Slipped deadlines, sample size shortfalls, and a proposed Bayesian remedy using an informative prior distribution

Stephen Simon1,2, Byron Gajewski3, Judy Champion4, Vidya Sharma4, Susan Carlson3

1P.Mean Consulting, 2University of Missouri-Kansas City, 3University of Kansas Medical Center, 4Children’s Mercy Hospitals and Clinics

Background: The most common reason why clinical trials fail is that they fall well below their goals for patient accrual. Researchers will frequently overpromise and underdeliver on the number of patients they can recruit during the proposed time frame. The result is studies that take far longer than planned and/or that end with fewer patients than planned. This raises serious economic and ethical issues.

Our research efforts have focused on (1) obtaining reliable data on the scope and magnitude of problems with slow patient accrual in clinical trials, and (2) developing a Bayesian model for accrual that will encourage careful planning of accrual rates as well as allow regular monitoring of accrual patterns during the conduct of the clinical trial.

Methods: A random sample of 130 prospective studies approved by the Children’s Mercy Hospital (CMH) IRB and completed between January 2001 through December 2005. A study was excluded if
• it was a retrospective chart review;
• it received expedited review or was declared to be exempt from IRB review;
• it was part of the Children’s Oncology Group (COG).

Retrospective studies, of course, could not have issues with accruing patients. Expedited and exempt studies were almost always retrospective. Studies in the Children’s Oncology Group were almost always studying rare conditions across multiple centers and were expected to accrue only one or two patients per site, at best.

For each calendar year, a randomized list of all studies submitting a final report was created. The protocols, interim reports, and final reports of the first 25 studies on this list were reviewed. If a study was found to be ineligible after review, then the next study on the randomized list was selected until 25 eligible studies per year were found. Any study receiving internal funding (Katherine B. Richardson foundation grants) was automatically included in the sample. The total sample size was 130 studies.

For each study selected, the following information was collected, if available.
• Proposed start date
• Proposed end date
• Actual start date
• Actual end date
• Proposed sample size
• Actual sample size

Based on these data, we calculated, if the appropriate data was available, the following statistics:
• the proportion of studies whose actual duration was longer than the proposed duration;
• the average increase in duration of these studies;
• the proportion of studies that collected a smaller sample size than proposed; and
• the average shortfall in sample size among these studies.

At the same time, a Bayesian model for accrual was applied to a clinical trial at Kansas University Medical Center to produce monthly reports projecting estimated final sample sizes with uncertainty limits given the initial projection and currently available enrolment data.

The Bayesian model required the researcher to provide an informative prior distribution on the accrual rate. The initial prior distribution was elicited by getting responses to two questions:
1. How long will it take to accrue n subjects?
2. On a scale of 1-10, how confident are you of this answer?

As data on accrual is collected during the course of the study, this data can be combined with the prior distribution to produce a revised estimate of the time duration for a given target sample size or the total sample size for a given target deadline.

These two responses were combined to provide an initial prior distribution for the accrual rate (Gajewski and Simon 2008). A similar approach appears in Anisimov and Fedorov (2007). This prior is then checked for reasonableness by reviewing uncertainty limits for the duration of the trial or for the total sample size. Adjustments can then be made if the prior distribution implies too much certainty about the final outcome or too little uncertainty.

Results: Documentation on proposed and actual time frames was not reported in most studies. Only 11 studies (8%) submitted to the IRB had sufficient information to allow calculation of whether the actual duration of the study was longer or shorter than the proposed duration (Table 1). Of these 11 studies, 8 took more time than proposed and the average increase in duration in these 8 studies was 100%.

There were 109 studies that included both sufficient information to estimate sample size shortfalls (Table 1). Most failures involved no documentation of the final sample size, but there were some ambiguous proposed sample sizes (e.g., up to 40).

Among the studies with complete sample size information, 59 (54%) fell short of the proposed sample size. The average shortfall across these 59 studies was 55%.

The informative prior distribution selected for the Bayesian model was consistent with accrual patterns observed during the study. This prevented the researchers from overreacting to a slow initial start to patient accrual. The projected sample sizes produced by the Bayesian model (Figure 1, solid line) were on target throughout most of the course of the study and the uncertainty limits quickly converged (dotted lines).

In contrast, the projections produced without the informative prior (Figure 2, solid line) placed too much weight on the limited accrual patterns observed in the first year of the study and produced projections that were smaller than what was actually observed.

The Bayesian model, however, also has the capability of correcting an inaccurate prior distribution as the accumulated accrual patterns provide contradictory results. Informative priors provide an ideal way to incorporate prior knowledge and experience into the monitoring of accrual patterns.

Additional applications of the Bayesian model include projections of the time duration of a clinical trial until a target sample size is met and projections of the number of patients needed to screen in order to reach a target number of patients who are eligible and who agree to participate in the clinical trial (Simon 2007).

Table 1. Studies with incomplete information.

<table>
<thead>
<tr>
<th>Total</th>
<th>130</th>
</tr>
</thead>
<tbody>
<tr>
<td>No proposed start date</td>
<td>- 2</td>
</tr>
<tr>
<td>No proposed end date</td>
<td>- 25</td>
</tr>
<tr>
<td>No proposed start/end date</td>
<td>- 90</td>
</tr>
<tr>
<td>Proposed start/end date</td>
<td>13</td>
</tr>
<tr>
<td>No actual start date</td>
<td>- 1</td>
</tr>
<tr>
<td>No actual end date</td>
<td>-</td>
</tr>
<tr>
<td>Complete data</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 2. Complete sample size information.

<table>
<thead>
<tr>
<th>Total</th>
<th>130</th>
</tr>
</thead>
<tbody>
<tr>
<td>No proposed sample size</td>
<td>- 2</td>
</tr>
<tr>
<td>Ambiguous proposed sample size</td>
<td>- 2</td>
</tr>
<tr>
<td>No actual sample size</td>
<td>- 17</td>
</tr>
<tr>
<td>Complete data</td>
<td>109</td>
</tr>
</tbody>
</table>

Conclusions: A large number of research studies failed to meet the specified sample sizes and the average shortfall among these studies was considerable. The Bayesian model for accrual produced useful reports for a particular study and provided reassurance to the researchers that their accrual rates were on target.

Future work: We hope to partner with other organizations that conduct large numbers of clinical trials to get more data on existing problems with slipped deadlines and sample size shortfalls and to test the Bayesian accrual model on a wide range of clinical trials. We also plan to extend the Bayesian model to more complex settings (e.g., multicenter trials).

Bibliography


Figure 1. Projected final sample size (vertical axis) over time in years (horizontal axis) as information about actual accrual patterns is observed. This projection uses an informative prior.

Figure 2. Projected final sample size (vertical axis) over time in years (horizontal axis) as information about actual accrual patterns is observed, without an informative prior.