Received: 7 June 2023 | Revised: 30 August 2023 | Accepted: 10 September 2023

DOI: 10.1002/pst.2345

TUTORIAL

WILEY

Conditional power and information fraction calculations at an interim analysis for random coefficient models

Sandra A. Lewis[1] | **Kevin J. Carroll**[2] | **Todd DeVries**[1] | **Jonathan Barratt**[3]

[1]Chinook Therapeutics, Novartis Company, Seattle, Washington, USA

[2]KJC Statistics Ltd, Cheadle Hulme, Cheadle, UK

[3]Department of Cardiovascular Sciences, University of Leicester, Leicester, UK

Correspondence
Kevin J. Carroll, KJC Statistics Ltd, Cheadle Hulme, Cheadle, UK.
Email: kevin.carroll@kjcstatistics.com

Abstract

Random coefficient (RC) models are commonly used in clinical trials to estimate the rate of change over time in longitudinal data. Trials utilizing a surrogate endpoint for accelerated approval with a confirmatory longitudinal endpoint to show clinical benefit is a strategy implemented across various therapeutic areas, including immunoglobulin A nephropathy. Understanding conditional power (CP) and information fraction calculations of RC models may help in the design of clinical trials as well as provide support for the confirmatory endpoint at the time of accelerated approval. This paper provides calculation methods, with practical examples, for determining CP at an interim analysis for a RC model with longitudinal data, such as estimated glomerular filtration rate (eGFR) assessments to measure rate of change in eGFR slope.

KEYWORDS

conditional power, IgA nephropathy, information fraction, interim analysis, longitudinal data, random coefficients model

1 | INTRODUCTION

The recent meta-analyses in immunoglobulin A nephropathy (IgAN) trials by Inker et al.,[1] showing the association of treatment effects in change in urine protein and glomerular filtration rate (GFR) slope, have paved the way in IgAN clinical trials for the use of an early surrogate endpoint for possible accelerated approval with verification of a longer term confirmatory endpoint. As a result, an upsurge of IgAN clinical trials have been initiated, typically designed with urine protein creatinine ratio (UPCR) as the surrogate endpoint evaluated at an interim timepoint, and the rate of change in estimated glomerular filtration rate (eGFR) as the confirmatory longer term endpoint evaluated at the end of the study. Estimates of conditional power (CP) and information fraction for the confirmatory endpoint at the time of the interim analysis may provide further confidence that the confirmatory endpoint will be significant in the final analysis. In addition, the use of random coefficient (RC) models to estimate rates of change in longitudinal data is a common approach in clinical trials across various therapeutic areas, including chronic lung diseases, chronic kidney diseases, diabetes, and neurodegenerative disorders. Understanding CP and information fraction calculations may help in the design of clinical trials as well as offer support to a possible filing for accelerated approval. This paper will describe the calculation methods for determining CP at an interim analysis for a RC model with longitudinal data, such as eGFR slope.

2 | BACKGROUND

To investigate the relationship between a response variable and time, a well-established approach is the RCs model (Littell, 2006).[2] Denoting the dependent variable as y_{ij}, the general form of the RC model is as follows:

$$y_{ij} = (\alpha + a_i) + (\beta + b_i)x_{ij} + e_{ij}, \tag{1}$$

where

- y_{ij} is the dependent variable value for subject i at assessment time j;
- x_{ij} is the time of the jth assessment for subject i;
- α and β are the overall fixed effects of intercept (α) and time (β);
- a_i and b_i are the random intercept and slope effect associated with subject i, with a_i *iid* distributed $N\left(0,\sigma_a^2\right)$ and b_i *iid* distributed $N\left(0,\sigma_b^2\right)$;
- e_{ij} is the residual error for subject i at time j, with e_{ij} *iid* distributed $N\left(0,\sigma_e^2\right)$.

The RC model therefore effectively fits a simple regression line to each subject, with the resulting intercept and slope estimates then combined across subjects to provide overall estimates of the intercept, $\widehat{\alpha}$, and slope, $\widehat{\beta}$.

In this paper, we will assume a randomized clinical trial with two treatment groups, drug (D) and placebo (P), where $i = 1\,to\,n$ subjects will be randomized to each treatment group for a total of $N = 2n$ subjects. Each subject is scheduled to have a set of $j = 1$ to v longitudinal values (e.g., eGFR assessments), y_{ij}, measured at times x_{ij} over a follow-up period of F months. The slope estimate for subject i is denoted β_i with variance $V(\beta_i) = \sigma_e^2/Sxx_i$, where σ_e^2 is the residual error around regression, and Sxx_i is the sum of the squared differences of measurement times minus the mean time. Thus Sxx_i is calculated as:

$$Sxx_i = \sum_{j=1}^{v}\left(x_{ij} - \overline{x}_{i\cdot}\right)^2 \text{ and } \overline{x}_{i\cdot} = \frac{\sum_{j=1}^{v} x_{ij}}{v}. \tag{2}$$

Assuming a balanced design with no dropouts and allowing the β_i estimates to vary between subjects with $\beta_i \sim N\left(\beta,\sigma_b^2\right)$, then $\overline{\beta} = \left(\sum_{i=1}^{n}\beta_i/n\right)$ with variance $V\left(\overline{\beta}\right)$, where:

$$V\left(\overline{\beta}\right) = 1/\sum_{i=1}^{n}\left\{\left(\frac{\sigma_e^2}{Sxx_i} + \sigma_b^2\right)^{-1}\right\}. \tag{3}$$

If all subjects have the same number of eGFR values, v, spaced at m monthly intervals, then $Sxx_i = Sxx$ for all i, where

$$Sxx = \sum_{j=1}^{v}\left(x_{ij} - \overline{x}_{i\cdot}\right)^2 = (m - \overline{m})^2 + (2m - \overline{m})^2 + \ldots + (vm - \overline{m})^2 \tag{4}$$

and where

$$\overline{m} = \frac{\sum_{j=1}^{v} x_{ij}}{v} = \frac{\sum_{j=1}^{v} m \cdot j}{v} = \frac{m(v+1)}{2} \text{ and } vm = F. \tag{5}$$

Then Sxx reduces to

$$Sxx = \frac{m^2 v(v^2 - 1)}{12}. \tag{6}$$

Therefore, a generalized formula for $V(\overline{\beta})$ can be written as

$$V(\overline{\beta}) = 1/\sum_{i=1}^{n}\left\{\left(\frac{\sigma_e^2}{Sxx_i}+\sigma_b^2\right)^{-1}\right\} = \frac{2}{N}\left(\frac{\sigma_e^2}{Sxx}+\sigma_b^2\right) = \frac{2}{N}\left(\frac{\sigma_e^2}{\left\{\frac{m^2v(v^2-1)}{12}\right\}}+\sigma_b^2\right). \tag{7}$$

3 | FISHERS INFORMATION

The preceding section provides the basis for computing the total amount of Fishers Information, I, expected at the end of the trial with respect to the estimated difference in slopes between drug and placebo as

$$I = \left\{V(\overline{\beta}_D)+V(\overline{\beta}_P)\right\}^{-1} = \left\{2V(\overline{\beta})\right\}^{-1} = \frac{N}{4}\left(\frac{\sigma_e^2}{\left\{\frac{m^2v(v^2-1)}{12}\right\}}+\sigma_b^2\right)^{-1}, \tag{8}$$

where $V(\overline{\beta}_D)$ and $V(\overline{\beta}_P)$ are the variances of the estimated slope values for the drug and placebo groups, respectively (Ly, 2017).[3]

3.1 | Information at the interim

As previously indicated, N subjects will be randomized in the trial. Each subject will be followed for a total duration of F months for the confirmatory endpoint. It is planned that the last subject will be randomized at time A, so that the trial will be completed with full information at time $A+F$. An interim takes place with a data cut at time B months, typically at the time of the primary surrogate endpoint analysis for clinical trials, with N_φ subjects randomized. Of interest is the mean follow-up time, $\overline{F}_\varphi$, for the N_φ subjects included in the interim. The $\overline{F}_\varphi$ can be calculated based on an approximated amount of information based on the average number of visits per subject at the interim or by the actual number of visits by subject at the interim.

3.2 | Approximate amount of information at the interim

To approximate the amount of information at the interim, the timing of the interim, B, in relation to F, total months a subject is planned to be followed for the confirmatory or final endpoint should be considered. If an interim is conducted at time B, where $(F<B<A)$ such that the interim is done before all subjects have been randomized and after some subjects have completed all eGFR assessments for the confirmatory endpoint (Figure 1), the mean follow-up time, $\overline{F}_\varphi$ can be calculated as follows:

$$\text{Mean follow-up} = \overline{F}_\varphi = \frac{F}{B}\left(B-\frac{F}{2}\right). \tag{9}$$

Hence the average number of visits per subject at the interim is given by, $v_\varphi = \overline{F}_\varphi/m$ and, therefore, the estimated variance of the mean slope at the interim is approximately,

$$V(\overline{\beta}_\varphi) = \frac{2}{N_\varphi}\left(\frac{\sigma_e^2}{Sxx_\varphi}+\sigma_b^2\right) = \frac{2}{N_\varphi}\left(\frac{\sigma_e^2}{\left\{\frac{m^2v_\varphi\left(v_\varphi^2-1\right)}{12}\right\}}+\sigma_b^2\right). \tag{10}$$

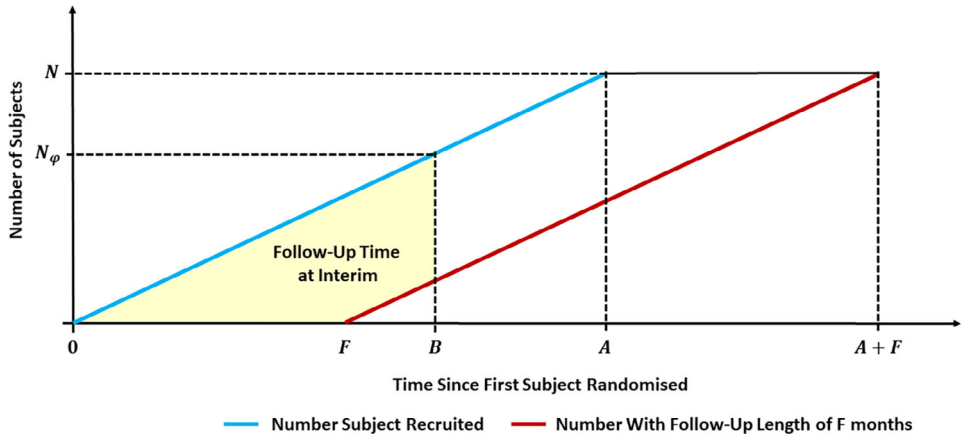

FIGURE 1 Timing and follow-up at an interim analysis conducted prior to completion of randomization but with some subjects having been followed for time *F*.

Thus, Fishers information at the interim is approximately

$$I_\varphi = \left\{V\left(\overline{\beta}_{\varphi D}\right) + V\left(\overline{\beta}_{\varphi P}\right)\right\}^{-1} = \left\{2V\left(\overline{\beta}_{\varphi}\right)\right\}^{-1} = \frac{N_\varphi}{4}\left(\frac{\sigma_e^2}{\left\{\frac{m^2 v_\varphi\left(v_\varphi^2-1\right)}{12}\right\}} + \sigma_b^2\right)^{-1}. \tag{11}$$

So that the fraction of information at the interim is approximately

$$\lambda_\varphi = \frac{I_\varphi}{I}. \tag{12}$$

For completeness, it is possible the interim is conducted at time $B \le F$ such that the interim is done before all subjects have been randomized, see Figure 2A; or at time $B > A$ such that the interim is done after all subjects have been randomized, see Figure 2B. For the former, mean follow-up time is $\overline{F}_\varphi = (B/2)$ and, for the latter, $\overline{F}_\varphi = (\mathrm{F}/\mathrm{A})(A - F/2) + ((B-A)/A)(F - (A-B)/2)$.

3.3 | Accurate information at the interim

To better compute the information at the interim we need to consider Sxx_i by subject. We know that $F = mv$; assume $B = mv + mr$ with $F < B < A$ as above. Then, as shown in Appendix 1, the sum of *Sxx* per arm is as follows:

$$\sum_i^n Sxx_i = \frac{(n/A)m^3 \cdot v \cdot (v^2-1) \cdot (v + 4r - 2)}{48}. \tag{13}$$

Hence, we can estimate the variance of the mean slope at the interim with n_φ subjects as

$$V\left(\overline{\beta}_\varphi\right) = 1/\sum_{i=1}^{n_\varphi}\left\{\left(\frac{\sigma_e^2}{Sxx_i} + \sigma_b^2\right)^{-1}\right\} \tag{14}$$

using the Sxx_i by subject values displayed in Appendix Table A1. If we assume a common slope per subject, so that $\sigma_b^2 = 0$, then the complexity reduces to,

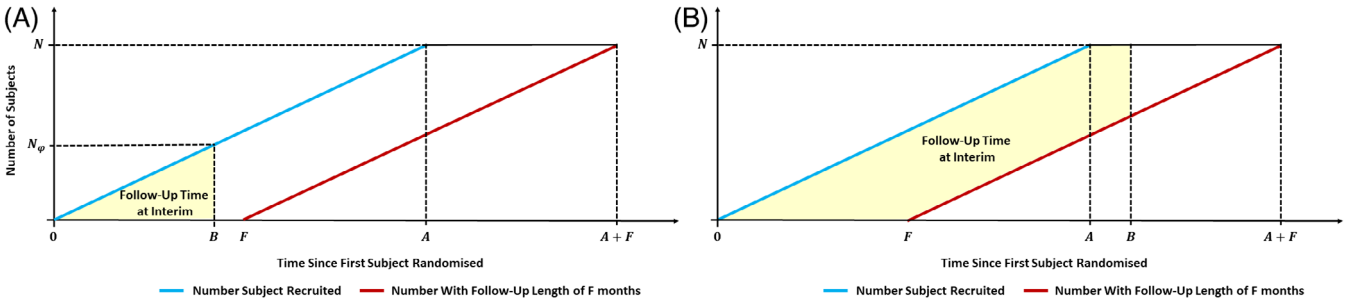

FIGURE 2 (A) Total follow-up for an interim conducted before time *F*. (B) Total follow-up for an interim conducted after time *A*.

$$V(\overline{\beta}_\varphi) = 1/\sum_{i=1}^{n_\varphi}\left\{\left(\frac{\sigma_e^2}{Sxx_i}\right)^{-1}\right\} = \frac{\sigma_e^2}{\sum_{i=1}^{n_\varphi} Sxx_i} = \frac{\sigma_e^2}{\left\{\frac{(n/A)\cdot m^3\cdot v\cdot(v^2-1)\cdot(v+4r-2)}{48}\right\}}. \tag{15}$$

Thus, Fishers information at the interim is

$$I_\varphi = \left\{V(\overline{\beta}_{\varphi D}) + V(\overline{\beta}_{\varphi P})\right\}^{-1} = \left\{2V(\overline{\beta}_\varphi)\right\}^{-1} = \frac{N}{4}\left(\frac{\sigma_e^2}{\left\{\frac{(1/A)\cdot m^3\cdot v\cdot(v^2-1)\cdot(v+4r-2)}{48}\right\}}\right)^{-1}. \tag{16}$$

So that the fraction of information at the interim is

$$\lambda_\varphi = \frac{I_\varphi}{I} = \frac{m(v+4r-2)}{4A} \tag{17}$$

and the information post interim is

$$I^* = I - I_{\varphi.} \tag{18}$$

4 | CONDITIONAL AND PREDICTIVE POWER

In a trial sized at the one-sided α level with power $1-p$, CP for the trial computed at the interim as outlined by Jennison and Turnbull[4] is given by

$$\mathrm{CP} = 1 - \Phi\left\{\frac{z_\alpha\sqrt{\lambda_\varphi} - z_\varphi}{\sqrt{\lambda_\varphi(1-\lambda_\varphi)}}\right\}, \tag{19}$$

where $z_\varphi = \Delta_\varphi\sqrt{I_\varphi}$ and Δ_φ are the observed difference in mean slope at the interim.

An alternative, and arguably better, measure of CP is to compute the expected CP, averaging across the distribution of treatment effect size implied by the interim data. This measure is referred to as predictive power (PP).[4] The PP for the trial computed at the interim is given by,

$$\mathrm{PP} = 1 - \Phi\left\{\frac{z_\alpha\sqrt{\lambda_\varphi} - z_\varphi}{\sqrt{1-\lambda_\varphi}}\right\}. \tag{20}$$

5 | EXAMPLE CALCULATIONS AT THE DESIGN STAGE

Assume the confirmatory endpoint is eGFR slope and the desire is to test the hypothesis that the annualized difference between drug and placebo is $\Delta = -1.695$ mL/min/1.73 m^2/year (i.e., 0.1413 mL/min/1.73 m^2 per month) with a one-

sided $\alpha = 0.025$ and power $= 1-\beta = 0.80$. We will assume a common slope per subject so that $\sigma_b^2 = 0$ and a residual error $\sigma_e^2 = 6^2$ mL/min/1.73 m^2. Follow-up of all subjects is planned to be $F = 24$ months with quarterly visits such that $m = 3$ monthly eGFR assessments, hence $v = 24/3 = 8$ observations per subject.

In a typical Phase III trial, the hypothesis to be tested is usually,

$$H_0 : \theta_{\text{TRUE}} = 0 \,\text{vs}\, H_1 : \theta_{\text{TRUE}} \neq 0, \tag{21}$$

where θ_{TRUE} denotes the true treatment effect. For the purposes of sizing, a positive effect $\theta_{\text{TRUE}} = \theta$ (>0) is usually assumed under the alternative. Let x be a sufficient statistic for θ with distribution $f(x|\theta) \sim N(\theta, \nu^2)$, where $N(.,.)$ represents the normal distribution. Trial size is then governed by the one-sided Type I and Type II errors, α and β, and the need to deliver the required information content, $1/\nu^2 = (z_\alpha + z_\beta)^2/\theta^2$, where $z_u = \Phi^{-1}(1-u)$ and $\Phi^{-1}(.)$ represents the inverse standard Normal distribution function. The null hypothesis is rejected in favor of the alternative when $x > z_\alpha \nu$.

In the case of eGFR slope analysis over time, we have that $\theta = \beta_D - \beta_P$ and $\nu^2 = (4/N)\left(\left(\sigma_e^2/Sxx\right) + \sigma_b^2\right)$, thus we have

$$\left\{\frac{4}{N}\left(\frac{\sigma_e^2}{Sxx} + \sigma_b^2\right)\right\}^{-1} = (z_\alpha + z_\beta)^2/(\beta_D - \beta_P)^2 \tag{22}$$

And hence,

$$N = \frac{4(z_\alpha + z_p)^2}{(\beta_D - \beta_P)^2}\left(\frac{\sigma_e^2}{Sxx} + \sigma_b^2\right), \tag{23}$$

which is consistent with Zhao and Edland (2020).[5] Furthermore, substituting Sxx from Equation (6) then,

$$N = \frac{4(z_\alpha + z_p)^2}{(\beta_D - \beta_P)^2}\left(\frac{\sigma_e^2}{\left\{\frac{m^2 v(v^2-1)}{12}\right\}} + \sigma_b^2\right), \tag{24}$$

$$N = (1.96 + 1.28)^2 \frac{4}{(0.1413)^2}\left(\frac{6^2}{\left\{\frac{3^2 \cdot 8 \cdot (8^2-1)}{12}\right\}} + 0\right) = 200. \tag{25}$$

Thus, it is planned that $N = 200$ subjects will be randomized over $A = 48$ months so the recruitment rate is $200/48 =$ approximately four subjects per month. An interim is planned at time $B = 30$ months when $N_\varphi = B \cdot N/A = 0.625N = 125$ subjects have been randomized, meaning that $r = 2$. At the final analysis, using Equation (8),

$$I = \frac{N\left\{\frac{m^2 v(v^2-1)}{12}\right\}}{4\sigma_e^2} = \frac{200\left\{\frac{3^2 \cdot 8 \cdot (8^2-1)}{12}\right\}}{4 \cdot 6^2} = 525. \tag{26}$$

At the interim using Equation (11),

$$I_\varphi = \frac{200}{4}\left(\frac{6^2}{\left\{\frac{(1/48) \cdot 3^3 \cdot 8 \cdot (8^2-1) \cdot (8+4 \cdot 2-2)}{48}\right\}}\right)^{-1} = 114.8438. \tag{27}$$

So that the fraction of information at the interim is

$$\lambda_{\varphi} = \frac{I_{\varphi}}{I} = \frac{m(v+4r-2)}{4A} = \frac{3\cdot(8+4\cdot 2-2)}{4\cdot 48} = 0.2188 \tag{28}$$

Note, also that

$$\lambda_{\varphi} = \frac{I_{\varphi}}{I} = \frac{114.8438}{525} = 0.2188. \tag{29}$$

If the observed result at the interim is a difference in mean slopes of $\Delta = 0.12\,\text{mL/min/1.73 m}^2$ per month (1.44 mL/min/1.73 m^2 per year) and recall $z_{\varphi} = \Delta_{\varphi}\sqrt{I_{\varphi}}$, then

$$\text{CP} = 1 - \Phi\left\{\frac{z_{\alpha}\sqrt{\lambda_{\varphi}} - z_{\varphi}}{\sqrt{\lambda_{\varphi}(1-\lambda_{\varphi})}}\right\} = 1 - \Phi\left\{\frac{1.96\sqrt{0.2188} - 0.12\sqrt{114.8438}}{\sqrt{0.2188(1-0.2188)}}\right\} = 0.814. \tag{30}$$

Appendix 2 contains SAS software code and two Excel files that provide the basis for computing the interim information, the interim information fraction, and conditional and PP. The assumption of no subject to subject variability in random slopes, $\sigma_b^2 = 0$, to simplify the above equations and example calculations may be unrealistic and result in anti-conservative estimates in these calculations. In IgAN nephropathy, analyses from the Leicester IgA nephropathy patient registries estimated σ_b^2 ranging from 50 to 60 mL/min/month. When designing studies in other therapeutic areas, σ_b^2 may be unknown and understanding the impact of ranges of σ_b^2 on design elements, such as sample size and power, may be helpful. The two Excel files included in Appendix 2 can be used to evaluate the impact of varying assumptions. Excel file 1 can be used for specialized cases for equally spaced visits and no subject to subject variability in random slopes. Excel file 2 can be used for more generalized cases, allowing for subject to subject variability in random slopes and unequally spaced observations.

6 | EXAMPLE CALCULATIONS AT AN INTERIM ANALYSIS

Calculations of conditional and predictive power for the difference in slopes at the interim analysis can be easily performed using parameter estimates generated from the RCs model using SAS software. The calculations require the difference in slopes, estimated from the model, and the information fraction, $\lambda_{\varphi} = I_{\varphi}/I$, which can be calculated from other parameter estimates generated. The actual information at the interim analysis, I_{φ}, can be calculated as the reciprocal of the square of the standard error for the difference in slopes from the RCs analysis. The estimated information at the final analysis, $I = \{V(\overline{\beta}_D) + V(\overline{\beta}_P)\}$, can be calculated using the estimated variance of the slopes σ_b^2, and the estimated residual error variance, σ_e^2, from the RCs model at the interim analysis and then computing the following for each treatment group $V(\overline{\beta}) = 1/\sum_{i=1}^{n}\left\{(\sigma_e^2/Sxx_i) + \sigma_b^{2^{-1}}\right\}$, where Sxx_i is the estimated Sxx for each subject at the final analysis (computed assuming no additional dropouts).

Appendix 2 provides SAS software code for calculating the conditional and predictive power for the difference in slopes at an interim analysis using estimates from a RCs model based on a small example dataset. The example dataset has 10 subjects for each of 2 treatment groups with eGFR assessments at timepoints of 0, 2, 4, 6, 12, 18, and 24 months. At the time of the interim analysis 10 subjects had completed the study, 2 subjects had dropped out of the study, and 8 subjects were ongoing. Subjects 12 and 15 each had an intermittent missing value. In this example dataset, the information at the interim analysis, $I_{\varphi} = 15.2606$, the estimated information at the final analysis $I = 19.8437$, resulting in the information fraction at the interim analysis, $\lambda_{\varphi} = I_{\varphi}/I = 0.769$. The conditional and predictive power for the example dataset is 0.96 and 0.94, respectively, assuming a significance test at the one-sided 0.025 level based on the z-statistic for the difference in slopes from the RCs model of 2.4673.

7 | DISCUSSION

The equations in this paper have been provided in order to have useful formulae, reduced to simplified forms, for planning and designing clinical trials with RC models to detect differences in mean rates of change between treatment

groups over time as well as providing insight for the determination of CP at a planned interim analysis. Broadly, these formulae can be used at any givern interim to evaluate the likelihood of success of clinical trials that implement RC models for the analysis of longitudinal data. While presented in terms of IgAN clinical trials where accelerated approval is often sought based on UPCR, with eGFR slope as the confirmatory endpoint, the methodologies presented in this paper can be applied across multiple therapeutic areas in which rates of change over time are of interest. The methodology presented is easily adapted to accomodate unequal allocation of sample sizes between treatment arms, subject attrition, and unequal variances across groups.

CONFLICT OF INTEREST STATEMENT

Sandra A. Lewis is an employee Chinook Therapeutics, Inc. Todd Devries is an employee of Chinook Therapeutics, Inc. Kevin J. Carroll is a consultant for Chinook Therapeutics, Inc. Jonathan Barratt is a consultant and an advisory board member for Chinook Therapeutics, Inc.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Sandra A. Lewis https://orcid.org/0009-0007-4519-8689
Kevin J. Carroll https://orcid.org/0000-0002-3500-9692
Todd DeVries https://orcid.org/0009-0006-5645-9373
Jonathan Barratt https://orcid.org/0000-0002-9063-7229

REFERENCES

1. Inker LA, Heerspink HJL, Tighiouart H, et al. Association of treatment effects on early change in urine protein and treatment effects on GFR slope in IgA nephropathy: an individual participant meta-analysis. *Am J Kidney Dis*. 2021;78(3):340-349.e1.
2. Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Oliver S. *SAS for Mixed Models*. 2nd ed. SAS Publishing; 2006.
3. Ly A, Marsman M, Verhagen J, Grasman R, Wagenmakers E. A tutorial on Fisher information. *J Math Psychol*. 2017;80:40-55.
4. Jennison C, Turnbull BW. *Group Sequential Methods with Applications to Clinical Trials*. 1st ed. Chapman and Hall/CRC; 2000.
5. Zhao Y, Edland SD. Power formulas for mixed effects models with random slope and intercept comparing rate of change across groups. *Int J Biostat*. 2021;18(1):173-182.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lewis SA, Carroll KJ, DeVries T, Barratt J. Conditional power and information fraction calculations at an interim analysis for random coefficient models. *Pharmaceutical Statistics*. 2023;1-8. doi:10.1002/pst.2345