Predicting accrual in clinical trials with Bayesian posterior predictive distributions

Byron J Gajewski, University of Kansas Medical Center
Stephen Simon
Susan Carlson

Available at: http://works.bepress.com/byron_gajewski/14/
Predicting accrual in clinical trials with Bayesian posterior predictive distributions

Byron J. Gajewski1,∗, †, Stephen D. Simon2 and Susan E. Carlson3

1Schools of Allied Health and Nursing, Center for Biostatistics and Advanced Informatics, The University of Kansas Medical Center, Kansas City, KS 66160, U.S.A.
2Office of Medical Research, Children’s Mercy Hospital, Kansas City, MO 64108, U.S.A.
3Department of Dietetics and Nutrition, School of Allied Health, The University of Kansas Medical Center, Kansas City, KS 66160, U.S.A.

SUMMARY

Investigators need good statistical tools for the initial planning and for the ongoing monitoring of clinical trials. In particular, they need to carefully consider the accrual rate—how rapidly patients are being recruited into the clinical trial. A slow accrual decreases the likelihood that the research will provide results at the end of the trial with sufficient precision (or power) to make meaningful scientific inferences. In this paper, we present a method for predicting accrual. Using a Bayesian framework we combine prior information with the information known up to a monitoring point to obtain a prediction. We provide posterior predictive distributions of the accrual. The approach is attractive since it accounts for both parameter and sampling distribution uncertainties. We illustrate the approach using actual accrual data and discuss practical points surrounding the accrual problem. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: prior elicitation; exponential; inverse gamma; Bayesian; sample size

1. INTRODUCTION

The most common reason why clinical trials fail is that they fall well below their goals for patient accrual. According to one source [1], more than 80 per cent of all clinical trials fall short of their accrual goals. The net result is too many studies that appear with inadequate sample sizes and result in confidence intervals that are so wide (or power so low) that they are effectively uninterpretable.

∗Correspondence to: Byron J. Gajewski, Schools of Allied Health and Nursing, Center for Biostatistics and Advanced Informatics, The University of Kansas Medical Center, Mail Stop 4043, 3901 Rainbow Blvd., Kansas City, KS 66160, U.S.A.
†E-mail: bgajewski@kumc.edu

Contract/grant sponsor: United States National Institute of Health; contract/grant number: 1 R01 HD047315-01A2

Received 30 March 2007
Copyright © 2007 John Wiley & Sons, Ltd.
Accepted 1 October 2007
Accrual rates are often developed in an ad hoc fashion. If they are monitored, accrual rates are examined using a purely subjective approach. There are not many good objective quantitative tools for planning and monitoring accrual rates.

With good planning tools, researchers would be able to construct realistic targets for their sample sizes rather than promising a sample size that could not be delivered in a reasonable time frame and within a limited research budget. With good monitoring tools, researchers would get an early warning when accrual rates are suffering. This would allow them to take appropriate corrective action before too much harm was done.

In this paper we propose a method for predicting accrual across a fixed time period and accrual to a target sample size. We will do this through the use of Bayesian prior distributions and Bayesian posterior predictive distributions. Careful and thoughtful elicitation of a prior distribution for accrual rates will force the researcher to confront issues about future expectations regarding patient accrual patterns. Posterior predictive distributions provide an ongoing monitor of the accrual process that accounts for the fact that the accrual process is a random distribution, but also accounts for uncertainty associated with parameters of this distribution.

Before any actual accrual datum is collected, the predictive distribution uses only the prior distribution. As data come in on the actual accrual rates in the trial, the predictive distribution becomes a weighted average of information from the prior distribution and information from the distribution of the observed accrual times. This prevents a researcher from overreacting to a small bit of bad news early in the trial. As more and more data on accrual appear, the weight on the prior distribution decreases while the weight on the actual data increases. This is a feature of Bayesian inference (e.g. [2–9]). If the actual accrual times are very slow, a sufficient number of them will appropriately pull down even the most optimistic of initial projections.

2. METHODOLOGY

2.1. Assumptions

Our primary objective is to provide a method for predicting accrual in a clinical trial by developing a model for the average waiting time between patients. For the purposes of our presentation, we assume that the sample size (n) is fixed after being justified by some statistical approach either Bayesian or frequentist. We also assume that a target value for the average waiting time has been agreed upon by all stakeholders. The direct stakeholders include all investigators and sponsors of the study. The indirect stakeholders might include investigators of competing trials (from an accrual point of view), internal review boards, and data safety monitoring boards (DSMBs).

A model for the waiting time will incorporate both prior expectations for the waiting time and actual waiting times that occur as patients are enrolled in the study. The waiting time is just the difference in dates when patients join the study. Suppose that the study was started on 1st January and the first two subjects were enrolled on 4th and 9th January, respectively. Then the first two waiting times are $w_1 = 3$ and $w_2 = 5$.

Suppose that a stakeholder wishes to assess the accrual process after m patients have been recruited. Let $t_0$ represent the time the study started and $t_1, t_2, \ldots, t_m$ represent the times when each new patient enters the trial. Without loss of generality, we can assume that the study starts at time $t_0 = 0$. Compute waiting times $w_i = t_i - t_{i-1}$. Note that the $w_i$ form a telescoping sum, hence
The goal of the accrual monitoring process is to develop a model for the yet to be observed waiting times \( W_{m+1}, W_{m+2}, \ldots, W_n \).

We will assume that \( w_i | \theta \) is exponentially distributed with mean \( \theta \). The exponential distribution is chosen for its simplicity since it has one parameter. This is convenient since most protocols have prior information on this single parameter. Additionally, we have found that the exponential distribution fits the time gap data we have observed in practice. There is nothing in this approach, however, that bars the use of a more complex waiting time distribution. In the Discussion section we further explore this idea.

A conjugate prior distribution for \( \theta \) is the inverse gamma distribution. We will assume that \( \theta \sim IG(k, V) \). The expected value and standard deviation for the inverse gamma are \( V/(k-1) \) and \( V/((k-1)\sqrt{(k-2)}) \), respectively. It may be useful to note that the coefficient of variation (CV) for the inverse gamma distribution is \( 1/\sqrt{(k-2)} \). These moments can be useful for the process of eliciting a prior distribution. To fully specify the prior distribution, we need to specify at least two moments of the inverse gamma distribution.

Assuming the sample size is established (say with power or precision arguments), we will ask the investigator two questions:

1. How long will it take to accrue \( n \) subjects?
2. On a scale of 1–10, how confident are in your answer to (1)?

Let \( T \) represent the answer to question (1) and \( P \) represent the answer to question (2) divided by 10. We will set the prior distribution for \( \theta \sim IG(nP, TP) \). The value \( nP \) in the inverse gamma distribution represents the ‘prior sample size’. If a researcher boldly selects \( P = 1 \), then the prior sample size is the target sample size. At the other extreme, the prior sample size is one-tenth the target sample size. Although this framework does not allow for ‘flat’ priors, a value of \( P = 0 \) would effectively produce a flat prior. The second parameter, \( TP \), scales the response so that the expected waiting time of \( V/(k-1) = TP/(nP-1) \approx T/n \). The standard deviation of the waiting time is \( V/((k-1)\sqrt{(k-2)}) \approx (T/n)/\sqrt{nP} \), which is the expected accrual rate divided by the square root of the prior sample size. From an exponential distribution point of view, this is the standard error for an average of \( nP \) exponential random variables, each with mean \( T/n \).

Recent literature on the elicitation of prior distributions suggests that the investigator’s expected prior accrual rate \( T/n \) should be reliable since means and medians can be specified by expert opinion fairly accurately [2, 3]. This is convenient for design purposes since most protocols report \( T \) and \( n \). What might not be readily available is the answer to question (2) (\( P \)). The investigator’s opinion needs to be elicited.

However, in some instances, the investigator can look back at historical data from a similar population. If this historical trial has a small to moderate sample size (smaller than \( n \)), then the prior can be weighed equal to the size of the previous trial. This is reasonable because the sample size from the new study will dominate the predictions fairly quickly because the investigator will eventually observe more data in the new protocol. In the presence of historical data on the sample size is bigger than or equal to \( n \), we wish to weigh this prior information less so as to not overwhelm (or swamp) the posterior distribution with data that are not directly related to the current trial. In the latter case, we propose weighing the prior sample size by \( \frac{1}{2} \). This allows the new data to carry at least equal weight to the prior around halfway through the new trial’s accrual. Other investigators with other trials might want to weigh the prior data a little differently. Or better, they could look at alternative choices for the weight.
Once a researcher specifies a value for \( P \), we can compute the CV of the prior distribution and use that to cross-check the original selection. If the research believes that the CV for the prior distribution is unreasonably large, a stakeholder can then increase the level of confidence to produce a smaller and more realistic CV. If the researcher believes that the CV is unreasonably small, a stakeholder can make a similar adjustment in the level of confidence in the opposite direction. This process of cross-checking the prior distribution by offering feedback about the prior on a different scale will help to insure the validity of the elicitation of prior information.

2.2. Algorithm

The algorithm for predicting the end of a trial will incorporate the uncertainty in the parameter estimate as well as the uncertainty in the random process of time dates. To obtain these estimates we will use simulation. First, we will obtain a simulated draw from a posterior distribution of \( \theta \) and then, using the \( \theta \) we drew, we will obtain a draw from the posterior predictive distribution, \( W \). Note that the posterior distribution is

\[
\theta | w \sim \text{IG}(nP + m, TP + tm)
\]

The posterior mean is

\[
E[\theta | w] = \frac{TP + tm}{nP + m - 1} \approx \left( \frac{nP}{nP + m} \right) \frac{T}{n} + \left( \frac{m}{nP + m} \right) \frac{tm}{m}
\]

which is the weighted average of the prior mean (\( T/n \)) and the mean of the observed data (\( tm/m \)). The weights are proportional to the prior and observed sample sizes, respectively.

Another interesting fact is that the predictive distribution of an unobserved \( W_{m+1} \) can be calculated in closed form. To calculate, we need to compute a weighted average of exponential distributions with weights determined by the posterior distribution. This is accomplished through integration. Let \( g(W) \) be the predictive distribution, then

\[
g(W) = \int_{0}^{\infty} \frac{1}{\theta} \exp(-W/\theta) V^k \theta^{-(k+1)} \exp(-V/\theta) / \Gamma(k) d\theta
\]

The density is a location-shifted version of the Pareto distribution. For large values of \( k \), it is not far from an exponential distribution. The expected value and variance are \( E(W) = V/(k+1) \) for \( k>1 \) and \( \text{var}(W) = E(W)^2 k/(k-2) \) for \( k>2 \). The CV of this distribution, \( \sqrt{k/(k-2)} \), is always larger than 1, but approaches 1 as \( k \) increases. The CV for an exponential distribution is exactly 1. Therefore, this formula tells us that uncertainty in the parameter \( \theta \) leads to an increase in variation of the waiting times relative to an exponential distribution.

We are interested in predicting the next \( n - m \) data \( W_{m+1}, \ldots, W_n \). While in the exponential case direct simulation from the posterior distribution is also possible, a simple and easily generalizable approach for each \( W \) is to randomly select \( \theta_1 \) from the posterior distribution and then randomly select waiting time \( n - m \) random variables from \( W_{m+1,1}, \ldots, W_{n,1} \) from an exponential distribution with parameter \( \theta_1 \). We use a simulation approach because it can be easily extended to other distributions. Repeat this process for \( \theta_2, \theta_3, \ldots, \theta_b \), where \( b \) is a large number (typically 1000). The sum of observed and simulated waiting times, \( S_b(n) = w_1 + w_2 + \cdots + w_m + W_{m+1,b} + \cdots + W_{n,b} \) will represent \( b \) estimates of the total duration of the clinical trial of size \( n \).
We can also use this process to obtain a posterior predictive sample size. Let $T$ represent the time point at which the study must end. Compute partial sums $S_b(m+1), S_b(m+2), \ldots$ until the partial sum exceeds $T$. The values $n_b^P$ that represent the largest values where the partial sums do not exceed $T$, provide a predictive distribution of sample sizes.

This simulation approach also allows one to examine more complex accrual patterns, such as a study where the goal is to accrue 50 patients, or until 6 months have elapsed, whichever one comes first.

Once the model is accepted and the simulation is complete, histograms of the posterior distributions can be viewed in order to use the full strength of the Bayesian approach. These posterior distributions can be summarized using 2.5, 50, and 97.5 percentiles. Note that these posterior predictions will be fixed if the trial is stopped or if accrual achieves $n$.

3. EXAMPLE

To illustrate the proposed method, consider a current phase III clinical trial (randomized, double-blind, and placebo-controlled) used to examine the efficacy and safety of 600mg/day of docosahexaenoic acid (DHA) for the last two trimesters of pregnancy. Kansas University DHA Outcome Study (KUDOS) was planned and accrual started prior to our development of these methods, but still serves to illustrate how this approach would work. This study is an extension of a previous clinical trial. The current protocol requires $n = 350$ subjects, with balanced randomization to either treatment or placebo control. Based on information from a previous study, investigators feel confident in being able to recruit 350 subjects across 3 years.

At the design phase of the study, it was felt that the previous clinical trial offered strong prior information. Setting $P = 0.5$ results in $k = 175$ and $V = 1.5$ years. This corresponds to a prior mean of $V/(k-1) = 0.0086$ years or 3.1 days.

After the study was funded and the protocol approved, the investigative team began recruiting subjects. After 239 days the project director compiled a report that displays enrolled dates of 41 subjects. This represents an average waiting time of 5.8 days, much longer than expected.

Using this information, we calculated predictive distributions under various conditions:

(A) Only the information at the beginning of the study (informative prior);
(B) non-informative prior ($P = 0$) with the observed data; and
(C) informative prior discussed above with the observed data.

<table>
<thead>
<tr>
<th>Situation</th>
<th>Parameter</th>
<th>2.5 Percentile</th>
<th>50.0 Percentile</th>
<th>97.5 Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Prior</td>
<td>Waiting time (days)</td>
<td>2.7</td>
<td>3.1</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>Total trial duration (years)</td>
<td>2.5</td>
<td>3.0</td>
<td>3.7</td>
</tr>
<tr>
<td>(B) Posterior</td>
<td>Waiting time (days)</td>
<td>4.4</td>
<td>5.9</td>
<td>8.1</td>
</tr>
<tr>
<td>(non-informative)</td>
<td>Total trial duration (years)</td>
<td>4.3</td>
<td>5.6</td>
<td>7.6</td>
</tr>
<tr>
<td>(C) Posterior</td>
<td>Waiting time (days)</td>
<td>3.2</td>
<td>3.6</td>
<td>4.2</td>
</tr>
<tr>
<td>(informative)</td>
<td>Total trial duration (years)</td>
<td>3.3</td>
<td>3.7</td>
<td>4.4</td>
</tr>
</tbody>
</table>

Table I. Posterior distribution of waiting times and predicted total trial duration.

Copyright © 2007 John Wiley & Sons, Ltd.

The parameter estimates are summarized in Table I. The median predicted trial duration is 3.0 years using the informative prior at the beginning of the study, 5.6 years using a non-informative prior with the observed data, and 3.7 years using an informative prior with the observed data.

The posterior $\theta|\mathbf{w}$ under different conditions shows that the smallest rate (fastest trial accrual) is associated with the informative prior distribution at the beginning of the study and the largest
rate (slowest trial accrual) is associated with the observed data and a non-informative prior. The informative posterior distribution indicates a compromise of the prior and the data.

Using these posterior distributions from simulation, we were able to graphically present the probable outcomes of the study (Figure 1). Note that the greatest degree of uncertainty occurs with the non-informative prior, even more than the informative prior at the start of the study.
Figure 3. Prediction distributions (based on 1000 simulations) for minimum detectable difference (0.80 power) using distribution of accrued patients for an informative prior at the start of the study (top left), for a non-informative prior after observing 41 patients through the first 239 days of the study (top right), and for an informative prior after observing 41 patients through the first 239 days of the study (bottom left). We assume a two-sided two-sample \( t \)-test (\( \alpha = 0.05 \)) with alternative hypothesis effect size having 0.80 power at \( n = 350 \).

This reflects the fact that the prior sample size of 175 provides greater precision than 41 patients observed. Also note that the median estimated total trial duration under the non-informative prior is roughly the same as a proportional extrapolation of the 41 patients (5.6 = \( \frac{350}{41} \) * \( \frac{239}{365} \)).

The prior information suggests that the current observed accrual rate (41 subjects in 239 days) was slower than the prior accrual rate. However, we are still early in the current accrual process. Therefore, in actuality, at this point we will utilize the informative prediction. This is a compromise of what we observed and what the prior belief is. Of course, this means that we believe that the deviation in the observed accrual rate is due, at least in part, to sampling variability.

The biggest benefit of the entire method is that our prediction has uncertainty. Since we have yet to observe all of the data, this uncertainty mimics reality. Incorporating uncertainty into the prediction is a big improvement over traditional approaches.

If the original plan of the research study was to collect patients until a fixed point in time, you can predict the expected number of patients that you would expect to accrue by the end of the trial. Figure 2 shows the predicted number of patients if the trial were required to be complete in 3 years. The graph on the top left shows the number of accrued patients expected at the start of the study using the informative prior. The median prediction is 350 patients, of course, but
also of interest is that the central 95 per cent probability of the predicted number of patients is 290–421. This information should be provided to the researcher during the elicitation of the prior distribution to provide feedback about the implications of the certainty for a given value of \( P \).

The top right plot shows the predicted number of patients for a 3-year trial using a non-informative prior after 239 days and 41 patients have been observed. The median prediction is 187, which is a simple linear extrapolation from current time point to 3 years. The 95 per cent central probability is 142–242 patients, well short of the desired 350 patients.

The plot on the bottom left shows the predicted number of patients for a 3-year trial using an informative prior. The median prediction is 275 patients and the central 95 per cent probability is 235–318 patients. This estimate is not as pessimistic as the estimate under the non-informative prior because the prior distribution still carries considerable weight in this analysis.

The original protocol was designed using a frequentist power calculation. Accrual decisions might be improved with the help of this power calculation. This produces an example of a hybrid of Bayesian and frequentist methods. The predicted number of patients and the power of the study can result in a posterior predictive minimum detectable difference for 80 per cent power. In our example, the primary outcome is assessed for statistical significance using a two-sided two-sample \( t \)-test at \( \alpha = 0.05 \). In the protocol, the assumed standardized effect size in the alternative hypothesis resulted in a power \((1 - \beta)\) of 0.80 when \( n = 350 \). But as shown in Figure 3, the minimum detectable difference to achieve 80 per cent power of the study actually varies at each of the iterations of the simulation. We will call this the posterior predictive minimum detectable difference. The non-informative posterior indicates a higher minimum detectable difference than the prior prediction. The informative posterior indicates a compromise in minimum detectable difference. If the minimum detectable difference is unrealistically large, and unlikely to be achievable with the treatment being studied, then the researcher should revise the experiment or stop early for futility.

4. DISCUSSION

As far as we know, no statistical models for predicting accrual rates in a clinical trial have been developed. Some investigators monitor accrual very carefully and some are more lax. However, even among those who monitor carefully, decisions about changing the protocol to accelerate recruitment are made on an \textit{ad hoc} basis without any quantitative rigor. Current practice relies on assuming that observed accrual rates have no uncertainty. We demonstrated a method for predicting accrual that accounts for the uncertainty in the model parameter as well as the uncertainty in the accrual process. The parameter uncertainty applies to the parameter \( \theta \) that is modeled using an inverse gamma distribution. The inverse gamma distribution is specified using a combination of prior knowledge and observed accrual gap data. We model the sampling distribution uncertainty (accrual process) with an exponential distribution.

Although Bayesian data analysis methods are increasing in popularity [4], a commonly cited complaint about this approach is the requirement that the data analyst must specify prior distributions. A prior distribution reflects current knowledge in the form of a statistical model. In our case the prior distribution reflects what is currently known about the anticipated accrual rate of a clinical trial. To obtain a valid prior distribution one needs knowledge about the projected accrual rate as well as the uncertainty of the accrual rate.
Critics of the Bayesian approach believe that researchers should take a disinterested and objective perspective on their research. While this point certainly merits extensive debate (e.g. [10, 11]), we come down on the side of the Bayesian statistics. From the perspective of accrual rates, however, there should be no debate. A researcher would never start a clinical trial if he/she did not have at least an inkling of how many patients are likely to be recruited during the course of the study. The accrual rates are reported in almost all protocols.

Identifying parameters for a prior distribution can involve elicitation from the investigator (expert). Eliciting expert opinion is itself a science that uses tools from statistics and psychology to transform opinion to a probability distribution. The literature on eliciting expert opinion is vast. Fortunately, O’Hagan et al. [2] have summarized key issues in the elicitation of expert opinion. In addition, the literature on prior distributions in clinical trials is large. For the sample size of phase II clinical trials, Tan and Machin [5] offer designs based on non-informative priors and Mayo and Gajewski [6, 7] have an approach based on informative priors. Spiegelhalter et al. [4] offer practical guidance on the use of expert opinion in clinical trials. They suggest pessimistic and optimistic priors as alternative analyses on the same data.

The trick is not getting researchers to offer an opinion about accrual rates, but instead getting them to recognize the uncertainty inherent in the accrual rates and quantify it in terms of a prior distribution. Eliciting a prior distribution in general is a difficult task. Most researchers tend to understate the true amount of uncertainty when you ask them about prior distributions [2].

There are at least two possible sources for eliciting a prior distribution for accrual rates. Researchers might rely on accrual rates from previous clinical trials. This is a reasonable approach as long as the proposed trial has similar accrual properties as the historical trial. When the current trial has different properties than the past trials we must elicit a prior distribution from expert opinion. A common statement in the elicitation literature is that psychometric methods are needed to design a valid and reliable tool for eliciting prior distributions from clinical researchers [2].

After specifying an informative prior, one can use our method to generate a predictive distribution of the sample size in a specified time period. This process can be done before any accrual data are collected. The predictive distribution accounts for both parameter and sampling distribution uncertainties—typically investigators do not account for either in planning a study. Accounting for this uncertainty in the planning stages might encourage investigators to allow more time for the recruitment of subjects for their study and may encourage more study centers or study sites to achieve their accrual goals.

The prediction algorithm may be sensitive to departures of the exponential distribution. To assess for model adequacy, we suggest the use of a Bayesian chi-square test for goodness of fit [8]. This chi-square test is conceptually the same as the classic version. But rather than computing a single \( p \)-value, this method provides a posterior distribution of the \( p \)-values because we are using the posterior distribution of the model parameter from the exponential. Computationally, this involves the calculation of a \( p \)-value at each of the iterations of the simulation. The Bayesian chi-square test will detect lack of fit from specifying an incorrect distribution, but also will detect lack of fit when the parameters of the prior distribution are inconsistent with the actual accrual data. In our case we had adequate fit since the posterior probability \( p \)-value for non-informative fit between 0.025 and 0.975 was 0.970. However, the informative fit’s probability was 0.0110, indicating that the current protocol may differ from the historical information.

This inconsistency could result from the prior distribution being flat out wrong at the start of the study (a combination of an incorrect mean and too great a degree of certainty about this mean).
Or perhaps, the actual accrual pattern could shift during the study. A change in patient profile could limit the number of eligible patients. A reduction in research staffing could limit the ability of the investigator to approach all eligible subjects. Patients may develop a preference for one arm of the study over another over time and may not wish to undergo randomization. The investigator or a DSMB may change the inclusion/exclusion criteria on the basis of emerging safety issues.

As a result of these issues, there are three major practical implication points that we would like to make.

**Point 1:** There is a learning curve. It is possible that the exponential distribution could be inadequate. Other obvious choices for alternative models include gamma, log-normal, or Weibull [9]. Conceptually, it is more difficult to elicit prior distributions from these parameters and beginning with an exponential model is a good strategy in the beginning (learning curve). However, after the investigators have observed a good amount of data (say \( n > 30 \)) and the Bayesian chi-square test for goodness of fit is inadequate for an exponential (or say a change point exponential model), it seems reasonable to try other distributions. This implies a posterior prediction using a non-informative prior. It seems to us that before the trial starts and early in a trial (or throughout a small trial), the exponential gamma distribution (our method) should be used.

One needs to be cautious of a non-significant \( p \)-value with a small sample size. However, as a matter of philosophy, we would not change to a more complex model beyond the exponential without strong prior information or strong observed values. A reviewer suggested that an alternative to the \( p \)-value might be to compare the actual accrual to date with the predicted accrual in an observational or graphical manner, so that departures can be observed.

We showed two examples of the use of a predictive distribution: predicting the amount of time that would take to accumulate 350 patients and predicting the number of patients that would accrue in trial that must end after 3 years. Although we do not show an example here, the predictive distribution can easily be extended to a composite completion criteria, such as 350 patients or 3 years, whichever comes first. Two complementary probabilities of interest with such a composite completion criterion would be the probability that the study finishes early with 350 patients and the probability that the study finishes at 3 years with fewer than 350 patients.

We have not yet investigated using the method in the case of adaptive designs (i.e. [12, 13]). For adaptive trials, an interim analysis is planned. First, suppose accrual interim analyses are planned during several fixed time points, then our method can be applied in stages: beginning of the study and interim analysis 1, interim analyses 1–3, and all the way to the final analysis. The accrual information in the previous stage can be used as the prior information in the upcoming stage. And with this raw gap data, one can use a much more flexible set of distributions.

Our paper suggests a simple model for accrual but it needs careful consideration in practice. The model might best be applied in pieces and adapted depending on strategy changes in accrual. In addition, one may want to use this information and incorporate in a subjective way into the prediction. This subjectivity might not necessarily be supported by the observed data. The utility of using pieces of an exponential model make elicitation of subjective opinion much easier.

**Point 2:** Predicting accrual can involve complex modeling. Accrual interim analyses might make investigators realize that the proposed exponential model is unrealistic. The interim analyses might suggest that the mean time between patients is not a fixed constant (as suggested by the exponential model), as it might vary considerably throughout the life of the trial, as recruitment decisions are made. For example, a reviewer stated that slow start ups are typical as the trial gets underway, and strategies may change if recruitment is slow, such as increasing payments to recruiting physicians, adding more centers, increasing efforts within each center to encourage recruitment, or the trial
may ‘self-market’ as subjects enjoy the benefits of participation. While the exponential model can be useful for approximately predicting future accrual and accounting for the uncertainty in these projections, it is important that investigators think deeply about why accrual is or is not adequate, and take further action, which then will typically invalidate predictions made by a simple model.

Another possibility is to model in the following way. Suppose researchers increase enrollment with new centers that come along. Then we can add hyperparameters with an exposure parameter to the centers and model hierarchically. This will allow us to do prediction equations and possibly drop different centers. Another possibility is that suppose that the protocol changes, and incentives are put in or a new inclusion/exclusion criterion is used. Then we might need to model this heterogeneity with an updated prior distribution for each of the centers.

**Point 3: Predicting accrual can assist in decision making.** In this paper we suggested using posterior predictive distributions of the minimum differences to assist in the decision-making process. This is not a full interim analysis because we do not unlock the outcome data and randomization code. A full interim analysis would include more than accrual updates. It would include peeks at the primary outcome and require one to unlock the randomization. One could define a fully Bayesian model of the outcome data and use the principles described in this paper to predict what the data will look like using both the prior opinion about the effect and the current data observed up to this point [13]. One could obtain a more informed distribution of the posterior predictive data using Bayesian tools for accrual simultaneously with outcome data to provide frequentist-operating characteristics. However, since one is peeking at the data, one would have to adjust the level as a penalty.

## 5. CONCLUSION

Predicting accrual across a fixed time period of a planned study is not an easy task. As demonstrated in this paper there are uncertainties that should be incorporated into this prediction. We hope that the method in this paper encourages investigators to account for these uncertainties when planning and monitoring accrual in a clinical trial. Readers may obtain R code from the authors.

### ACKNOWLEDGEMENTS

The work of B. J. G. and S. E. C. was partially funded by the United States National Institute of Health (1 R01 HD047315-01A2). We thank the editor, anonymous AE, and two anonymous reviewers for comments that helped in improving this manuscript.

### REFERENCES