollment on Day t. A regression spline is used to model the overall underlying enrollment λt with prior distribution for the spline parameters, $b \sim \boldsymbol{MVN}(\mu, \Sigma)$.

When $\lambda_i = \lambda$, the above reduces to a homogeneous Poisson process over time, i.e., constant underlying enrollment rate overall. This non-parametric model is flexible yet complicated.

A Simple and Robust Model

In plain English, the assumptions of the proposed model are:

- Sites initialize at different times during a trial.
- Once initialized, enrollment rate is stable at a particular site with random fluctuations.

In mathematical terms, we assume $\lambda_j > 0$ as the underlying enrollment rate at Site j once accrual starts at that site at time S_j. It follows that the number of patients enrolled on Day t at Site j, $n_{tj} \sim$ Poisson(λ_j), or

$$\Pr(n_{tj}|\lambda_j) = \frac{e^{-\lambda_j}\lambda_j^{n_{tj}}}{n_{tj}!}, \text{ for } t \geq S_j,$$
$$\text{or } n_{tj} = 0, \text{ for } t < S_j.$$

which means waiting time between two consecutive patients at Site j follows an exponential distribution, $Exp(\lambda_j)$.

Let us assume a Gamma prior, $\boldsymbol{\Gamma}(a, b)$, on the underlying enrollment rate at each site, i.e.,

$$\Pr(\lambda_j|a, b) = \frac{b^a}{\Gamma(a)}\lambda_j^{a-1}e^{-b\lambda_j}.$$

As a result, closed form of posterior distribution exists due to conjugate prior, i.e., λ_j follows $\boldsymbol{\Gamma}(a + \sum_{t=S_j}^{T} n_{tj}, b + T - S_j + 1)$, where S_j is the start of enrollment at Site j.

In the event a site has not started enrollment at the time of enrollment projection, we can simply assume the underlying enrollment rate as a random variable following $\boldsymbol{\Gamma}(a, b)$.

In an empirical Bayes paradigm, the values of a and b are data driven. One can use observed average daily enrollment, $\hat{\lambda}_j = \frac{\sum_{t=S_j}^{T} n_{tj}}{T - S_j + 1}$,

As a crude estimate of λ_j and then use the mean (a/b) and variance (a/b^2) of these estimates to estimate a and b.

RESULTS

Real Example 1

700 patients were randomized from 197 centers between 28Apr2016 and 09Jun2019. An enrollment projection made when enrollment is 100 patients shy of the target of 70% enrollment (achieved on Day 797) would be on Day 679 (Figure 1a).

Figure 2 shows enrollment projection made at various points when enrollment is 20, 60, 100, and 140 patients shy of the target using both NHPP (broken black line) and the proposed (broken green line) methods. The 90% confidence bands of the proposed method covers true observed enrollment beyond the projection date well (solid red line). As expected, the uncertainty in prediction reduces when prediction is made when more patients were enrolled as the 90% CI assumes a funnel shape. In all scenarios, the 90% CI of the proposed method covers the observed enrollment well, whereas the NHPP method slightly missed when the projection was made early (i.e., 140 patients shy of the target).

Further numerical comparisons of the two methods show the proposed method with narrower 90% CIs, smaller rMSEs and higher coverage probabilities than the NHPP method (Table 1).

Real Example 2

886 patients randomized from 156 centers between 23Mar2016 and 13Dec2017. An enrollment projection made when enrollment target of 70% enrollment (achieved on Day 512) is 100 patients shy would be on Day 469 (Figure 1b).

Similar conclusions can be drawn as in Real Example 1 (Figure 3 and Table 2).

Figure 1. Observed Enrollment in Two Real Oncology Trials

(a) 1L UC study (b) 1L RCC study

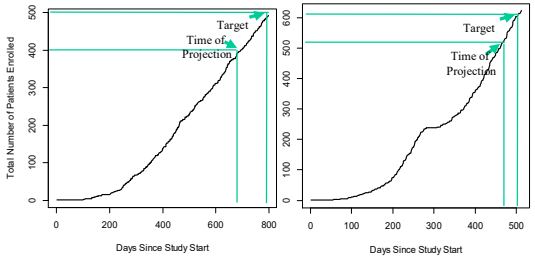

Figure 2. Projection When 20/60/100/140 Patients Shy of Target – Real Example 1

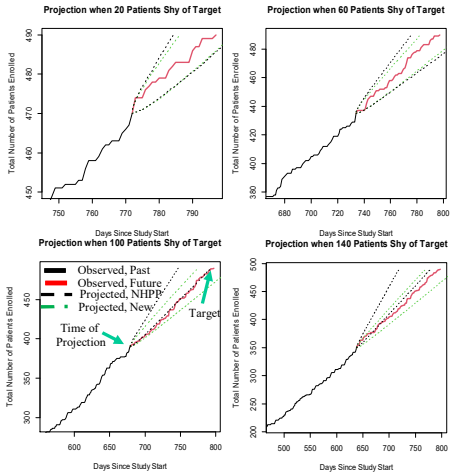

Table 1. Comparison of the New Method and the NHPP Method – Real Example 1

	# pts needed till target	Actual/Projected Time to Target (days)	90% CI Width (days)	rMSE (days)	Pr(Projected > Actual Time to Target)
NHPP	20	25/20.4	18.8	7.6	19%
	60	63/63.5	44.3	14.0	46%
	100	118/87.6	47.5	34.1	3%
	140	159/108.7	59.9	53.6	1%
NEW	20	25/21.7	16.2	6.0	24%
	60	63/63.1	30.7	9.2	48%
	100	118/118.4	47.5	14.4	49%
	140	159/159.7	58.2	17.6	50%

Figure 3. Projection When 20/60/100/140 Patients Shy of Target – Real Example 2

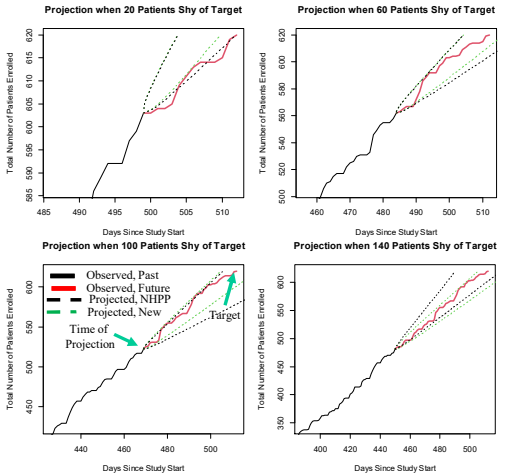

Table 2. Comparison of the New Method with NHPP Method – Real Example 2

	# pts needed till target	Actual/Projected Time to Target (days)	90% CI Width (days)	rMSE (days)	Pr(Projected > Actual Time to Target)
NHPP	20	13/8.2	8.1	6.6	5%
	60	28/27.8	17.0	5.3	43%
	100	43/49.8	26.2	10.5	79%
	140	62/52.0	28.4	13.3	13%
NEW	20	13/7.5	6.0	5.8	5%
	60	28/26.1	12.3	4.2	30%
	100	43/43.8	16.6	5.1	55%
	140	62/66.1	22.6	8.0	71%

CONCLUSIONS

In this poster, we presented a simple and robust empiric Bayes model for enrollment modeling. The proposed method properly models random fluctuation as well as site variability in enrollment and generally performs better than the more complicated NHPP method. We obtain narrower credible intervals with the proposed method as a result of parametric assumptions. However, if the model assumptions are off, prediction could be off on a streak. R function is available upon request.

REFERENCES

-Tang G, Kong Y, Chang CCH, Kong L, and Costantino JP, Prediction of Accrual Closure Date in Multi-center Clinical Trials with Discrete-time Poisson Process Models, *Pharmaceutical Statistics* 2012, 11 351-356.
-Lan Y, Tang G, and Heitjan DF, Statistical Modeling and Prediction of Clinical Trial Recruitment, *Statistics in Medicine* 2019, 38:945-955
-Urbas S, Sherlock C, and Metcalfe P, Interim Recruitment Prediction for Multi-centre Clinical Trials, manuscript 2020
-Gajewski BJ, Simon SD, and Carlson SE, Predicting Accrual in Clinical Trials with Bayesian Posterior Predictive Distributions, Statistics in Medicine 2008; 27:2328-2340.
-Jiang Y, Simon S, Mayo MS, and Gajeski BJ, Modeling and Validating Bayesian Accrual Models on Clinical Data and Simulations Using Adaptive Priors, Statistics in Medicine 2015;34:613-629.
-Anisimov VV and Federov VV, Modeling, Prediction and Adaptive Adjustment of Recruitment in Multicentre Trials, Statistics in Medicine 2007;26:4958-4975.
-Zhang X and Long Q, Stochastic Modeling and Prediction for Accrual in Clinical Trials, Statistics in Medicine 2010, 29(6) 649-658. https://doi.org/10.1002/sim.3847