Discussion on the paper "Real-Time Prediction of Clinical Trial Enrollment and Event Counts: A Review", by DF Heitjan, Z Ge, and GS Ying

Vladimir V. Anisimov

The paper by Heitjan *et al* [11] provides very interesting and useful review of the methods for predicting patient enrollment and event counts in clinical trials. The aim of this letter is to raise an additional discussion on some points and to provide readers with more comprehensive information and clarification of particular methods/techniques.

First, it would be useful to specify that there are two basic stages in predicting patient enrollment and various events:

1. Start-up (baseline) prediction before trials starts and therefore there is no real trial data available yet, and
2. Interim prediction where it is possible to use real trial data and update (re-project) trial behaviour for the remaining period.

At both stages, good predictive techniques potentially can use similar models, only input parameters will be evaluated differently.

Since the trial start-up stage is not reflected in detail in [11], it seems expedient to devote some time to this.

1 Trial start-up stage

This stage may also include an early stage of the trial where not many centers are initiated and not many patients have been recruited yet. Typically during this stage, the basic input information that is provided by clinical teams for enrollment predicting includes the following key elements:

(a) total number of randomized patients (sample size); (b) expected number of screened patients and screening duration; (c) list of regions and countries to be involved into the study: (d) planned number of centers to be initiated in each country and some information about the expected schedule of initiation; (e) expected enrollment rates in centers or countries (this may include screening/enrollment rates and dropout probabilities).

This information has many uncertainties. In particular, at the start-up stage we may not know the exact schedule of center's initiation and we especially cannot predict the exact screening/enrollment rates and dropout probabilities.

Therefore, one of the main problems at this stage is how to account for these uncertainties and evaluate trial enrollment feasibility. There is no universal approach since the solution may depend on data availability. If we have a similar historical study (similar therapeutic indication, inclusion/exclusion criteria, etc.) conducted in the same regions, then this information can be used to create the initial trial enrollment design. Specifically, the enrollment rates for a new trial can be treated as random variables with some prior distribution where parameters can be evaluated using historical data in these regions and some prior information. As the rates are positive, it is natural to use a gamma distribution.

For a new trial, we can also assume that the centers in the regions can be initiated in time according to some distributions where parameters are estimated using historical data. As usually teams for each country/region may provide some time intervals where a given number of centers is planned to be initiated, then at the first instance we can assume that the times of initiation are distributed uniformly in these intervals [3,4]. If some historical information about initiation dates is available, other types of distributions can be also used.

Note that during the start-up stage there can be a rather long transient period until most of the centers will be initiated. Thus, the total number of patients and centers may not be too large. Therefore, during this period it is important to account for the process of centers initiation and the methods based on modeling enrollment in the individual centers are more preferable compared to models based on global prediction.

1.1 Poisson-gamma enrollment model

On this way we are naturally coming to using a so-called Poisson-gamma enrollment model (P-G model) developed in [1-5]. This model assumes that the patients arrive at clinical centers according to delayed doubly stochastic Poisson processes where the variation in rates between different centers is modelled using a gamma distribution. The delays in center's initiation also can be random.

This model is very flexible as it provides the opportunity to model the enrollment on different levels (center, country, region, trial) and has many additional features, e.g. predicting with credibility bounds, predicting probability to complete in time, evaluate effects of changing the number of centers, etc. One of the additional advantages of P-G model is that most of these characteristics can be calculated using closed-form expressions, thus, there is no need to use Monte Carlo simulation.

Note that Carter *et al.* [10] also modelled variation in rates of corresponding Poisson processes but using a uniform distribution. However, this approach has some limitations as it assumes that the rates are bounded in some interval. Moreover, the analysis of many real trials shows that the empirical distributions of the rates are rather far from uniform distribution and heavy tailed.

In the framework of P-G model, at the start-up stage as input data it should be provided the expected means and standard deviations of the enrollment rates (on center or country level) to estimate the prior parameters of the rates used in prediction. These values can be evaluated using historical data from similar trials and information provided by clinical teams.

If there is no information from similar trials, then we can use the planned/expected rates provided by clinical teams weighted with some expert estimators. This data can be used as sample statistics for evaluating the prior parameters of P-G model (on country or regional level). Some discussion on using baseline estimates of rates at the trial start-up was provided in [3,5]. Bakhshi *et al.* [9] investigated P-G model further and suggested the empirical way to set the prior parameters by using the results of the meta-analysis.

As a separate set of input data for P-G model, the information about the process of the center's initiation should be provided. The case where the times of initiation have uniform distribution was considered in [3,4]. In this case, the closed-form expressions for predictive characteristics were derived.

Note that at start-up and early stages other approaches based on models for global enrollment, e.g. using Poisson models with global gamma distributed rate [13,17], and Brownian (Lai *et al* [13]) or fractional Brownian (Zhang & Lai [17]) motions may not be appropriate as in general at these stages there is a small number of active centers and patients recruited.

Therefore, on my opinion, P-G model is rather flexible and can be applied to the vast majority of trials at start-up and early stages.

2 Interim stage

At this stage, it is natural to use real data and re-estimate parameters of the model with the purpose to adjust to real data and improve accuracy of prediction of the remaining enrollment. Thus, it is typically assumed that there is already some number of active centers that enrolled a reasonable number of patients (enough to use statistical estimations). Therefore, the methods and results may depend on trial goals and data availability.

There can be other tasks at the interim stage including evaluating enrollment performance and other operational characteristics, detecting outliers, etc. However, this interesting direction may lead us outside the current discussion.

Most papers by other authors are mainly dealing with prediction of global enrollment and there are two basic directions. One is using mixed Poisson processes where the global

rate is modelled using different approaches [12,13,16,17]. Another one uses Brownian or fractional Brownian motions [14,18].

A brief review of the papers related to these directions is provided by the authors [11] in Sec. 3 “Predicting Accrual”. However, I would argue with the classification of the models (or two streams) proposed in Sec. 3.1. It seems rather artificial as actually the first stream should also involve modeling of enrolment. The second stream potentially can use modeling for predicting future trends and therefore time to reach targets, as well. In addition, the description of the papers related to using random-effect models in Sec. 3.3 is done rather schematically. As the use of P-G model is receiving further attention and development in papers of different authors, it seems expedient to provide more details here.

2.1 Use of a Poisson-gamma enrollment model

In the framework of P-G model [1-5], the enrollment processes at different levels are modelled as non-homogeneous Poisson processes with time-dependent and in general random rates, which are governed by the processes of center’s opening and closing as well as individual center’s data. Together with modeling enrollment at the start-up stage, P-G model can be efficiently applied to an interim prediction. The input is enrollment data (for each center, the duration of active enrollment and the number of patients recruited). Using this data, the parameters of a gamma distribution of the enrollment rates are estimated using ML procedure (on global or regional level). Then in each center the posterior rate is re-estimated using individual data and the Bayesian procedure. The posterior rates also have gamma distributions with different parameters depending on interim data due to the property of conjugate distributions (Poisson and gamma). These rates can be used to create the predictions of the remaining enrollment on different levels and evaluate other characteristics.

The technique based on using P-G model has several advantages compared to other approaches: it accounts for multiple center’s effects, different times for opening and closing centers, allows predicting in a closed form the mean number of recruited patients with credibility bounds (on different levels), predicting credibility bounds for time to complete enrollment and probability to complete in time. One of the essential features is the opportunity to evaluate the interim adaptive adjustment (if enrollment is going slower as expected, evaluate the number of new centers needed to be added with the purpose to complete enrollment in time with a given confidence). In addition, this technique has several other features that are available only in this framework, e.g., predicting center/country performance, number of “empty” centers, creating optimal enrollment design [1-5].

It also seems expedient to raise some discussion on using formulae compared to Monte Carlo simulation. As for rather general scenarios the most of characteristics can be calculated using closed-form expressions (explicit formulae), then there is no need to use Monte Carlo simulation. The availability of formulae has advantages as it allows to investigate the functional dependence on different parameters (number of sites, vector of rates, center’s delays, etc.) and, thus, analyse in real time the impact of various factors, perform sensitivity analysis and find the optimal solutions, which would be hard to archive using simulation. In addition, simulation may not work well for evaluating small tail and risk probabilities, P-values and also may lead to large errors in small regions.

Note also that in some cases of special restrictions on enrollment and more complicated assumptions, it may be difficult to derive formulae. In these cases Monte Carlo simulation can be the natural choice.

I would also like to correct the author’s statement in Sec. 3.3 [11] that “Mijoule *et al.* [15] proposed replacing the gamma with a Pareto mixture.” Actually in [15] the authors investigated further properties of P-G model, compared them with the Pareto-Poisson model using real datasets from [2], and also investigated the feasibility of the model. In final conclusions the authors “recommend the use of the Poisson-Gamma, which is easier to handle”, and also recommend using a uniform distribution for centers initiation when the opening dates of the centers are not known precisely, which is proposed in [3,4].

It would also be interesting to provide some parallel between Williford *et al* [17], Gajewski *et al* [13] and P-G model. In both papers the authors use a Poisson process with gamma distributed rate to model the global enrollment and also a Bayesian interim adjustment. Note that in the framework of P-G model the global enrollment in general is not described by a Poisson-gamma process as the sum of gamma distributed variables in general does not follow a gamma distribution. Nevertheless, this sum for a large number of summands can be well approximated by a gamma distributed variable [5]. Thus, P-G model can also serve as a justification of model [13,17] on a global level.

It is also worth noting that in the framework of P-G model, the interim prediction can account for the opportunity in the future to open or close some centers [5]. Thus, the predictive processes at different levels are in general non-homogeneous doubly stochastic Poisson processes where the global and individual rates depend on the processes of initiation, closing centers and individual rates. This feature of P-G model profitably differentiates it from the other models for global prediction based on using Poisson models [12,13,17] and Brownian and fractional Brownian motions [14,18], as in these papers it is assumed that the predictive process is time-homogeneous with constant parameters estimated at interim time.

2.2 Modeling trends

Here I would like to add some discussion to Sec. 3.4 [11] “Modeling trends in the Poisson rate“. Actually the author’s statement “The models described thus far all assume a constant mean enrollment rate per center, allowing the overall enrollment rate to change only as centers enter or leave the trial” related to cited papers on Poisson models with random effects does not reflect the state-of-art.

While using P-G model, the parameters of enrollment rates are re-estimated at any interim time using real data without regard to whether the number of centers is changed or not. Therefore, the posterior rates are actually time dependent as depend on interim data. In fact the idea of interim estimation for P-G model is in some sense similar to estimating the enrollment rates in [12] by using spline models and interim data, however the procedure of estimation is different as in P-G model data is used to re-estimate the parameters of gamma distributions of the rates (on global or regional levels), but in [12] data is used to re-estimate the spline parameters of the rates.

Therefore, P-G model can be naturally interpreted compared to spline models, and in addition it has many other features that cannot be directly incorporated by using global enrollment models, e.g. predicting individual center performance, adaptive adjustment of enrollment, evaluating the optimal number of centers, etc.

It is also worth to note that P-G model allows to consider time-dependent rates $\lambda_i = \lambda_i(t)$ as it is noted in [5]. The only question is - how to introduce the dependence on time.

In Sec. 3.4 [11] the authors reviewed a few papers devoted to this topic where researchers considered special types of dependence of enrollment rate on time, in particular Tang *et al* [16]. However, some types of dependence may essentially depend on the type of the trial (therapeutic area, region, seasonable variation, etc.). The verification of such models will require data from many similar trials with a specific type of dependence, which is not realistic to achieve. Moreover, at the trial start-up due to lack of data it would not be possible to estimate the additional parameters related to time dependence.

One more restriction of the model proposed in [16] is related to the assumption that the trial has a rather long stable period where the global rate is some unknown constant. However, for not so long trials the stable period may be reached only closer to the end or may never be reached. Thus, some additional discussion related to models in [16] is provided in [7]. Of course the notes above should not prevent researchers from further investigation of time-dependent models; however, the use of these models may be restricted to special types of trials having similar behaviour in the past.

2.3 Brownian motion models

Here I would like to raise some additional discussion related to Sec. 3.5 [11]. Actually the models based on using Brownian (Lai *et al* [14]) or fractional Brownian (Zhang & Lai [18]) motions potentially can behave well for predicting global enrollment in large trials during rather stable period where there are many centers and most of them are already initiated.

While using P-G model, the global enrollment process can be also approximated by a Brownian motion. Thus, both techniques should provide similar results for modeling the global trend. However, Brownian and fractional Brownian motions are not oriented to model enrollment in individual centers. Thus, these models are not appropriate at start-up and early stages (for small number of centers) and also may not work well at transient periods where a large proportion of centers is in the process of opening or closing. However, P-G model can handle these cases and has many other features described above in Sec. 2.1.

Also note that an interesting discussion on the argumentation of using Brownian and fractional Brownian motion in comparison with P-G model and approximations of the global enrollment [1] is given in (Zhang & Lai [18]).

2.4 Critique

Here I would like to add some additional discussion related to Sec. 3.6 [11] – "Critique". I fully agree with the author's statement "A key point in modeling accrual is the inclusion of information on centers." However, in large trials we should not expect that "adding a new center will have a detectable effect on accrual". For example, for a trial with 100 centers adding one center increases on average the global rate on 1% and thus reduces the remaining enrollment time only on a few days for one year enrollment duration.

Let me also comment on the applicability of techniques based on using P-G model (the authors call it "hierarchic models"). Actually it would be enough to have at least 10-20 centers for a reasonably good estimation of parameters of P-G model to be used for prediction. Practically all Phase III and the vast majority of Phase II trials satisfy this condition. For less number of centers P-G model can also be used where the variation in rates can be evaluated by using historical data or expert estimates for similar trials.

Here is some clarification to the last paragraph in Sec. 3.6 [11]. Actually the author's statement about the "prediction method" is true only if the prediction is based on the interim data without any knowledge about the future. However, in real trials clinical teams may know the schedule of future opening/closing of some centers and some information about enrollment rates in the new centers. As soon as this information is available, P-G model can be used for predicting future enrollment and the time to complete [5].

It would also be useful to distinguish between models for evaluating the expected enrollment rates for new centers (it can be regression models, weighted models using historical and current information, etc.) and the models for predicting enrollment where the rates estimated using other models are used as input data.

Consider also the comparison of some predictive methods. Actually during the stable period the basic models for the mean trend of global enrollment should behave similar. This includes Poisson type models [13,17], models with spline approximation [12], Brownian and fractional Brownian motions [14,18], and P-G model. However, these techniques except P-G model may not be suitable at the trial start-up and during the transient period and also for other types of analysis (enrollment performance, adaptive adjustment, etc.).

I also fully support the author's statement "Our sense is that prediction methods should actively discount older information." Thus, potentially the techniques at interim prediction should use some moving window for estimating parameters. This window should be rather long to have enough data for good estimators but realistically not longer than several months (say, up to 6 months) to use the latest information about the trial.

3 Predicting Event Counts

Sec. 4 of [11] is devoted to the review of the methods for predicting event counts. The authors provide a rather detailed survey and focus mainly on the results for predicting the number of clinical events (such as death or disease progression) in so-called event-driven trials using as example REMATCH study.

To provide readers with a wider picture of current state-of-art, it would be expedient to mention the results of paper [6] which is not reflected there. In this paper an analytic technique for predicting the number of events together with ongoing enrollment in event-driven trials is developed. The technique accounts for the events that may happen for patients already at risk in the trial and for events that may happen for patients that will be recruited in the future. The enrollment is modelled using P-G model [1-5], and the process of events is modelled on the top of enrollment. For multiple events (recurrence, death and lost-to-follow-up) the event process is described by finite Markov models. The predictive characteristics are derived in a closed form, thus, Monte Carlo simulation is not required. Some applications to interim prediction of events together with ongoing enrollment in oncology trials are considered.

Note that a couple of plots for real data in [6] show essentially non-linear curves for a long-term interim event prediction. Thus, it would be interesting to understand the nature of a linear trend for the mean number of events in Fig. 1,2 [11].

In addition, it would be useful to distinguish between different types of events in clinical trials as there can be many other events related either to patient reactions, e.g. adverse events, or other types like screen failures, different visits, dropout, consent withdrawn, etc. These events are associated to various trial operational characteristics.

Potentially this area is rather wide and is outside the current review, however as this is organically related to predicting events, it is worth to mention paper [8] where a new methodology for predictive modeling of various operational characteristics based on using so-called evolving hierarchic processes is proposed. This approach combines modeling enrollment and associated events. Some applications to modeling the number of follow-up patients, multiple visits, dropout and other operational events are also considered there.

4 General Discussion

I also would like to contribute to Sec 6 “Discussion” in [11], specifically, to Sec. 6.2. I agree with the authors that if we are talking only about a global prediction, then the global enrollment can be approximated by a Poisson process with some global rate, and this rate is the main factor that drives the future enrollment. Thus, on the global level potentially different models can work well as noted above. The global rate (or mean trend) can be evaluated using some global statistics and should behave similar for different models (during the stable period). The point is mainly about estimating the variation, where Bayesian re-estimation should work better as it uses shrinkage estimators.

However, in general the aim of clinical operation teams is not only to create global predictions but also evaluate predictions on different levels (country, region), predict enrollment performance of centers/countries and provide analysis of different operational characteristics associated with enrollment. For these tasks the models oriented to predicting global enrollment may not be appropriate.

From another side, P-G model [1-5] is created with the purpose to model enrollment in individual centers like starting building blocks using the nature of real processes, and then combine these blocks at different levels and use them to evaluate predictive characteristics. Therefore, this model can be used as the universal and adequate tool for predictive modeling of various enrollment characteristics at different levels. As the next level of hierarchy, it can also be extended to model different events and operational processes on the top of enrollment [8].

I trust these comments and discussion points add potentially useful information for the readers of the interesting review [11] and provide more food to think which models can be used in practical situations.

References

1. Anisimov VV., Fedorov VV. Design of multicentre clinical trials with random enrolment, in book "Advances in Statistical Methods for the Health Sciences. …", Series: Statistics for Industry and Technology, Balakrishnan, N.; Auget, J.-L.; Mesbah, M.; Molenberghs, G. Eds. Birkhäuser, 2006, Ch. 25, 393-406.
2. Anisimov VV, Fedorov VV. Modelling, prediction and adaptive adjustment of recruitment in multicentre trials. Statistics in Medicine 2007; 26: 4958–4975.
3. Anisimov VV. Downing D, Fedorov VV, Recruitment in multicentre trials: prediction and adjustment, In: mODa 8 – Advances in Model-Oriented Design and Analysis, Lopez-Fidalgo J, Rodriguez-Diaz JM, Torsney B (Eds), Physica-Verlag, 2007; 1–8.
4. Anisimov, VV. Predictive modelling of recruitment and drug supply in multicenter clinical trials, Proc. of the Joint Statistical Meeting, Washington, USA, August, 2009, 1248-1259.
5. Anisimov VV. Statistical modeling of clinical trials (recruitment and randomization). *Communications in Statistics — Theory & Methods* 2011; 40:3684–3699.
6. Anisimov VV. Predictive event modelling in multicenter clinical trials with waiting time to response. Pharmaceutical Statistics 2011; 10: 517–522.
7. Anisimov VV. Discussion on the paper 'Prediction of accrual closure date in multi-center clinical trials with discrete-time Poisson process models'. *Pharmaceutical Statistics* 2012; 11: 357–358.
8. Anisimov V. Predictive hierarchic modelling of operational characteristics in clinical trials, *Communications in Statistics - Simulation and Computation* (Dec 2014, online).
9. Bakhshi A, Senn S, Phillips A. Some issues in predicting patient recruitment in multi-centre clinical trials. *Statistics in Medicine,* 2013; 32: 5458–5468.
10. Carter RE, Sonne SC, Brady KT. Practical considerations for estimating clinical trial accrual periods: application to a multi-center effectiveness study. BMC Med. Res. Methodol. 2005: 5:11.
11. Heitjan DF, Ge Z, Ying GS. Real-Time Prediction of Clinical Trial Enrollment and Event Counts: A Review, *Contemporary Clinical Trials* 2015 (online)
12. Deng Y, Zhang X, Long Q. Bayesian modeling and prediction of accrual in multi-regional clinical trials. *Statistical Methods in Medical Research* 2014 (online);
13. Gajewski BJ, Simon SD, Carlson SE. Predicting accrual in clinical trials with Bayesian posterior predictive distributions. Statistics in Medicine 2008; 27: 2328–2340.
14. Lai D, Moy LA, Davis BR, Brown LE, Sacks FM. Brownian motion and long-term clinical trial recruitment. *Journal of Statistical Planning & Inference* 2001; 93: 239–246.
15. Mijoule G, Savy S, Savy N. Models for patients' recruitment in clinical trials and sensitivity analysis. *Statistics in Medicine*, 2012; 31: 1655–1674.
16. Tang G, Kong Y, Chang CCH, Kong L, Costantino JP. Prediction of accrual closure date in multi-center clinical trials with discrete-time Poisson process models. *Pharmaceutical Statistics* 2012; 11: 351–356.
17. Williford WO, Bingham SF, Weiss DG, Collins JF, Rains KT, Krol WF. The "constant intake rate" assumption in interim recruitment goal methodology for multicenter clinical trials. *Journal of Chronic Diseases* 1987; 40: 297–307.
18. Zhang Q, Lai D. Fractional Brownian motion and long-term clinical trial recruitment. *Journal of Statistical Planning & Inference* 2011; 141: 1783–1788.

Vladimir Anisimov
Quintiles & University of Glasgow, UK
E-mail: Vladimir.V.Anisimov@gmail.com