A Bayesian approach in design and analysis of pediatric cancer clinical trials

Jingjing Ye PhD1 | Gregory Reaman MD2 | R. Angelo De Claro MD3 | Rajeshwari Sridhara PhD1

1Division of Biometrics V, Office of Biostatistics, Office of Translational Sciences, Center of Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland
2Oncology Center of Excellence, U.S. Food and Drug Administration, Silver Spring, Maryland
3Division of Hematologic Malignancies I (DHM1), Office of Oncologic Diseases, Center of Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland

Correspondence
Jingjing Ye, Division of Biometrics V, Office of Biostatistics, Office of Translational Sciences, Center of Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD. Email: jingjingye@gmail.com

Summary
It is well recognized that cancer drug development for children and adolescents has many challenges, from biological and societal to economic. Pediatric cancer consists of a diverse group of rare diseases, and the relatively small population of children with multiple, disparate tumor types across various age groups presents a significant challenge for drug development programs as compared to oncology drug development programs for adults. Due to the different types of cancers, limited opportunities exist for extrapolation of efficacy from adult cancer indications to children. Thus, innovative study designs including Bayesian statistical approaches should be considered. A Bayesian approach can be a flexible tool to formally leverage prior knowledge of adult or external controls in pediatric cancer trials. In this article, we provide in a case example of how Bayesian approaches can be used to design, monitor, and analyze pediatric trials. Particularly, Bayesian sequential monitoring can be useful to monitor pediatric trial results as data accumulate. In addition, designing a pediatric trial with both skeptical and enthusiastic priors with Bayesian sequential monitoring can be an efficient mechanism for early trial cessation for both efficacy and futility. The interpretation of efficacy using a Bayesian approach is based on posterior probability and is intuitive and interpretable for patients, parents and prescribers given limited data.

KEYWORDS
Bayesian operating characteristics, Bayesian sequential monitoring, incorporating prior knowledge, pediatric oncology/hematology, registrational cancer trial

1 | INTRODUCTION

A 1968 editorial article in the Journal of Pediatrics referred to children as “therapeutic orphans”1 to express the frustration of many clinicians over the lack of pediatric prescribing information for approved drugs.2 The Pediatric Research Equity Act (PREA),3 which provides the requirement for studies, and the Best Pharmaceuticals for Children Act (BPCA),4 which provides the incentive of additional exclusivity for products of sponsors who conduct requested studies in the pediatric population, were enacted in 2003 and 2002, respectively, following the initial
legislative provision for exclusivity in 1997 to correct this serious deficiency in drug development for young patients.

Until recently, the requirement for pediatric evaluation of most oncology products for adult cancers was waived, because children typically do not have adult-type cancers (eg, lung cancer) or because the indication or drug had been granted orphan designation. PREA, therefore, has had no impact on pediatric anticancer drug development. Pediatric studies for labeling updates are largely done through BPCA by the fulfillment of a Written Request (WR), issued by the FDA. The WR asks sponsors to voluntarily submit data from specified studies to determine whether a drug provides meaningful health benefits to the pediatric population. On 18 August 2017, Congress amended PREA to include provisions of the Research to Accelerate Cure and Equity (RACE) for Children Act to require early evaluation of oncology products in children if the molecular target of the drug is considered relevant to childhood cancer. In addition, the Food and Drug Administration Reauthorization Act of 2017 (FDARA) eliminates the exemption for required pediatric studies based on orphan designation when the molecular target of the drug is relevant to one or more pediatric cancers.

The paradigm for cancer drug development now focuses on molecularly targeted agents and even utilizes histology agnostic approaches. This shift to requirements based on molecular mechanism of action results in even smaller pediatric subpopulations and requires more complex trial designs in terms of enrichment procedures for dose-finding and pharmacokinetic studies, safety monitoring, meaningful endpoint selections, and recruitment strategies. Given the challenges in pediatric cancer, new provisions to require pediatric studies and the small patient numbers, consideration of innovative study designs including Bayesian statistical approaches is required. In this article, we use a case study to illustrate how Bayesian design and analysis can be applied to pediatric registration trial setting to help address the challenges.

2 | REGISTRATIONAL TRIAL CONSIDERATIONS

Because cancers in the pediatric and adult populations generally do not share the same biology, natural history and disease progression, full extrapolation is unlikely, and requirements for the pediatric studies can vary greatly according to the disease indications. Pediatric registration trials frequently fall into two scenarios: partial extrapolation or full efficacy and safety evaluations for a pediatric specific indication.

Partial extrapolation is considered when adult data are available and are thought to be relevant to pediatric patients. The determination of partial extrapolation is based on the decision tree, illustrated in FDA guidance.

For pediatric-specific indications, diseases that occur predominantly in children, for example, Wilms’ tumor and neuroblastoma complete efficacy and safety data in children are required, whereas cancers that have relatively high prevalence in adults and pediatrics, for example, acute lymphoblastic leukemia (ALL) may possibly lend themselves to extrapolation approaches.

For either partial extrapolation or full efficacy and safety studies for pediatric specific indications, uniform international master protocol for biomarker-directed studies, platform trial, and common control would be encouraged to facilitate the pediatric development including possible registration strategies as much as possible. Because of the limited patient population and challenges discussed earlier, these designs provide efficiency and avoid duplication and competition.

In addition, limited opportunities for extrapolation of efficacy and the small patient numbers require consideration of innovative study designs including Bayesian statistical approaches to incorporate either prior knowledge from adult or external and historical controls.

3 | BAYESIAN APPROACHES

In general, Bayesian approaches can account for uncertainty in prior knowledge and formally incorporate prior knowledge either from adult studies or other sources. Decision making using a Bayesian approach is based on posterior probability of efficacy, and therefore provides direct measure of evidence on a clinical scale. When incorporating prior information, if unsure of the relevance of prior information, a probability of relevance can be incorporated as part of the prior distribution. However, even if prior information seems very relevant, sufficient skepticism about potential efficacy exists, therefore requiring that prior information be discounted.
For a pediatric cancer registration trial, Bayesian analysis can be a flexible tool for either partial extrapolation or pediatric specific indication. For the case of partial extrapolation, because prior knowledge on adult trials is available, Bayesian analysis can incorporate prior adult information in the pediatric setting. If pediatric older age group information, for example, adolescents, is available, in the case when deemed biologically reasonable, both adult and older age group information can be incorporated.

For pediatric-specific indications, Bayesian analysis can incorporate prior information from older age group or control. The control can be an external, historical control.

3.1 | Case study

We will illustrate the possible use of Bayesian analysis using a case study. Nilotinib was approved by the US Food and Drug Administration (FDA) in 2018 to treat pediatric patients greater than or equal to 1 year of age with newly diagnosed Philadelphia chromosome positive chronic myeloid leukemia (Ph + CML) in chronic phase. The approval decision was not based on a Bayesian approach. By reanalyzing the data, however, we propose it as a case example to illustrate how a Bayesian approach can be useful. Table 1 summarizes the clinical trial designs and results of the primary endpoint major molecular response (MMR) at 12 months used in the approval of Nilotinib for the treatment of newly diagnosed Ph + CML in chronic phase. The approval in adults was based on a randomized clinical trial, where patients were randomized 1:1 to treat with either Nilotinib or imatinib with sample size of 282 and 283, respectively. The pediatric approval was based on a single-arm study with 25 patients. The 12-month MMR results are summarized: 44%, 95% confidence interval (CI) 38.4% to 50.3% in adults and 60%, 95% CI 38.7% to 78.9% in children were observed. The full prescriber information is available in https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/022068s027lbl.pdf.

Here, we considered this case study to use partial extrapolation of available adult study data to pediatric studies.

3.2 | Construction of priors

In the Bayesian framework, prior distribution refers to distribution of our prior expectation and the posterior distribution is our updated estimate as data are prospectively accumulated.

For pediatric cancer registration studies, the pediatric development program historically commences after the adult studies have been completed, and the drug in question has been or is in the process of being approved for that indication. Assuming borrowing the adult data is possible, the applicability of the adult data to the pediatric setting must be discounted in order to account for the difference between adults and pediatric patients. Some common methods for discounting are weighted average of the means for the randomized and historical controls to control bias,11 meta-analytic predictive approaches,12,13 the power prior approach14 where the prior distribution is taken to some fraction power, $(f(\theta))^a$ to flatten the distribution, the commensurate prior approach,15 and test and pool method.16 Lim et al (2018)17 provided review and summary of available Bayesian and frequentist methods for the use of the historical controls.

In this article, we will use the prior discounting method proposed by Greenhouse and Waserman, (1995).18 The prior for the pediatric study is formulated as follows:

$$\text{Prior} = (1-a) \cdot f(D) + a \cdot g(D),$$

\[ \text{TABLE 1} \quad \text{Comparison of adult and pediatric approval of Nilotinib} \]

<table>
<thead>
<tr>
<th></th>
<th>Adult</th>
<th>Pediatric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design</td>
<td>Randomized</td>
<td>Single arm</td>
</tr>
<tr>
<td>Sample size</td>
<td>Nilotinib N = 282 vs imatinib N = 283</td>
<td>N = 25</td>
</tr>
<tr>
<td>Age</td>
<td>&gt;18</td>
<td>2–18</td>
</tr>
<tr>
<td>MMR at 12 months$^a$</td>
<td>44% (38.4, 50.3)</td>
<td>60% (38.7, 78.9)</td>
</tr>
</tbody>
</table>

$^a$Results are for primary endpoint, for other information refer to USPI.
where \( f(D) \) is skeptical prior distribution (adult data are completely different from pediatric data and minimal response rate expected in pediatric patients), \( g(D) \) is the enthusiastic prior distribution (adult data reflects pediatric data) and relevance factor, \( a = \text{Pr(applicability of adult results)} \). When \( a = 0 \) or 0%, the prior distribution corresponds to \( f(D) \), which indicates 0% applicability of the adult study and the pediatric study should be a stand-alone study. When \( a = 1 \) or 100%, the prior distribution corresponds to \( g(D) \), which indicates 100% applicability of the adult study. A number between 0 and 1 for the relevance factor corresponds to the amount of information borrowed from the adult study, thus the interpretation of applicability of the adult study.

To illustrate the use of the prior in the case example, because the primary endpoint is binary, beta distribution is used as prior distribution. Figure 1 plots four different prior distributions: the red dashed line is the adult study result for the same indication, blue dashed line is the noninformative prior, black line is the skeptical prior, and the green dashed line is the 50%/50% mixture of the skeptical prior and the adult study result.

The skeptical \( f(D) \) is the skeptical pediatric stand-alone prior. This \( f(D) \) distribution is parameterized as Beta(1.2, 12). The distribution is centered around a response rate of roughly 9%, indicating no meaningful clinical benefit and places a low probability to observe a high response rate. The \( g(D) \) distribution is the adult prior and is Beta(8, 13) distribution. As indicated previously, the adult prior is centered around slightly less than the reported response rate and the variability is also discounted because the distribution is not centered at the 95% confidence interval. The 50%/50% mixture is when \( a = 0.5 \) in the mixture prior, and this indicates adult study is 50% applicable to pediatric study.

As a comparison, the noninformative prior distribution is plotted. The noninformative prior is a flat line, indicating prior knowledge has no value and offers no preference on response rate scale, for example, 10% response rate is equally likely as 90% response rate and equally likely as any other possible response rates.

### 3.3 Analyzing data

When the trial data become available, the posterior distribution is calculated by combining the prior distribution with the data. Since our prior distribution is a beta distribution, the posterior distribution can be easily updated based on Beta-Binomial conjugate distributions. For the posterior update based on conjugate mixture beta prior distribution, refer to Appendix and the References [13,19]. There are R packages available for calculations based on mixture prior distributions. For example, RBest, refer to https://cran.r-project.org/web/packages/RBesT/RBesT.pdf for details.

Figure 2 plots the posterior distribution given the four different prior distributions. The red dashed line indicates the likelihood, which depends only on the observed pediatric trial data, and it is the same for all plots. The green dashed line is the posterior distribution, superimposed with the prior distribution.
For noninformative prior, the posterior overlaid with the data because the prior information offers no value and we ignore the information. This case is equivalent to frequentist design to let the data speak for itself.

In the case of skeptical prior, where our prior belief is no benefit of the drug, and the pediatric trial does show benefit as measured based on response rate centered at 60%, it shifts our prior belief to the right toward the likelihood.

For the prior distributions of 100% adult and 50%/50% mixture, similarly, when the prior and likelihood are combined to form the posterior, the posterior is pulled toward the likelihood.

Suppose 30% response rate is considered minimally clinically meaningful in this population. If we consider efficacy to be defined as the probability of response rate greater than 30%, Table 2 provides the Bayesian estimate (using the prior and the observed data) of the response rate and the posterior probability of efficacy when different prior distributions are considered. Note using Bayesian approach, we can calculate the posterior distribution by essentially any target response rate because we use the same posterior probability distributions. However, the target clinically meaningful response rate should be prespecified.

In this case example if a threshold of posterior probabilities of 97.5% is considered as a success criterion, for 50% mixture (50% applicability of adult study), the posterior probability of efficacy in pediatric patients is greater than 99% (Table 2).

### 3.4 Sensitivity of priors

In general, the robustness of the study conclusion to changes in the proportion of adult data usage (e.g., 30% applicability of adult data instead of 50%) needs to be evaluated using a sensitivity analysis. This can be achieved by plotting the

---

**FIGURE 2** Posterior distributions for the case study of newly diagnosed Ph + CML in chronic phase

**TABLE 2** Bayesian estimates and 95% credible interval and posterior probability of efficacy given four prior distributions

<table>
<thead>
<tr>
<th>Prior</th>
<th>Bayesian estimate (95% cred. int.)</th>
<th>Posterior prob. of efficacy Pr(response &gt; 30%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skeptical</td>
<td>42.3 (27.4, 58.2)</td>
<td>94.5%</td>
</tr>
<tr>
<td>100% adult</td>
<td>50 (35.8, 64.2)</td>
<td>99.8%</td>
</tr>
<tr>
<td>50% mixture</td>
<td>50 (35.6, 64.2)</td>
<td>99.7%</td>
</tr>
<tr>
<td>Noninformative</td>
<td>59.5 (40.6, 76.6)</td>
<td>100%</td>
</tr>
</tbody>
</table>
posterior probability of efficacy in pediatric group against the applicability of adult data (relevance factor, \(a\) in Equation (1)) from 0% to 100% in increments of 5%.

In the case example above, Figure 3 shows the applicability of adult study against the posterior probability of efficacy in pediatric group. With 5% applicability of adult study, the posterior probability is greater than 98%. The result is very robust to the choice of probability of applicability and shows strong support for efficacy in pediatric group for this case example.

### 3.5 Sequential monitoring

Children as a vulnerable population in clinical research require special considerations to minimize the risk and assure prospect of direct clinical benefit when the risk to the child is more than minimal. This requirement impacts both study design and conduct. Monitoring for safety is particularly important during pediatric trials.\(^{20}\) Because of the limited patient population for study, sequential monitoring for efficacy to potentially stop a trial early can be a valuable tool to improve efficiency in pediatric cancer trials as well as assure that a minimum number of children are not needlessly exposed to ineffective drugs.

Bayesian analysis is a flexible tool to monitor trial outcome as data accumulate. For the above case example, we will show the execution of Bayesian sequential monitoring of response rate as data accumulates.

The 50%/50% mixture prior distribution is used to illustrate the approach. Updating the posterior probability of efficacy, defined as posterior probability of response rate greater than 30%, is first calculated when data from five patients become available. We update the posterior probability of efficacy given data from every patient thereafter becomes available.

Figure 4 illustrates the Bayesian sequential monitoring of the case example. The black solid line is the 50%/50% mixture prior illustrated earlier. The red dashed line is the observed trial data, that is, likelihood. The gray dashed line is the interim updating of posterior distribution of efficacy as data accumulates. The green dashed line is the posterior distribution when response evaluation from all 25 patients on the trial are known. The posterior distribution curve starts out by closely reassembling the shape of prior distribution. As more and more data become available, the posterior distribution curve moves gradually to the right and finally ends up as the same as the green dash curve, the final posterior distribution.

To illustrate further, using the 50%/50% mixture prior the Bayesian sequential monitoring on the posterior probability of efficacy as data accumulate is shown in Figure 5. The total number of patients, \(n\) enrolled in the trial at each evaluation and the number of responders, \(y\) out of the number of patients enrolled, are shown on the x-axis. The y-axis shows the calculated posterior probability of response rate greater than 30%. The red horizontal line is the 97.5% threshold for the posterior probability \(\geq 30\%\).
When only data on five patients were available, one responder was observed. The posterior probability of efficacy is slightly above 36% at this time point. When the next patient data were available, there were two responders out of the six patients. The posterior probability goes up to over 60%. When four out of eight patients were responders, the posterior probability of efficacy increases to 85%. But, the next three patients were nonresponders; therefore, the posterior probability of efficacy comes down. When the 19th patient data were available and there were 10 responders out of the 19, the posterior probability of efficacy crosses the 97.5% threshold. The posterior probability of efficacy stays on top of the 97.5% threshold until the end of the trial, where 15 responders were observed out of 25 patients. In this example, potentially the trial could have been stopped early for efficacy when 10 responders were observed out of 19 patients.

The Bayesian sequential monitoring shown is an extreme example of updating of the information. In practice, we can update for every two, five, or 10 patients as logistics allow. For example, if 20 pediatric patients are needed for the safety database, the timing of the first update on the posterior probability of efficacy can be coordinated when data from 20 patients become available.

### 4 | Designing a Study with Bayesian Approach

In this section, we will illustrate how to design a study with Bayesian approach incorporating sequential monitoring. We will use the idea that was introduced as community of priors.\(^{21}\) As described in Spiegelhalter et al (1994)\(^ {22}\) and Spiegelhalter et al (2004),\(^ {23}\) the community of priors is important because “the purpose of a trial is to collect data that
bring to conclusive consensus at termination opinions that had been diverse and indecisive at the outset.” This can be achieved by bringing in two prior distributions that represent two extreme ends of prior beliefs, one skeptical and the other enthusiastic.

Figure 6 shows the skeptical prior \( f(D) \) in Equation (1)). The \( f(D) \) is formulated as distribution centered at response rate with no meaningful clinical benefit and places very low probability to achieve a high response rate. In the case example, \( f(D) \) is Beta(1.2, 12), which centers at response rate roughly 9%, indicating no meaningful clinical benefit. The probability to observe a high response rate, for example, over 30% is 2.5% in \( f(D) \). The effective sample size (ESS) reflects the amount of borrowing by incorporating prior information.\(^{24}\) Because beta distribution is used, for the case of skeptical prior, the ESS can be easily derived to be 13.2.

At the other extreme is the enthusiastic prior \( g(D) \). In the case where the results of the adult study resulted in approval, the adult study prior would be an optimistic view that the drug would work equally well in the pediatric population. When adult study data are not available, in the case where the disease only occurs in children, the enthusiastic prior can be formulated from other sources of information, for example, external and historical controls, for example, prior clinical trials in other indications or literature. The idea is that adult or enthusiastic prior would center at response rate that would show clinical benefit and places low probability of no clinical benefit.
Figure 7 shows the adult/enthusiastic prior \( g(D) \) in Equation (1)). The \( g(D) \) is Beta(8, 13), which centers at response rate of about 38%, as indicated above, slightly discounted from the adult study trial. It places low probability of observing response rate less than 20% (2.5% chance). The ESS for adult prior is 21. Both the skeptical prior and adult prior are designed to have ESS in the same range as the possible trial sample size and is not intended to overwhelm the trial data.

There are three components in designing the trial incorporating skeptical, enthusiastic prior and sequential monitoring:

a) The maximum sample size accrual is 50.

b) We will stop the trial for efficacy if the posterior probability \( \Pr(RR > 0.3|\text{data, skeptical prior}) \geq 0.975 \).

c) We will stop the trial for futility if the posterior probability \( \Pr(RR < 0.35|\text{data, enthusiastic prior}) \geq 0.85 \).

The posterior probability is updated first when data on five patients are available and update the two posterior probabilities as data accumulates. The idea is that for the above component b), if we start out being skeptical about the drug to be efficacious, as data accumulates, if the posterior probability of observing response rate \( \geq 30\% \) is high (97.5%), the data have convinced us that the drug has efficacy and therefore, stop the trial early for efficacy. On the other hand, if the drug is ineffective, this is component c), if we start out being enthusiastic that the drug should work, as data accumulates, the posterior probability of observing response rate less than 35% is high, for example, 85%, then it suggests strong evidence against the drug, and therefore stop the trial early for futility.

As a note, this design is applicable to other rare disease situations, where the number of patients is limited and shares the same challenges as pediatric trials.

Figure 8 shows the simulation results of the Bayesian design properties if the trial is designed with the above three components. The true response rate (TrueRR) simulated is on the x-axis, from 15% to 60%. The average numbers of patients enrolled before stopping for either efficacy or futility are displayed on the second row of the x-axis.

The black curve is the probability of stopping for efficacy. When the true response rate is low, 15% to 30%, the chance of stopping for efficacy is 0%. This is the false positive probability and is similar to the frequentist version of the type I error rate. As the response rate increases, the chance of stopping for efficacy increases. When the true response rate is about 55%, 78% of times the trial may be stopped for efficacy, this is similar to power of the hypothesis test and the average number of patients when stopped is approximately 36.

The red dashed line is the probability of stopping for futility. When response rate is low, for example, 15%, the chance of stopping for futility is high, 98%. As the response rate increases, the chance of stopping for futility decreases. When the response rate is higher than 45%, the chance that the trial would be stopped for futility is less than 3%.

The green curve is the chance of inconclusive results leading to full accrual. For example when the true response rate is between 35% and 45%, there is no strong evidence based on the posterior probability to stop for either efficacy or futility.
The design incorporating community of priors with skeptical and enthusiastic representing both ends of the extreme beliefs can reasonably control the false-positive rate. The parameters in skeptical prior quantify the belief that large treatment effects are unlikely. For example, if the skeptical prior is selected to be Beta(3, 149) corresponding to ESS of 152, the probability of observing response rate $\geq 30\%$ would be close to 0. Because this strong belief of treatment not efficacious and the corresponding ESS will dominate the trial data, the posterior probability of response rate $\geq 30\%$ given true response rate of 60% is $<0.0001$. Similarly, the parameters selected in enthusiastic priors should also be compatible to the expected pediatric trial size. Otherwise, the posterior probability will be dominated by the strong enthusiastic prior, leading to higher chance of false positive.

Because the pediatric population is generally limited and there is high unmet medical need, alternative designs using noninformative prior instead of skeptical prior may be justified and implemented by replacing component b) in the design as: stop the trial for efficacy if the posterior probability $\Pr(RR > 0.3|\text{data, noninformative prior}) \geq 0.975$. Given other design components are the same, Figure 9 shows the simulation results of Bayesian properties using noninformative prior together with enthusiastic prior.

The notations and displays are the same as Figure 8. As a comparison, using noninformative and enthusiastic prior, the trial can be stopped for futility at similar probability levels as skeptical and enthusiastic prior since the futility decision is driven by the posterior probability given enthusiastic prior at this part of the trial.

The power is improved for noninformative and enthusiastic prior when the true response rate is $\geq 40\%$: the power is 52% for non informative-enthusiastic vs 10% skeptical-enthusiastic priors at 40% true response rate, 77% vs 25% respectively at 45% and 93% vs 52% respectively at 50% true response rates. When the true response rates are high, the power is about the same since the evidence in the trial data becomes the main driver for the efficacy.

The chance of inconclusive results leading to full accrual is also reduced for noninformative-enthusiastic priors for true response rates between 30% and 45%, for example, 49% inconclusive in noninformative-enthusiastic priors vs 78% in skeptical-enthusiastic priors at 35% true response rates. This is because the noninformative prior did not discount the trial data as the skeptical prior; therefore, the trial data drive the efficacy.

The average sample size is similar between the two designs when the true response rates are low, for example, $\leq 20\%$. Again the same enthusiastic prior is used to stop the trial for futility. The average sample size stays smaller for
response rates >20% in noninformative-enthusiastic priors than the skeptical-enthusiastic because there is no discount on trial data in noninformative prior.

However, the false-positive rates are higher in noninformative and enthusiastic priors when the true response rates are between 20% and 30% and can be as high as 13% at 30% true response rate. For the pediatric registration trial where no other confirmatory trial would be designed, this increased false-positive rate would need to be weighed against the potential risk, safety, treatment options or other relevant considerations.

The Bayesian estimates and 95% credible intervals are included in Figure 10. The plot compares side by side the designs using skeptical-enthusiastic priors and the noninformative-enthusiastic priors. The Bayesian estimates are close to the true response rates for both designs. The wider credible intervals for the design based on noninformative-enthusiastic priors reflect the average smaller sample sizes when the trial is stopped early for efficacy given true response rates are greater than 55%.

5 | CONCLUSIONS AND DISCUSSIONS

Challenges exist in the development of new pediatric anticancer drugs, particularly in the current era of targeted therapies. Potential study populations are small, and the evaluation of new agents based on the presence of specific molecular phenotypes rather than a specific disease or histology may warrant investigation of an agent across multiple cancer indications and/or age groups. Under FDARA, pediatric studies for oncology products will no longer be exempt based on orphan designation, and pediatric studies will be required for drugs directed at molecular targets observed in pediatric cancers. Plans for pediatric cancer investigations will be required for original marketing applications in the US for certain adult oncology drugs that are submitted on or after 18 August 2020. Under the changing landscape and regulation of pediatric anticancer drug development, innovative study designs including Bayesian approaches may help address current and future challenges.

We have described possible uses of Bayesian approaches in registrational pediatric cancer trials: design, sequential monitoring, and analysis of the trial using a case study. In the example, the mixture prior combining skeptical and
enthusiastic factors is appropriate because an important concern is to stop a trial as soon as possible if the treatment effect is small or in the adverse direction, thereby minimizing exposure of children to ineffective therapy. At the same time, the Bayesian approach also allows one to evaluate efficacy as data accumulates and stop the trial early when there is compelling efficacy evidence so that effective drugs are available to pediatric patients early. This is of particular advantage when accrual to a trial may be prolonged in the case of rare diseases.

The applicability in the mixture prior to weigh between the skeptical and enthusiastic prior is important and should be prespecified as part of the study design and analysis plan. There are two methods that can be employed to determine the amount of applicability of adult data, \( a \), by using existing data for similar drugs, which have completed studies in adults/pediatric or other external/historical controls or by eliciting experts who have experience treating the indications. When eliciting expert opinion, areas of consideration should include but not be limited to mechanism of action, pharmacological properties of the drug, dosing, disease process and diagnosis criteria, tolerability, outcome assessment and response criteria differences between adults and pediatrics. For basic references on methods for eliciting prior distributions from experts see chapter 5 in Spiegelhalter et al (2004)\(^23\) and Garthwaite et al (2005).\(^{25}\) For an example of the process involved in the elicitation of a Bayesian prior using expert elicitation, see Dallow et al (2018) [26] and Ye and Travis (2017) [27]. Analysis plans to stress test the proportion of applicability of adult data should be built in the trial statistical analysis plan as sensitivity analysis on the prior distribution by varying the applicability proportion from 0% to 100%.

A major benefit of the Bayesian approach is the flexibility it provides. The Bayesian analysis at any stage of the trial is not dependent on the stopping rules applied up to that point. Bayesian analysis remains valid even when a trial is terminated early for reasons other than the prespecified stopping rules. The efficacy can be calculated as posterior probability using the data to that point as usual.\(^{28}\) Additional flexibility of Bayesian sequential analysis is that it offers options to stop a trial early for both efficacy and futility when combined with skeptical and enthusiastic priors. There can be greater sample size saving when the data match or appear to be even stronger in the direction suggested by either skeptical or enthusiastic priors, but would be limited when the data are in between the skeptical or enthusiastic priors.

When designing a trial with Bayesian sequential monitoring, the accrual rate of the trial would be an important factor to consider. In general, the trial may stop enrollment at a given point in time, but all enrolled patients will be followed for outcome ascertainment. The benefit of Bayesian sequential monitoring is diminished if the accrual to the trial is fast relative to the outcome assessment. A useful analysis to evaluate the impact of accrual rate to final enrollment number can be done on average number of patients with outcome ascertained and ongoing at interim sequential analysis.

In summary, we believe Bayesian analysis is another approach that can be useful and allows flexibility to address some of the challenges to pediatric anticancer drug development. Bayesian designs and analysis can formally incorporate prior knowledge from adult studies into the trials, quantify uncertainty in the prior knowledge, and base the interpretation on a measure of probability of success. The sequential monitoring offers flexible monitoring of trial results as data accumulate. When combined with skeptical and enthusiastic priors, the analysis plan offers good trial design properties, including frequentist-equivalent controlled false-positive rate and sufficient power.

ACKNOWLEDGEMENTS
The authors thank Professor Frank Harrell for valuable comments. They thank the associate editor and two reviewers for their insightful comments to improve the manuscript.

DATA AVAILABILITY STATEMENT
Exempt for main paper. Use company data which is confidential to the company.

ORCID
Jingjing Ye https://orcid.org/0000-0003-0086-2228

REFERENCES


19. Bennett M. *Improving the Efficiency of Clinical Trial Designs by Using Historical Control Data or Adding a Treatment Arm to an Ongoing Trial.* St Catharine’s College, University of Cambridge; 2018.


**APPENDIX**

When the prior distribution is a conjugate mixture beta distribution, the prior can be written as \(\sum_{i} w_i C_i \text{Beta}(a_i, b_i)\). For the likelihood or the observed trial data, if observe \(r\) responders from \(n\) observations (binomial), then the posterior is also a mixture of beta distributions:

\[
\sum_{i} \frac{w_i C_i}{\sum w_j C_j} \text{Beta}(a_i + r, b_i + n - r),
\]

where \(C_i = \binom{n}{r} \frac{\Gamma(a_i + b_i) \Gamma(a_i + r) \Gamma(b_i + n - r)}{\Gamma(a_i) \Gamma(b_i) \Gamma(a_i + b_i + n)}\) and is the marginal likelihood of the data under prior \(\text{Beta}(a_i, b_i)\).

---

**How to cite this article:** Ye J, Reaman G, De Claro RA, Sridhara R. A Bayesian approach in design and analysis of pediatric cancer clinical trials. *Pharmaceutical Statistics.* 2020;1–13. [https://doi.org/10.1002/pst.2039](https://doi.org/10.1002/pst.2039)