Considerations for Pediatric Trial Designs and Analyses

Meehyung Cho, Zhiying Qiu, Jenny Ye, Hui Quan,
Peng-Liang Zhao
Biostatistics and Programming, Sanofi, Bridgewater, NJ 08807

Abstract

Pediatric trials are often conducted to obtain extended marketing exclusivity or to satisfy regulatory requirements. There are many challenges in designing and analyzing pediatric trials arising from special ethical issues and the relatively small accessible patient population. The application of conventional phase 3 trial designs to pediatrics is generally not realistic in some therapeutic areas. In this paper we review regulatory guidance and existing research in pediatrics. We then examine different approaches for designing a pediatric trial and analyzing outcomes. In particular, we consider weighted combination methods utilizing available adult data such as James-Stein shrinkage estimates, empirical shrinkage estimates and Bayesian methods. The performance of these methods is assessed through simulation.

Key Words: sample size, pediatric, James-Stein shrinkage estimator, empirical shrinkage estimator, Bayesian method, simulation

1. Introduction

Pediatric efficacy and safety assessments are often conducted either to obtain extended marketing exclusivity or to satisfy requirements from health authorities due to the potential for off-label use in pediatric patients. Challenges in the design and analysis of Phase 3 pediatric trials include addressing specific ethical issues and taking into consideration the relatively small candidate patient population. The application of a well-powered conventional phase 3 trial design for a pediatric population is not realistic in some therapeutic areas, including pediatric Type 2 diabetes mellitus (T2DM).

A recent example arose in response to a health authority request for a brief description of a Phase 3 pediatric trial design in T2DM where change in HbA1c, a blood factor associated with diabetes status, is the primary endpoint. The recruitment of T2DM pediatric patients is known to be very challenging due to the small pool of patients and competing trials from comparators. Thus, it is not practical to design a full scale trial with enough power that can be completed in the required timeframe. Alternative approaches should be considered.

The purpose of this paper is to summarize guidance from health authorities and related articles and to assess the performance of different approaches that could be considered and used for estimating sample size in designing a pediatric efficacy and safety trial with a continuous variable as the primary interest and its subsequent analyses. We discuss three weighted combination methods for borrowing information from multiple existing adult trials for designing a pediatric trial. The first is a fixed effects model with James-Stein (J-S) shrinkage estimation. The second is a random effects model with empirical

shrinkage estimation and the last uses Bayesian estimation. In the Bayesian method, we apply the idea of a power prior [1, 2, 3] by using a *discount factor* to quantify how much information to borrow from adult trials. But unlike Hobbs *et al*. [2] and Hobbs, Carlin and Mueller [3] who considered the discount factor to be a random variable, we assume that the discount factor is a fixed value.

2. Regulatory guidance and literature review

This section briefly summarizes guidance, guidelines or concept papers in this area from health authorities, particularly FDA or European Medicines Agency (EMA). Related articles are also briefly summarized.

The ICH Topic E11 guideline [4] provides a general outline of critical issues and approaches to the safe, efficient and ethical study in pediatric drug development. This guidance briefly mentions the possible extrapolation from adult efficacy data but does not discuss any details.

FDA has released draft industry guidance for planning pediatric studies [5]. It mainly covers the content of and process for submitting initial and amended pediatric study plans. It does not contain statistical details for designing a pediatric efficacy and safety trial.

EMA released the guideline on clinical trials in small populations [6] briefly discusses the statistical methods to consider in a limited patient population, suggesting combining knowledge from previous data through Bayesian method in data analysis. From a study design perspective, it mainly suggests using adaptive design, such as sequential design and response-adaptive design. EMA also issued a draft concept paper in 2012 on the extrapolation of efficacy and safety in medicine development which can be applicable to pediatrics [7]. It gives a general discussion with no particular statistical details.

The adhoc group for the development of implementing guidelines for Directive 2001/20/EC relating to good clinical practice in the conduct of clinical trials on medicinal products for human use, chaired by the European Commission released a document in 2008 on the topic of ethical considerations for clinical trials on medicinal products conducted with the paediatric population [8]. It provides a general review of a pediatric study design, but no specific statistical related context is discussed.

In the area of T2DM, EMA issued the report of a workshop on pediatric investigational plans in T2DM that was held in London on February 2013 [9]. One of the main objectives of the meeting was to identify possible approaches to improve practicability of conducting a pediatric trial inT2DM. One suggestion was to reduce the sample size: comments from some Biostatistics experts include possibly adopting Bayesian methods where adult information is employed as a prior or changing to a lower significance level (for example, change from a 0.05 to a 0.1 significance level).

In terms of the modeling and simulation approaches, Dunne *et al*. [10] reviewed 370 pediatric studies submitted to the FDA between 1998 and 2008. They examined the cases in which efficacy was extrapolated from adult data or other data and classified the type of pediatric data required to support extrapolation. Manolis and Pons [11] presented the approaches to implement modeling and simulation in pediatric drug development in the

European regulatory environment based on the EMA guidelines. Zisowsky, Krause and Dingemanse [12] reviewed regulatory guidelines from EMA and FDA for pediatric drug development from industry perspective, mainly focusing on the role of modeling and simulation in pediatrics, i.e., PK/PD modeling. Jadhav, Zhang and Gobburu [13] discussed a case study to use prior exposure response data from adult patients in PK/PD modeling and simulation for designing a pediatric trial in developing rationale for dosing recommendation in pediatrics.

For borrowing historical information while assessing treatment effects, several approaches have been proposed [1, 2, 3, 14, 15]. Further detailed review of this research area can be found in [16]. Two pediatric case studies which utilize adult data using the Bayesian approach have been discussed in the literature [17, 18].

3. Method for estimating sample size and analysis

Let h be the number of available adult trials. Suppose the pediatric and all adult trials are 2-arm with equal sample size $n_i, i = 0, 1, \cdots, h$ per arm. Let n_0 be the sample size per arm in the pediatric trial that we plan to design. Then $N = \sum_{i=1}^{h} n_i$ is the total number of patients per arm in all adult trials. Suppose $\delta_i, i = 1, \cdots, h$, is the true treatment effect in the existing adult trial i. Let δ_0 be the true treatment effect in the pediatric trial. We define the treatment effect as the mean difference between 2 treatment groups. Then the estimator $\hat{\delta}_i$ of the treatment effect (defined as the sample mean difference) given the true treatment effect for trial i, $i = 0, \cdots, h$, follows

$$\left(\hat{\delta}_i \middle| \delta_i\right) \sim N\left(\delta_i, \frac{2{\sigma_i}^2}{n_i}\right),$$

where ${\sigma_i}^2$ is the variance of the treatment effect in the ith trial.

In the following sections, we consider both fixed and random effect models that allow borrowing information from existing adult trials for designing the pediatric trials. In fixed effect models, we consider two options for δ_0 to calculate the sample size, assuming either a known fixed value or following a normal distribution with mean and variance γ^2. In random effect models, the existing adult trials and the pediatric trial we design are assumed to be similar in the sense that the treatment effect $\delta_0, \delta_1, \cdots, \delta_h$ are from a common normal distribution

$$\delta_i \sim \mathrm{N}\left(\delta, \tau^2\right), i = 0, \cdots, h$$

where τ^2 represents between-study variation. Then $\hat{\delta}_i$ $(i = 0, \cdots, h)$ can be calculated as following the normal distribution with mean δ and variance $\tau^2 + \frac{2{\sigma_i}^2}{n_i}$. That is,

$$\hat{\delta}_i \sim N\left(\delta, \tau^2 + \frac{2{\sigma_i}^2}{n_i}\right).$$

3.1 Classical and typical fixed-effect approach

We first consider the classical and typical sample size determination approach without borrowing any adult information. In order to better compare with all extrapolation methods, we also consider that the true pediatric treatment effect is either a known fixed value or a random variable in calculating the sample size and power.

a. True pediatric treatment effect is a known fixed value

Suppose the true pediatric treatment effect δ_0 is a known fixed value. Then, for a given δ_0 and the desired power $1-\beta$, the required sample size n_0 per arm in the pediatric trial is

$$n_0 = \frac{2{\sigma_0}^2(z_{1-\alpha/2}+z_{1-\beta})^2}{{\delta_0}^2}, \quad (1)$$

where $z_{1-\alpha/2}$ represents the $1-\alpha$ percentile of the standard normal distribution and is 1.96 when $\alpha = 0.05$.

b. True pediatric effect is a random variable with variability

The known fixed value δ_0 in (1) is an estimate from historical data. It certainly has variability. Instead of a known fixed value of δ_0, suppose the true pediatric treatment effect is distributed as

$$\delta_0 \sim N(\delta, \gamma^2),$$

where δ is a known value that can be estimated from the adult data. Then the sample size n_0 for the desired power $1-\beta$ in the pediatric trial can be obtained by solving $f(n_0) = 1-\beta$, where

$$f(n_0) = 1 - E_{\delta_0}\left(\Phi\left(z_{1-\alpha/2} - \frac{\delta_0}{\sqrt{\frac{2{\sigma_0}^2}{n_0}}}\middle|\delta_0\right)\right).$$

3.2 Methods incorporating data from adult studies

3.2.1 James-Stein (J-S) shrinkage fixed-effects estimator

Quan *et al*. [19] proposed to use the J-S shrinkage estimator for multi-regional clinical trial design to quantify regional treatment effects. We apply the concept of J-S shrinkage estimator to pediatric trial design. Let the overall treatment effect be $\bar{\delta} = \frac{n_0\delta_0+\sum_{i=1}^{h} n_i\delta_i}{\sum_{i=0}^{h} n_i} = \frac{n_0\delta_0+\sum_{i=1}^{h} n_i\delta_i}{N+n_0}$. The J-S shrinkage estimator $\check{\delta}_0$ is written as

$$\check{\delta}_0 = b_0\hat{\delta}_0 + (1-b_0)\bar{\delta},$$

where b_0 is obtained using the results shown in Quan *et al*. [19]

$$b_0 = \frac{\sum_{i=0}^{h} n_i(\delta_i-\bar{\delta})^2}{\sum_{i=0}^{h} n_i(\delta_i-\bar{\delta})^2+\sum_{i=0}^{h} 2{\sigma_i}^2}. \quad (2)$$

Given the observed treatment effects ($\hat{\delta}_i$) from adult trials, we can replace δ_i with $\hat{\delta}_i, i = 1, \ldots h$ in (2). Let $\check{\delta}^* = \frac{\sum_{i=1}^{h} n_i\hat{\delta}_i}{N+n_0}$. In data analysis after the pediatric trial is completed, we can use $\hat{\delta}_0$ to replace δ_0 for estimating b_0. The estimator $\hat{\delta}_0$ is unknown at the time of the pediatric study design, and suppose $\hat{\delta}_0|\delta_0 \sim N\left(\delta_0, \frac{2{\sigma_0}^2}{n_0}\right)$. Then, the estimator of b_0 can be written as

$$\hat{b}_0 = \frac{E_{\hat{\delta}_0}\left(\sum_{i=1}^{h} n_i\left(\hat{\delta}_i - \check{\delta}^* - \frac{n_0\hat{\delta}_0}{N+n_0}\right)^2 + n_0\left(\hat{\delta}_0 - \check{\delta}^* - \frac{n_0\hat{\delta}_0}{N+n_0}\right)^2\right)}{E_{\hat{\delta}_0}\left(\sum_{i=1}^{h} n_i\left(\hat{\delta}_i - \check{\delta}^* - \frac{n_0\hat{\delta}_0}{N+n_0}\right)^2 + n_0\left(\hat{\delta}_0 - \check{\delta}^* - \frac{n_0\hat{\delta}_0}{N+n_0}\right)^2\right) + \sum_{i=0}^{h} 2{\sigma_i}^2}.$$

The J-S shrinkage estimator $\check{\delta}_0$is

$$\check{\delta}_0 = \left(\hat{b}_0 + (1-\hat{b}_0)\frac{n_0}{N+n_0}\right)\hat{\delta}_0 + (1-\hat{b}_0)\check{\delta}^*$$

with variance

$$var(\check{\delta}_0|\delta_0) = \left(\hat{b}_0 + (1-\hat{b}_0)\frac{n_0}{N+n_0}\right)^2 \frac{2{\sigma_0}^2}{n_0} + (1-\hat{b}_0)^2 \frac{\sum_{i=1}^{h} n_i 2{\sigma_i}^2}{(N+n_0)^2}.$$

Note that if the observed variability across observed treatment effects $(\hat{\delta}_i, i = 1, \ldots h)$ from adult trials is large, then the J-S estimator will be close to $\hat{\delta}_0$ and $\hat{b}_0$ will approach to 1.

Similar to the discussion in Section 2.1, we consider two options for the true pediatric treatment effect in calculating the sample size and power: either a known fixed value or a random variable.

a. True pediatric treatment effect is a known fixed value

Suppose the true pediatric treatment effect δ_0 is a known fixed value. Then, for a given δ_0 and then desired power $1-\beta$, the required sample size n_0 per arm in the pediatric trial is calculated by solving$f(n_0) = 1-\beta$, where

$$\begin{aligned} f(n_0) &= \Pr\left(\frac{\check{\delta}_0}{sd(\check{\delta}_0|\delta_0)} > z_{1-\alpha/2}\middle|\,\delta_0\right) \\ &= \Pr\left(\frac{\left(\hat{b}_0 + (1-\hat{b}_0)\frac{n_0}{N+n_0}\right)\hat{\delta}_0 + (1-\hat{b}_0)\check{\delta}^*}{sd(\check{\delta}_0|\delta_0)} > z_{1-\alpha/2}\middle|\,\delta_0\right) \\ &= \Pr\left(\frac{\hat{\delta}_0 - \delta_0}{\sqrt{\frac{2{\sigma_0}^2}{n_0}}} > \frac{\frac{z_{1-\alpha/2}sd(\check{\delta}_0|\delta_0) - (1-\hat{b}_0)\check{\delta}^*}{\left(\hat{b}_0 + (1-\hat{b}_0)\frac{n_0}{N+n_0}\right)} - \delta_0}{\sqrt{\frac{2{\sigma_0}^2}{n_0}}}\middle|\,\delta_0\right) \\ &= 1 - \Phi\left(\frac{\frac{z_{1-\alpha/2}sd(\check{\delta}_0|\delta_0) - (1-\hat{b}_0)\check{\delta}^*}{\left(\hat{b}_0 + (1-\hat{b}_0)\frac{n_0}{N+n_0}\right)} - \delta_0}{\sqrt{\frac{2{\sigma_0}^2}{n_0}}}\middle|\,\delta_0\right). \end{aligned}$$

b. True pediatric effect is a random variable with variability

Instead of a known fixed value of δ_0, suppose the true pediatric treatment effect is distributed as

$$\delta_0 \sim N(\delta, \gamma^2),$$

where δ is a known value that can be estimated from the adult data. Then the required sample size n_0 per arm for the desired power 1-β in the pediatric trial can be calculated by solving $f(n_0) = 1-\beta$, where

$$
\begin{aligned}
f(n_0) &= E_{\delta_0}\left(\Pr\left(\frac{\check{\delta}_0}{sd\left(\check{\delta}_0|\delta_0\right)} > z_{1-\alpha/2}\middle|\,\delta_0\right)\right)\\
&= E_{\delta_0}\left(\Pr\left(\frac{\left(\hat{b}_0+\left(1-\hat{b}_0\right)\frac{n_0}{N+n_0}\right)\hat{\delta}_0+\left(1-\hat{b}_0\right)\check{\delta}^*}{sd\left(\check{\delta}_0|\delta_0\right)} > z_{1-\alpha/2}\middle|\,\delta_0\right)\right)\\
&= E_{\delta_0}\left(\Pr\left(\hat{\delta}_0 > \frac{z_{1-\alpha/2}sd\left(\check{\delta}_0|\delta_0\right)-\left(1-\hat{b}_0\right)\check{\delta}^*}{\left(\hat{b}_0+\left(1-\hat{b}_0\right)\frac{n_0}{N+n_0}\right)}\middle|\,\delta_0\right)\right)\\
&= E_{\delta_0}\left(\Pr\left(\frac{\hat{\delta}_0-\delta_0}{\sqrt{\frac{2\sigma_0{}^2}{n_0}}} > \frac{\frac{z_{1-\alpha/2}sd\left(\check{\delta}_0|\delta_0\right)-\left(1-\hat{b}_0\right)\check{\delta}^*}{\left(\hat{b}_0+\left(1-\hat{b}_0\right)\frac{n_0}{N+n_0}\right)}-\delta_0}{\sqrt{\frac{2\sigma_0{}^2}{n_0}}}\middle|\,\delta_0\right)\right)\\
&= 1-E_{\delta_0}\left(\Phi\left(\frac{\frac{z_{1-\alpha/2}sd\left(\check{\delta}_0|\delta_0\right)-\left(1-\hat{b}_0\right)\check{\delta}^*}{\left(\hat{b}_0+\left(1-\hat{b}_0\right)\frac{n_0}{N+n_0}\right)}-\delta_0}{\sqrt{\frac{2\sigma_0{}^2}{n_0}}}\middle|\,\delta_0\right)\right).
\end{aligned}
$$

3.2.2 Empirical shrinkage estimator under random effects model

For generalization purpose, we consider a random effect model in this section, taking into account between-study variability. One overall estimator of treatment effect $\tilde{\delta}$ from a random effects model follows a normal distribution. That is,

$$\tilde{\delta} = \frac{\sum_{i=0}^{h} w_i\hat{\delta}_i}{\sum_{i=0}^{h} w_i} \sim N\left(\delta,\ \ \frac{1}{\sum_{i=0}^{h} w_i}\right),$$

where $w_i = \frac{1}{\tau^2+2\,\sigma_i{}^2/n_i}$.

Denote $a_0 = \frac{\tau^2}{\tau^2+2\sigma_0{}^2/n_0} = \tau^2 w_0$. Then, the empirical shrinkage estimate ($\tilde{\delta}_0$) of the pediatric trial treatment effect δ_0 borrowing information from adult trials can be written by incorporating the adult data as

$$
\begin{aligned}
\tilde{\delta}_0 &= a_0\hat{\delta}_0 + (1-a_0)\tilde{\delta}\\
&= a_0\hat{\delta}_0 + (1-a_0)\frac{\sum_{i=0}^{h} w_i\hat{\delta}_i}{\sum_{i=0}^{h} w_i}\\
&= \left(a_0+\frac{(1-a_0)w_0}{\sum_{i=0}^{h} w_i}\right)\hat{\delta}_0 + (1-a_0)\frac{\sum_{i=1}^{h} w_i\hat{\delta}_i}{\sum_{i=0}^{h} w_i}\\
&= A_0\hat{\delta}_0 + (1-A_0)\hat{\delta}^*, \qquad (3)
\end{aligned}
$$

where $A_0 = a_0 + \frac{(1-a_0)w_0}{\sum_{i=0}^{h} w_i} = \frac{\tau^2 w_0\sum_{i=1}^{h} w_i + w_0}{w_0+\sum_{i=1}^{h} w_i}$ and $\hat{\delta}^* = \frac{\sum_{i=1}^{h} w_i\hat{\delta}_i}{\sum_{i=1}^{h} w_i}$. It should be noted that $1-A_0$ is considered a weight to the information from adult studies.

The conditional mean and variance of empirical shrinkage estimator $\tilde{\delta}_0$ in (3) given the observed adult data are calculated as

$$E(\tilde{\delta}_0|\hat{\delta}^*) = A_0 E(\hat{\delta}_0) + (1 - A_0)\hat{\delta}^* = A_0\delta + (1 - A_0)\hat{\delta}^* \text{ and}$$
$$Var(\tilde{\delta}_0|\hat{\delta}^*) = {A_0}^2\left[Var\left(E(\hat{\delta}_0|\delta_0)\right) + E\left(Var(\hat{\delta}_0|\delta_0)\right)\right] = {A_0}^2\left[\tau^2 + \frac{2{\sigma_0}^2}{n_0}\right].$$

Then the required sample size n_0 per arm for the desired power 1-β in the pediatric trial is calculated by solving $f(n_0) = 1 - \beta$, where

$$\begin{aligned} f(n_0) &= \Pr\left(\frac{\tilde{\delta}_0}{sd(\tilde{\delta}_0|\delta_0)} > z_{1-\alpha/2}\middle|\, \delta_0, \hat{\delta}^*\right) \\ &= \Pr\left(\frac{A_0\hat{\delta}_0 + (1-A_0)\hat{\delta}^*}{sd(\tilde{\delta}_0|\delta_0)} > z_{1-\alpha/2}\middle|\, \delta_0, \hat{\delta}^*\right) \\ &= \Pr\left(\hat{\delta}_0 > \frac{z_{1-\alpha/2}sd(\tilde{\delta}_0) - (1-A_0)\hat{\delta}^*}{A_0}\middle|\, \delta_0, \hat{\delta}^*\right) \\ &= \Pr\left(\frac{\hat{\delta}_0 - \delta_0}{sd(\hat{\delta}_0|\delta_0)} > \frac{\frac{z_{1-\alpha/2}sd(\tilde{\delta}_0|\hat{\delta}^*) - (1-A_0)\hat{\delta}^*}{A_0} - \delta_0}{sd(\hat{\delta}_0|\delta_0)}\middle|\, \delta_0, \hat{\delta}^*\right) \\ &= \mathbf{1} - \boldsymbol{\Phi}\left(\frac{\frac{z_{1-\alpha/2}sd(\tilde{\delta}_0|\hat{\delta}^*) - (1-A_0)\hat{\delta}^*}{A_0} - \delta_0}{\sqrt{\frac{2{\sigma_0}^2}{n_0}}}\right) \end{aligned}$$

3.2.3 Bayesian method

We generalize and extend the Bayesian method proposed by Schoenfeld, Zheng and Finkelstein [18]. Suppose$\Pr(\delta) \propto 1$, a non-informative prior for δ of the treatment effect. The posterior distribution of δ based on the existing adult data is calculated as

$$\Pr\left(\delta|\hat{\delta}_i, i = 1, \cdots h\right) \propto \prod_{i=1}^{h} L\left(\hat{\delta}_i|\delta\right)\Pr(\delta).$$

That is, the updated prior distribution of treatment effect δ before the pediatric study data follows

$$\delta|\hat{\delta}_i, i = 1, \cdots h, \tau^2 \sim N\left(\frac{\sum_{i=1}^{h} w_i\hat{\delta}_i}{\sum_{i=1}^{h} w_i}, \frac{1}{\sum_{i=1}^{h} w_i}\right). \quad (4)$$

Let $\hat{\delta}^* = \frac{\sum_{i=1}^{h} w_i\hat{\delta}_i}{\sum_{i=1}^{h} w_i}$. Then the predictive distribution of the pediatric treatment effect δ_0 given the adult data $\hat{\delta}_i, i = 1, \cdots h$ and the between-study variability τ^2 can be calculated from (4) as

$$\delta_0|\tau^2, \hat{\delta}^* \sim N\left(\hat{\delta}^*, \ \frac{1}{\sum_{i=1}^{h} w_i} + \tau^2\right). \quad (5)$$

This predictive distribution of δ_0 in (5) is used as the prior distribution of δ_0 before the pediatric trial for calculating the conditional posterior distribution of $\delta_0|\tau^2, \hat{\delta}_0, \hat{\delta}^*$ which is of interest for calculating the sample size of the pediatric trial.

Suppose d is the discount factor on the prior distribution, adapting the power prior concept for borrowing the information from adult studies. Based on Bayes theorem, the conditional posterior distribution of δ_0 given the between-study variability $\tau^2$and existing adult data along with the data from the pediatric study we plan to design is written as

$$\Pr\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right) \propto L\left(\hat{\delta}_0|\delta_0\right) Pr\left(\delta_0|\tau^2,\hat{\delta}^*\right)^d \quad (6)$$

The posterior in (6) follows a normal distribution with mean

$$E\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right) = \frac{\left(\frac{2{\sigma_0}^2}{n_0}\right)^{-1}\hat{\delta}_0 + \left(\frac{1}{\sum_{i=1}^{h} w_i d} + \frac{\tau^2}{d}\right)^{-1}\hat{\delta}^*}{\left(\frac{2{\sigma_0}^2}{n_0}\right)^{-1} + \left(\frac{1}{\sum_{i=1}^{h} w_i d} + \frac{\tau^2}{d}\right)^{-1}}$$

$$= C_0\hat{\delta}_0 + (1-C_0)\hat{\delta}^*$$

and variance

$$Var\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right) = C_0\frac{2{\sigma_0}^2}{n_0},$$

where

$$C_0 = \frac{\left(\frac{2{\sigma_0}^2}{n_0}\right)^{-1}}{\left(\frac{2{\sigma_0}^2}{n_0}\right)^{-1} + \left(\frac{1}{\sum_{i=1}^{h} w_i d} + \frac{\tau^2}{d}\right)^{-1}}$$

First, $\Pr(\delta_0 < 0) = \Pr\left(\frac{\delta_0 - E\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right)}{sd\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right)} < \frac{-E\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right)}{sd\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right)}\right) \leq \alpha/2$, implying $\frac{E\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right)}{sd\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right)} > z_{1-\alpha/2}$. Then the required sample size n_0 per arm for the desired power 1-β in the pediatric trial is obtained by computing$f(n_0) = 1-\beta$, where

$$\begin{aligned}
f(n_0) &= \Pr\left(\frac{E\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right)}{sd\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right)} > z_{1-\alpha/2}\middle|\, \delta_0,\hat{\delta}^*\right)\\
&= \Pr\left(\frac{C_0\hat{\delta}_0 + (1-C_0)\hat{\delta}^*}{sd\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right)} > z_{1-\alpha/2}\middle|\, \delta_0,\hat{\delta}^*\right)\\
&= \Pr\left(C_0\hat{\delta}_0 > z_{1-\alpha/2}sd\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right) - (1-C_0)\hat{\delta}^*\middle|\delta_0,\hat{\delta}^*\right)\\
&= \Pr\left(\hat{\delta}_0 > \frac{z_{1-\alpha/2}sd\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right) - (1-C_0)\hat{\delta}^*}{C_0}\middle|\, \delta_0,\hat{\delta}^*\right)\\
&= \Pr\left(\frac{\hat{\delta}_0-\delta_0}{sd\left(\hat{\delta}_0|\delta_0\right)} > \frac{\frac{z_{1-\alpha/2}sd\left(\delta_0|\tau^2,\hat{\delta}_0,\hat{\delta}^*\right) - (1-C_0)\hat{\delta}^*}{C_0} - \delta_0}{sd\left(\hat{\delta}_0|\delta_0\right)}\middle|\, \delta_0,\hat{\delta}^*\right)\\
&= \mathbf{1} - \boldsymbol{\Phi}\left(\frac{\frac{z_{1-\alpha/2}\sqrt{C_0\frac{2{\sigma_0}^2}{n_0}} - (1-C_0)\hat{\delta}^*}{C_0} - \delta_0}{\sqrt{\frac{2{\sigma_0}^2}{n_0}}}\right)
\end{aligned}$$

4. Simulation results

Simulation studies were performed to investigate the performance of the aforementioned methods in terms of required sample size or power. We also conducted simulations comparing the weights for adult trials between the empirical shrinkage estimator and the Bayesian method, under various between-study variability τ^2, given a pediatric sample size n_0. The weights for adult trials in the J-S shrinkage fixed estimator depend on the variability of the observed treatment effects in existing adult trials; the higher the variability is, the lower the weight is.

All simulations were conducted in the T2DM setting where the primary efficacy endpoint is often the treatment effect in HbA1c change from baseline after 6 months of treatment.

In the simulation, we set $h = 3$, the number of existing adult trials, Without loss of generality, we assumed a balanced design, with equal sample size $n_i = 102$ per arm for all 3 adult trials and the same common standard deviation σ_i of 1.1%, $i = 1, \cdots, 3$. We set the same observed treatment effect of 0.5% in all 3 adult trials for the HbA1c change from baseline in all simulations except for J-S shrinkage fixed effect estimator: that is, $\hat{\delta}_1 = \hat{\delta}_2 = \hat{\delta}_3 = 0.5\%$. The sample size of 102 patients per arm corresponds to the required size for obtaining 90% power to detect the treatment effect of 0.5% between treatment arms with a common standard deviation of 1.1% (2-sided, alpha=0.05) using the classical sample size calculation method.

For J-S shrinkage fixed effect estimator method, unlike other methods in which sample size depends on only the overall observed treatment effect across existing adult trials, the sample size reduction depends on the variability of the observed treatment effects in existing adult trials. If the estimated treatment effect from the existing adult trials is almost homogeneous, then the required sample size for pediatric trial from J-S shrinkage fixed effect estimator method can be drastically reduced. Thus, we considered the following two scenarios for observed treatment effects from 3 adult trials for the J-S shrinkage estimator method: (1) adult studies with a larger observed sample variance: $\hat{\delta}_1 = 0.3, \hat{\delta}_2 = 0.5,$ and $\hat{\delta}_3 = 0.7$; (2) adult studies with a smaller observed sample variance: $\hat{\delta}_1 = 0.45, \hat{\delta}_2 = 0.5,$ and $\hat{\delta}_3 = 0.55$.

For exploring the Bayesian method, we considered a discount factor d of 0.2, 05 or 1.

For the pediatric trial that we aimed to calculate the sample size, we assumed a 2-arm, equal randomization ratio and the common standard deviation of 1.1% (i.e.,$\sigma_0 = 1.1$). For both the classical method and J-S shrinkage fixed effect estimator, the true pediatric treatment effect $\delta_0$is set to 0.5% if it is assumed to be a known fixed value or set to follow a normal distribution with mean 0.5% and variance γ^2 (ie, $\delta_0 \sim N\ (0.5\%, \gamma^2)$) when it is considered a random variable. For the empirical shrinkage estimator and the Bayesian method, the true pediatric treatment effect was assumed to follow a normal distribution with mean of common treatment effect of 0.5% and the variance τ^2, between-study variability (i.e., $\delta_i \sim N\ (0.5\%, \tau^2)$), $i = 0, 1, \cdots, 3$.

4.1 Comparing sample sizes given a study power

We first compared sample sizes among the classical approach, J-S shrinkage fixed effect estimator and Bayesian method for the pediatric trial, given a study power of 80% and 85%, respectively, in Figure 1. As mentioned above, the J-S shrinkage fixed effect

estimator was explored under two scenarios for the observed treatment effect from 3 adult trials, and the Bayesian method was examined with a small discount factor $d = 0.2$.

We considered the variability (between-study variability τ^2 for the Bayesian method or variance γ^2 of the true pediatric treatment effect δ_0) up to 0.04% in Figure 1. In T2DM, the between-study variability (τ^2) of the estimated treatment effect of HbA1c change or the variability around the true pediatric treatment effect is expected to be small. For example, 0.04% of variability (i.e., 0.2% of standard deviation) indicates the lower limit of the 95% confidence interval for the overall estimated treatment effect is around 0.1%.

As seen from Figure 1, both the J-S shrinkage fixed effect estimator and Bayesian method utilizing adult information reduced the required sample size for the pediatric trial, compared to the classical approach. The Bayesian method that utilized about 20% of the information from adult trials resulted in more than 30% reduction in required sample size, compared to the classical approach; the difference in sample size between these methods was bigger when τ^2 became larger. Also, the Bayesian approach appeared to show a similar sample size with increased τ^2. When the observed treatment effect in 3 existing adult trials was very similar to each other ($\hat{\delta}_1 = 0.45, \hat{\delta}_2 = 0.5, \hat{\delta}_3 = 0.55$), the required sample size for the pediatric trial was drastically reduced, even lower than the number from the Bayesian method.

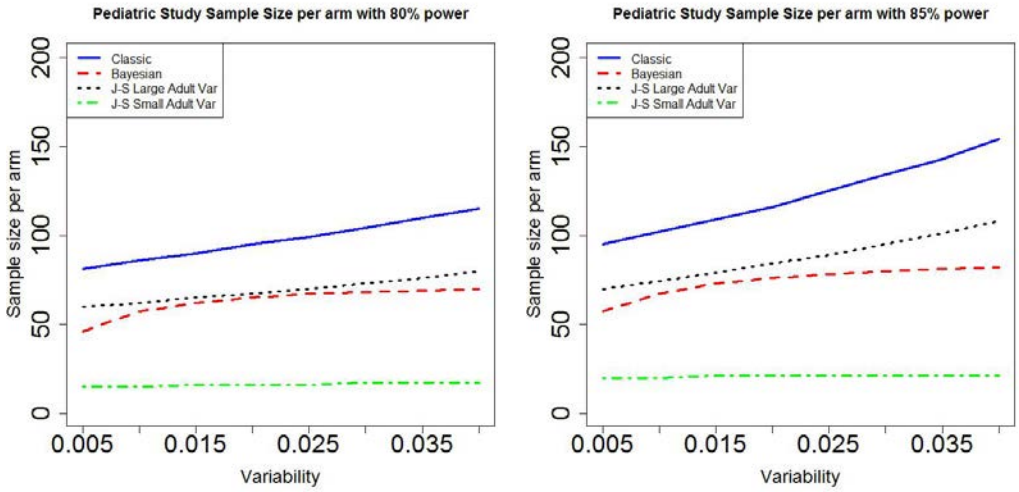

Figure 1: Required sample size per arm in the pediatric trial to obtain 80% and 85% power from three methods (2-sided and α=0.05): Classic indicates the classical method with $\delta_0 \sim N(0.5\%, \gamma^2)$, where γ^2 represents variability; Bayesian indicates the Bayesian method with $\delta_0 \sim N(0.5\%, \tau^2)$ and $d = 0.2$, where τ^2 represents variability; J-S Large Adult Var indicates the J-S shrinkage fixed effect estimator with $\hat{\delta}_1 = 0.3, \hat{\delta}_2 = 0.5, \hat{\delta}_3 = 0.7$; J-S Small Adult Var indicates the J-S shrinkage fixed effect estimator with $\hat{\delta}_1 = 0.45, \hat{\delta}_2 = 0.5, \hat{\delta}_3 = 0.55$. Both J-S estimators assumed $\delta_0 \sim N(0.5\%, \gamma^2)$, where γ^2 represents variability.

4.2 Comparing powers given sample sizes in Bayesian method under different d values

Table 1 to Table 3 present power calculations of Bayesian method, given a sample size, under different d values. Overall, the study power was decreasing when τ^2 increased, given a sample size.

Table 1 shows the powers, under various sample sizes per arm, with $d = 1$ (fully utilize the existing adult data). The results may be bias due to the dominance of adult data, that is, too much information is borrowed. Thus, we considered $d = 0.5$ and $d = 0.2$, in which the contribution of the adult data was reduced to a half or 20% (Table 2 and Table 3). As expected, the study powers, given a sample size, were decreased, more noticeably with increased τ^2. Despite the reduced power with larger τ^2, the resulted study power was greater than the classical approach. The discount value $d = 0.5$ may still cause agency's concern of borrowing too much of the adult information, while $d = 0.2$ which limits to 20% of information from existing adult studies may be acceptable to health authorities.

As shown in Table 3, with the discount value of 0.2 (i.e., $d = 0.2$), about 80 pediatric patients per arm was needed to obtain the 90% power when τ^2=0.005. When τ^2ranged from 0.02% to 0.03%, between 80 and 100 pediatric patients per arm was required to obtain the 90% power. It is known that the required sample size is 102 per arm for achieving 90% power with a common standard deviation of 1.1%, when the true pediatric treatment effect δ_0 is assumed to be a known fixed value of 0.5%. As discussed in Section 3.1, given a fixed sample size with at least n_0=60, the study power remained similar with increased τ^2.

Table 1: Power calculations using Bayesian method with $d = 1$, α=0.05 and 2-sided, under various pediatric sample sizes(n_0) per arm and between-study variability (τ^2)

	τ^2=0.005	τ^2 =0.01	τ^2 =0.02	τ^2 =0.025	τ^2 =0.03	τ^2=0.04
n_0=20	>99.9%	99.6%	89.7%	82.1%	75.2%	64.5%
n_0=40	>99.9%	99.6%	94.0%	90.3%	86.7%	80.7%
n_0=60	>99.9%	99.7%	96.8%	94.8%	92.8%	89.4%
n_0=80	>99.9%	99.8%	98.3%	97.2%	96.2%	94.3%
n_0=100	>99.9%	99.9%	99.1%	98.5%	98.0%	96.9%
n_0=120	>99.9%	99.9%	99.5%	99.3%	99.0%	98.4%

Table 2: Power calculations using Bayesian method with $d = 0.5$, α=0.05 and 2-sided, under various pediatric sample sizes(n_0) per arm and between-study variability (τ^2)

	τ^2=0.005	τ^2 =0.01	τ^2 =0.02	τ^2 =0.025	τ^2 =0.03	τ^2 =0.04
n_0=20	95.3%	80.8%	59.9%	54.2%	50.1%	44.9%
n_0=40	96.9%	89.6%	77.9%	74.2%	71.4%	67.4%
n_0=60	98.3%	94.4%	87.8%	85.5%	83.7%	81.1%
n_0=80	99.1%	97.0%	93.3%	92.0%	90.9%	89.3%
n_0=100	99.5%	98.4%	96.4%	95.6%	95.0%	94.1%
n_0=120	99.7%	99.2%	98.1%	97.7%	97.3%	96.8%

Table 3: Power calculations using Bayesian method with $d = 0.2$, α=0.05 and 2-sided, under various pediatric sample sizes(n_0) per arm and between-study variability (τ^2)

	τ^2=0.005	τ^2 =0.01	τ^2 =0.02	τ^2 =0.025	τ^2 =0.03	τ^2 =0.04
n_0=20	58.1%	47.7%	39.8%	38.0%	36.7%	35.1%
n_0=40	76.8%	69.6%	63.2%	61.5%	60.3%	58.7%
n_0=60	87.1%	82.5%	78.1%	76.9%	76.0%	74.8%
n_0=80	92.9%	90.2%	87.4%	86.6%	86.0%	85.2%
n_0=100	96.2%	94.6%	92.9%	92.4%	92.1%	91.6%
n_0=120	98.0%	97.1%	96.1%	95.8%	95.6%	95.3%

4.3 Comparing weights to the adult data given a sample size of the pediatric trial between empirical shrinkage estimator and Bayesian method

Figure 2 shows that for both the empirical shrinkage estimator and Bayesian method, weights for adult studies decreases as τ^2 increases, given a fixed pediatric sample size n_0. For empirical shrinkage estimator, the weight was the largest in the sense that it had very high degree of shrinkage and the data from the existing adult trials almost dominated the final pediatric results, especially when τ^2 was very small. For Bayesian method, the more adult information was borrowed (i.e., the greater the discount value d is), the larger the weight for adult data was.

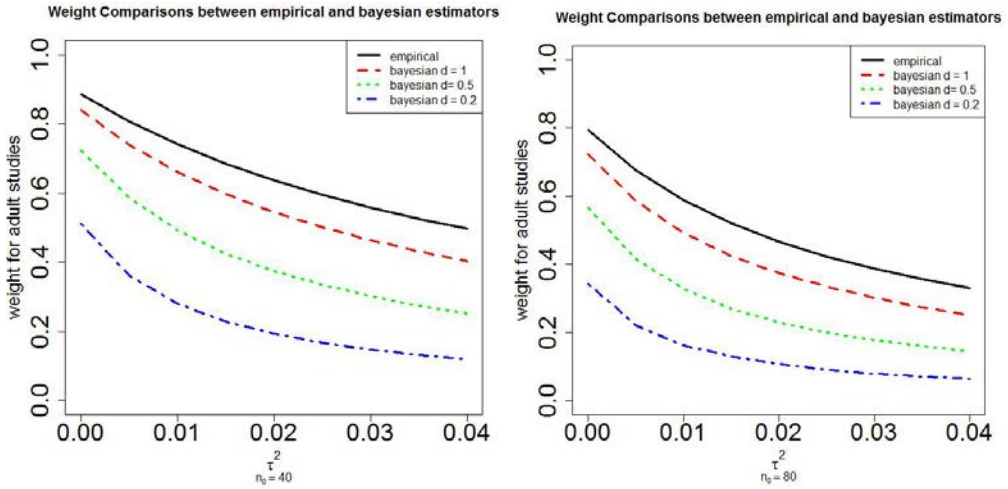

Figure 2: Weights for existing adult trials versus between-study variability τ^2 in empirical shrinkage estimator and Bayesian method with $d = 0.2, 0.5, 1$ for the pediatric sample size per arm $n_0 = 40$ or $n_0 = 80$.

5. Discussion

Conducting well-powered conventional efficacy and safety trials in pediatric patients is known to be challenging. Using the typical classical study design could require a sizeable number of patients which is not practical in some therapeutic areas, including T2DM.

Here we have considered and compared the performance of three methods for incorporating information from adult data in designing a Phase 3 pediatric trial that lead to a reduced sample size: J-S fixed effect estimates, empirical shrinkage estimates and Bayesian methods using a discount factor determined by the amount of adult data. If the observed treatment effect from existing adult trials is similar to each other, the J-S

shrinkage fixed-effect estimation can substantially reduce the required sample size for a pediatric trial. The fixed-effect method is commonly used, but a random effects model may be considered if generalization is necessary.

Assuming similar PK, PD and safety effects in adult and pediatric populations, an alternative approach to that explored here is to consider the concept of consistency with the treatment effect observed in the adult population instead of aiming to show a significant treatment difference in the pediatric patients. This concept was originally proposed in the MHLW guidance by PMDA for multi-regional clinical trials [20]. Multiple methodologies including sample size considerations have been developed for showing consistency of treatment effects. Additional work on the application of the consistency concept to the design and analysis of Phase 3 pediatric trials is ongoing.

References

1. J. G. Ibrahim and M. H. Chen. 2000. Power prior distributions for regression models. *Statistical Science*, 15(1), 46-60.

2. B. P. Hobbs, B. P. Carlin, S. J. Mandrekar, and D. J. Sargent. 2011. Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical. *Biometrics*, 67(3), 1047-1056.

3. B. P. Hobbs, D. J. Sargent, and B. P. Carlin. 2012. Commensurate priors for incorporating historical information in clinical trials using general and generalized linear models. *Bayesian Analysis*, 7(3), 639-674.

4. ICH Topic E11 Clinical Investigation of Medicinal Products in the Pediatric Population. Geneva, International Conference on Harmonisation, July 2000. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073143.pdf.

5. FDA draft guidance July 2013. Guidance for industry Pediatric study plans: content of and process for submitting initial pediatric study plans and amended pediatric study plans.

6. European Medicines Agency. Guideline on clinical trials in small populations. CHMP/EWP/83561/2005. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003615.pdf.

7. European Medicines Agency June 2012. Concept paper on extrapolation of efficacy and safety in medicine development (draft).

8. Ethical considerations for clinical trials on medicinal products conducted with the paediatric population. 2008. Available at: http://ec.europa.eu/health/files/eudralex/vol-10/ethical_considerations_en.pdf.

9. European Medicines Agency April 2013. Report of the workshop on paediatric investigation plans in type 2 diabetes mellitus, European Medicines agency, London, 25 February 2013.

10. J. Dunne, W. J. Rodriguez, M. D. Murphy, B. N. Beasley, G. J. Burckart, J. D. Filie, L. L. Lewis, H. C. Sachs, P. H. Sheridan, P. Starke and L. P. Yao. 2011. Extrapolation of adult data and other data in pediatric drug-development programs. Pediatrics, 128(5), 1242-1249.

11. E. Manolis and G. Pons, EMA. 2009. Proposals for model-based paediatric medicinal development within the current European Union regulatory framework. British Journal of Clinical Pharmacology, 68(4), 493-501.

12. J. Zisowsky, A. Krause and J. Dingemanse. 2010. Drug development for pediatric populations: regulatory aspects. Pharmaceutics, 2(4), 364-388.

13. P. R. Jadhav, J Zhang and J. V. S. Gobburu. 2009. Leveraging prior quantitative knowledge in guiding pediatric drug development: a case study. Pharmaceutical Statistics, 8, 216-224.

14. S. J. Pocock 1976. The combination of randomized and historical controls in clinical trials. *Journal of Chronic Diseases*, 29, 175-188.

15. B. Neuenschwander, G. Capkun-Niggli, M. Branson, and D. J. Spiegelhalter. 2010. Summarizing historical information on controls in clinical trials. *Clinical Trials,* 7(1), 5-18.

16. K. Viele, S. Berry, B. Neuenschwander, B. Amzal, F. Chen, N. Enas, B. P. Hobbs, J. G. Ibrahim, N. Kinnersley, S. Lindborg, S. Micallef, S. Roychoudhury, and L. Thompson. 2014. Use of historical control data for assessing treatment effects in clinical trials. *Pharmaceutical Statistics*, 13(1), 41-54.

17. S. N. Goodman, and J. T. Sladky. 2005. A Bayesian approach to randomized controlled trials in children utilizing information from adults: the case of Guillain-Barre syndrome. *Clinical Trials*, 2(4), 305-310.

18. D. A. Schoenfeld, H. Zheng and D. M. Finkelstein. 2009. Bayesian design using adult data to augment pediatric trials. *Clinical Trials*, 6(4), 297-304.

19. H. Quan, X. Mao, J. Chen, J.W. Shih, P.S. Ouyang, J. Zhang, P.L. Zhao and B. Binkowitz. 2014. Multi-regional clinical trial design and consistency assessment of treatment effects. *Statistics in Medicine*, 33(13), 2191-2205.

20. Ministry of Health, Labour and Welfare of Japan September 2007. Basic principles on Global Clinical Trials, September 28, 2007. http://www.pmda.go.jp/kijunsakusei/file/guideline/new_drug /GlobalClinicalTrials_en.pdf.