Mixed-effects models for slope-based endpoints in clinical trials of chronic kidney disease

Supplementary Materials

Edward Vonesh, Hocine Tighiouart, Jian Ying, Hiddo Heerspink, Julia Lewis, Natalie Staplin, Lesley Inker and Tom Greene

WEB Appendix A: Sample size/power calculations using SAS macro *%GFR_Slope_Power*

This appendix provides a description of a SAS macro, *%GFR_Slope_Power*, presented in the Section 6. The SAS macro was developed for the purpose of determining sample size/power estimates for comparing slope-based endpoints in randomized controlled trials (RCT's) of patients with chronic kidney disease (CKD). The power calculations are based on the linear spline-mixed-effects model described in Section 2.1. Assuming just two treatment groups, a control group and a treated group, the macro computes power for detecting differences between the two treatment groups with respect to the acute slopes, $\beta_{1t} - \beta_{1c}$, the delta or change slopes, $\beta_{2t} - \beta_{2c}$, the chronic slopes, $\beta_{3t} - \beta_{3c}$, and the total slopes evaluated at t=2, 3 and 4 years of follow-up, $\beta_{4t}(t) - \beta_{4c}(t)$. The user is required to input values of the key model parameters using the macro variable naming convention shown in Table A.1.

Table A.1. Key parameter inputs into *%GFR_Slope_Power* and their SAS macro variable names

Key Input Parameters	Description of Population-based Parameters	SAS Macro Variable Name
β_{0c}	Control group intercept (mL/min/1.73m^2)	beta0[1]
β_{0t}	Treated group intercept (mL/min/1.73m^2)	beta0
β_{1c}	Control group acute slope (mL/min/1.73m^2/month)	beta1c
β_{1t}	Treated group acute slope (mL/min/1.73m^2/month)	beta1t
β_{2c}	Control group delta (change) slope (mL/min/1.73m^2/month)	beta2c
β_{2t}	Treated group delta (change) slope (mL/min/1.73m^2/month)	beta2t
σ^2	The within-subject variance component for $Var(\varepsilon_{ij}) = \sigma^2 \mu_{ij}^2(\beta_i)^\theta$	Var_e
ψ_{11}	The variance of the control group random intercept effect, u_{0i}	Var_u0
ψ_{22}	The variance of the control group random acute slope effect, u_{1i}	Var_u1
ψ_{33}	The variance of the control group random change slope effect, u_{2i}	Var_u2
ψ_{12}	The covariance between u_{0i} and u_{1i}	Cov_u0u1
ψ_{13}	The covariance between u_{0i} and u_{2i}	Cov_u0u2
ψ_{23}	The covariance between u_{1i} and u_{2i}	Cov_u1u2
θ	The Power-of-Mean parameter for $Var(\varepsilon_{ij}) = \sigma^2 \mu_{ij}^2(\beta_i)^\theta$	theta
κ	The treatment-dependent constant for Ψ_k (see Table 1 of paper)	kappa
t^*	The fixed-point knot at which the acute slope phase ends	knot
N	A proposed sample size per treatment group	N
α	The overall two-sided type I error	Alpha

[1] It is assumed that the two treatment groups have a common intercept, $\beta_{0c} = \beta_{0t} = \beta_0$, as would be expected based on the principles of randomization

The user can specify the input parameters to meet the specific needs and assumptions required for a given application. For example, results from a phase 2 trial can be used to provide input into what value a fixed-point knot t^* should be set equal to in the presence of an acute treatment effect. Likewise, evidence of homogeneous within- and/or between-subject variability can be incorporated into the power calculations by simply setting the macro variables THETA=0 and/or KAPPA=0.

The program uses calculations based on an approximation to the power of an F-test (or, equivalently, a two-sided t-test) described by Littell et al. and references therein.[30] Based on a given sample size of *N* subjects per treatment group along with hypothesized mean response profiles for the two treatment groups at *p* specified measurement occasions or time points (from the input model parameters given in Table A.1), one can create a single "dummy" dataset with replicates of the treatment-dependent mean response profiles taking the place of actual data values. By holding the within- and between-subject variance and covariance parameters fixed, one can then "fit" the linear spline mixed-effects model to the "dummy" data using PROC MIXED and test the null hypothesis $H_0 : \boldsymbol{L}'\beta = \boldsymbol{0}$ vs the alternative $H_1 : \boldsymbol{L}'\beta \neq \boldsymbol{0}$ with CONTRAST or ESTIMATE statements (with the ESTIMATE statement being preferred whenever $rank(\boldsymbol{L}) = 1$). One can then estimate power as

$$\text{Power} = \Pr(F_{NDF,DDF,\delta} > F_{NDF,DDF,0,\alpha})$$

where $NDF = rank(\boldsymbol{L})$ is the numerator degrees of freedom from the assumed estimable CONTRAST or ESTIMATE statement, DDF is the denominator degrees of freedom using the default value determined by either PROC MIXED or PROC NLMIXED (the latter of which is 2*N-3* where 2*N* is the total number of subjects and 3 is the number of random effects), $\delta = (\boldsymbol{L}'\boldsymbol{\beta})'\{\boldsymbol{L}'(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{L}\}(\boldsymbol{L}'\boldsymbol{\beta})$ is the non-centrality parameter of the F-distribution with $\boldsymbol{X}$ being the design matrix and $\boldsymbol{\Sigma}$ the marginal covariance matrix, and $F_{NDF,DDF,0,\alpha}$ is the critical F value for testing the null hypothesis at a type I error of α. Under the linear spline mixed-effects model (2) with SS-POM variances $\sigma^2(\mu_{ij}^2(\boldsymbol{\beta}_i))^\theta$, the overall marginal covariance matrix will be $\boldsymbol{\Sigma} = \boldsymbol{Z}\boldsymbol{\Psi}\boldsymbol{Z}' + \boldsymbol{R}$ where $\boldsymbol{R} = E_{\boldsymbol{b}}(\boldsymbol{R}(\boldsymbol{b}))$ and $\boldsymbol{R}(\boldsymbol{b})$ is a block diagonal matrix with diagonal matrices $\boldsymbol{R}_i(\boldsymbol{b}_i)$ having values of $\sigma^2(\mu_{ij}^2(\boldsymbol{\beta}_i))^\theta$ along their diagonals (see Section 6). One can approximate $\boldsymbol{R}$ using a first-order Taylor series expansion of the SS-POM variances $\sigma^2(\mu_{ij}^2(\boldsymbol{\beta}_i))^\theta$ around $\boldsymbol{b}_i = \boldsymbol{0}$. Doing so leads to the use of a population-average power-of-mean (PA-POM) variance structure in $\boldsymbol{R}$. To fit this structure using PROC MIXED, one must first form the weights $w_{ij} = 1/(\mu_{ij}^2(\boldsymbol{\beta}))^\theta$ where $\mu_{ij}(\boldsymbol{\beta}) = E(Y_{ij})$ is the marginal or population-average mean response based on the linear spline mixed-effects model defined by either (1) or (2) of Section 2.1. One can then simply use the WEIGHT statement of PROC MIXED together with the default R-side within-subject variance-covariance structure $\sigma^2\boldsymbol{I}$ to form an R-side diagonal covariance matrix $\boldsymbol{R}_i = \sigma^2\boldsymbol{W}_i^{-1}$ where $\boldsymbol{W}_i$ is a diagonal matrix with values $w_{ij} = 1/(\mu_{ij}^2(\boldsymbol{\beta}))^\theta$ on the diagonal and 0 values on the off-diagonals. We will refer to the weights $w_{ij} = 1/(\mu_{ij}^2(\boldsymbol{\beta}))^\theta$ as population-averaged power-of-mean (PA-POM) weights to distinguish them from the subject-specific power-of-mean (SS-POM) weights $w_{ij} = 1/(\mu_{ij}^2(\boldsymbol{\beta}_i))^\theta$ assumed in Section 2.1.

Table A.2 describes other key SAS macro variables along with their default values that the user can specify when running the macro (other macro variables controlling the printout etc. are described in the program itself). One important macro variable not shown in Table A.2 is the variable DIFF. This SAS macro variable allows the user to override whatever value one assigns to the delta (change) slope of the treatment group and, instead, allows the user to specify incremental differences between the treatment group and control group chronic slopes (NOTE: as input all slopes are expressed in mL/min/1.73m^2/month corresponding to time measured in months while the output expresses the slopes in mL/min/1.73m^2/year). This, in turn, induces incremental differences between the treatment and control group delta slopes and between the treatment and control

group total slopes. The values assigned to the acute slopes for the treatment and control groups remain fixed as does the value assigned to the control group delta slope. A typical specification might look like DIFF = 0.10 to 0.30 by 0.05 which specifies that the treatment effect on the chronic slope, $\Delta = \beta_{3t} - \beta_{3c}$, is defined by DIFF values ranging from 0.10 to 0.30 mL/min/1.73m2/month.

Table A.2. Other Key SAS Macro Variables as Inputs into *%GFR_Slope_Power*

SAS Macro Variables	Description	Default Value
Months=	Defines the occasions (months) when eGFR is to be measured. For the IDNT study we set this to be 0,3,6,12,18,24,30,36,42,48 and 54 months	See Program
POM=	Defines whether one uses a PA or EM power-of-mean (POM) structure	PA
REPS=	Defines number of replicated random effects to be used with POM=EM	500
LowerGFR=	A lower limit exclusion criterion for eGFR at entry (mL/min/1.73m^2)	15
UpperGFR=	An upper limit exclusion criterion for eGFR at entry (mL/min/1.73m^2)	120
Screen=	Defines an additional number of patients per treatment group (beyond that specified by the N= option) to be screened at entry to satisfy the above exclusion criteria. This is accomplished by generating a one-time sample of multivariate normal random effects based on the model input parameters (Table A.1) and selecting the first 2N subjects whose predicted baseline eGFR values fall within the boundaries defined by the lower and upper limit exclusion criteria.	100
Accrual=	Defines an accrual period (months)	0
Followup=	Defines a maximum follow-up period after accrual (months)	60
Dropout=	Defines an annual proportion of dropout which is then converted into a monthly dropout rate assuming non-informative dropout	0.00
InfCensor=	Defines whether we censor patients based on whether their subject-specific predicted eGFR falls <15 mL/min/1.73m^2 at a given time point	NO
DDF=	Defines what default denominator degrees of freedom one uses for power calculations (Choice: PROC MIXED or PROC NLMIXED)	NLMIXED

It should be noted that when one uses the default settings with POM=PA, LowerGFR=15, UpperGFR=120, Accrual=0, Followup=60, Dropout=0.00 and InfCensor=NO, the macro program will compute power for fixed sample sizes based on CONTRAST and ESTIMATE statements for a balanced complete study design utilizing whatever measurement occasions (time points) the user assigns to the macro variable MONTHS= but excluding any measurement occasions that exceed the default 60 month maximum follow-up. Using these default settings may lead to overly optimistic power estimates based on the fact that the default values ignore potential study design imbalances one would expect due to staggered patient entry, random patient dropout and informative censoring due to ESKD (defined within the program to occur when a subject's predicted mean eGFR falls below 15 mL/min/1.73m^2). To illustrate this and other aspects of the macro, we ran four different cases of the macro. In Case 1 (code shown below), we use the above default values in combination with model parameters taken from the IDNT study under SP model 1 (Table 5).

```
/*=============================================================================
Case 1: Assumes no censoring due to staggered entry or dropout nor informative censoring
due to ESKD. Slope and variance-covariance estimates are based on SP Model 1 results
(Table 5) from the IDNT study. Power is based on a PA-POM variance structure and study
design considerations like measurement times taken from the clinical trial design. We
employed a screening of an additional 100 subjects per treatment group (200 additional
subjects in total) to ensure subjects enrolled do not have starting predicted eGFR values
that fall below 15 mL/min/1.73m2 or above 120 mL/min/1.73m2.
=============================================================================*/
```

```
%GFR_Slope_Power(Data=CKD, knot=4, beta0=50.00, beta1c=-1.012, beta1t=-1.213, beta2c=0.593,
 beta2t=0.878, Var_e=3.501, Var_u0=325.825, Var_u1=1.088, Var_u2=0.876, Cov_u0u1=2.260,
 Cov_u0u2=-3.032, Cov_u1u2=-0.935, kappa=-0.061, theta=0.917, DIFF=,
 Months=%str(0,3,6,12,18,24,30,36,42,48,54), POM=PA, LowerGFR=15, UpperGFR=120,
 Screen=100, Accrual=0 , Followup=60, Dropout=0.00, InfCensor=NO,  Alpha=0.05, DDF=NLMIXED,
 N=300 to 600 by 100, Seed=95738, Close=close, Check=NO, Print=YES, Table_F=A.3,
 Table_t=A.4, Directory=d:\Stat in Med, Filename=GFR Slope Power Estimates - Case 1);
```

The results from this call to the macro are shown below in Table A.3 (F-tests) and Table A.4 (t-tests). The power estimates for the F-tests and t-tests are exactly the same as they should be. Table A.3 provides information related to the non-centrality parameters of the F-tests and the denominator degrees of freedom used. Table A.4 provides the assumed parameter values and the corresponding treatment effects (Treated — Control) on the acute, delta (change), chronic and total slopes for the hypothesized mean profiles one wishes to base power calculations on. Here we find that a sample of 300 patients per group or 600 patients total would be adequate to detect a 1.008 mL/min/1.73m^2/year difference in the chronic slopes with approximate power of 0.91. This is in contrast to results based on a sample of 1,148 subjects as reported for the irbesartan vs placebo control groups in which patients randomized to receive irbesartan had a 20% lower risk of reaching the primary composite endpoint of death, ESKD or a doubling of serum creatinine compared to the placebo control group, p-value=0.02.[13] In terms of its clinical relevance, a 1.008 mL/min/1.73m^2/year reduction in the chronic slope due to treatment translates into an extended time to ESKD of 1.344 years for the average or typical patient based on the assumed acute and delta (change) slopes for the control and treated subjects (see Section 6.0 for a discussion of the clinical benefit associated with reductions in the rate of decline in eGFR).

An alternative more conservative estimate of power is obtained by setting the following options: POM=EM, Reps=500, LowerGFR=15, UpperGFR=120, Accrual=36, Followup=24, Dropout=0.05 and InfCensor=YES. With these options, the macro program will generate a sample of M=500 multivariate normal vectors of subject-specific random intercept and slope effects for each treatment group under the assumptions described in Section 2.1 in combination with the same input parameters as used in the above call to *%GFR_Slope_Power*. The M=500 generated random effects per treatment group are then used obtain the estimated marginal POM (EM-POM) weights as

$$w_{ij} = 1/\sum_{m=1}^{M}(\mu_{ij}^2(\boldsymbol{\beta}_c + \boldsymbol{b}_{mc}))^{\theta} \text{ for the control group with } \boldsymbol{b}_{mc} \sim N(\mathbf{0}, \boldsymbol{\Psi}_c)$$

and

$$w_{ij} = 1/\sum_{m=1}^{M}(\mu_{ij}^2(\boldsymbol{\beta}_t + \boldsymbol{b}_{mt}))^{\theta} \text{ for the treated group with } \boldsymbol{b}_{mt} \sim N(\mathbf{0}, \boldsymbol{\Psi}_t)$$

as described in Section 6. The macro will also generate a single multivariate normal vector of random intercept and slope effects for each of the N+S subjects per treatment group (S is the additional number of subjects used for screening) under the assumptions described in Section 2.1. These, in turn, are used to 1) screen for the first 2N subjects who meet the inclusion/exclusion criteria, and 2) allow for patient dropout due to either informative censoring or random dropout, whichever occurs first. Informative censoring due to ESKD is defined to have occurred on the first measurement occasion where a subject's predicted eGFR is at or below 15 mL/min/1.73m^2, while random dropout is defined to have occurred when a patient's randomly generated dropout time occurs (random times to dropout are generated assuming an exponential distribution with a constant annual dropout rate equal to 5%). This is illustrated in Case 2 using the following call to *%GFR_Slope_Power*.

```
/*=============================================================================================
Case 2: Assumes administrative censoring due to staggered entry, random dropout and
censoring due to ESKD. We continue to use slope and variance-covariance estimates based
on SP Model 1 for comparison to Case 1 above. Power is based on a EM-POM variance structure
and the same study design considerations as in Case 1. We use the same starting seed so as
to keep the same set of random effects. We employed a screening of an additional 100
subjects per treatment group (200 additional subjects in total) to ensure subjects enrolled
do not have starting predicted eGFR values < 15 mL/min/1.73m2 or > 120 mL/min/1.73m2.
=============================================================================================*/

%GFR_Slope_Power(Data=CKD, knot=4, beta0=50.00, beta1c=-1.012, beta1t=-1.213, beta2c=0.593,
 beta2t=0.878, Var_e=3.501, Var_u0=325.825, Var_u1=1.088, Var_u2=0.876, Cov_u0u1=2.260,
 Cov_u0u2=-3.032, Cov_u1u2=-0.935, kappa=-0.061, theta=0.917, DIFF=,
 Months=%str(0,3,6,12,18,24,30,36,42,48,54), POM=EM, Reps=500, LowerGFR=15, UpperGFR=120,
 Screen=100, Accrual=36, Followup=24, Dropout=0.05, InfCensor=YES, Alpha=0.05, DDF=NLMIXED,
 N=300 to 600 by 100, Seed=95738, Close=close, Check=NO, Print=YES, Table_F=., Table_t=A.5,
 Directory=d:\Stat in Med, Filename=GFR Slope Power Estimates - Case 2);
```

The results based on the t-tests are summarized in Table A.5 below. In this case, the power to detect a 1.008 mL/min/1.73m^2/year difference in the chronic slopes based on 300 subjects per group is estimated to be only 0.71 reflecting a loss in precision due to incomplete data. Increasing the sample to 400 subjects per group overcomes this loss in precision with power now estimated to be 0.82.

Since the results in Tables A.5 reflect a one-time simulated dataset of subject-specific random effects used to screen patients and simulate nonignorable dropout, one may question whether the power approximations are subject to "sampling" error. To address this concern a second macro, *%Replicate*, was written in which one can call the macro *%GFR_Slope_Power* repeatedly from within *%Replicate* so as to obtain estimated power over a user-specified number of replications. The macro call to *%Replicate* is shown below with results summarized in Table A.6.

```
%macro Replicate(StartSeed=12345, Rep=100);
        data seeds;
         retain seed &StartSeed;
         do i = 1 to &Rep;
           call ranuni(seed,uni);
           output;
         end;
        run;
        data Avg_F_Tests;set _null_;run;
        data Avg_t_Tests;set _null_;run;
        %do i=1 %to &rep;
        data seed;
         set seeds;
         if i=&i;
        run;
        data seed;
         set seed;
         call symput('seed_i', seed);
        run;
        %GFR_Slope_Power(Data=CKD, knot=4, beta0=50.00, beta1c=-1.012, beta1t=-1.213,
          beta2c=0.593, beta2t=0.878, Var_e=3.501, Var_u0=325.825, Var_u1=1.088,
          Var_u2=0.876, Cov_u0u1=2.260, Cov_u0u2=-3.032, Cov_u1u2=-0.935, kappa=-0.061,
          theta=0.917, DIFF=, Months=%str(0,3,6,12,18,24,30,36,42,48,54), POM=EM, Reps=500,
          LowerGFR=15, UpperGFR=120, Screen=150, Accrual=36, Followup=24, Dropout=0.05,
          InfCensor=YES, Alpha=0.05, DDF=NLMIXED, N=400 to 400, Seed=&seed_i, Close=close,
          Check=NO, Print=NO, Table_F=., Table_t=., Directory=d:\Stat in Med, Filename=);
```

```
        data Avg_F_Tests;
         set Avg_F_Tests F_Tests(in=a);
         if a then do; replication=&i; seed=&seed; end;
        run;
        data Avg_t_Tests;
         set Avg_t_Tests t_Tests(in=a);
         if a then do; replication=&i; seed=&seed; end;
        run;
        %end;
%mend Replicate;
filename bootlog "d:\Stat in Med\MacroPower.log";
proc printto log=bootlog;
run;
%Replicate(startseed=82346, rep=100);
proc printto;
run;
```

The number of subjects per group was set at 400 so as to compare the power estimates shown in Table A.5 for N=400 subjects per group but based on a single "dummy" dataset with the power estimates for N=400 subjects per group but based on the average of 100 simulated "dummy" datasets. To ensure having 400 patients per group for each of the 100 "dummy" datasets, we increased the number of additional subjects to be screened from 100 to 150 per group.

The power estimates summarized in Table A.5 are in rough agreement with the means shown in Table A.6 as one might expect based on the fact that population parameters remain fixed across all 100 replicated "dummy" datasets (hence the small standard deviations and narrow ranges in the power estimates). This suggests that a single run of the SAS macro *%GFR_Slope_Power* based on a single "dummy" dataset with one-time simulated dropout times and subject-specific random effects will provide a reasonable initial estimate of power for a fixed sample size. Once one has narrowed down the choice of N to a single value using single calls to *%GFR_Slope_Power*, one can run the macro *%Replicate* to provide a range within which one can expect the power for any of the slope comparisons to lie.

Table A.3

Approximate power of F-tests for comparing eGFR slope parameters (mL/min/1.73m2/year) between treatment groups based on a linear spline mixed-effects model with specified model parameters, 0 month accrual period, 60 month follow-up, and designated sample sizes.

Estimated power for jointly testing the following null hypothesis at a type I error=0.05

H0: No Differences in Slopes Between Two Treatment Groups (F-Test based on 1 DF).

		F-test type I error		*0.05*		
N per Group	*Select Hypotheses*	*F-test NDF*	*F-test DDF*	*F-test Critical Value*	*F-test Noncentrality Parameter*	*Power*
300	H1: No Acute Slope Effect	1	597	3.8571	1.2924	0.205
	H2: No Delta Slope Effect	1	597	3.8571	2.6071	0.364
	H3: No Chronic Slope Effect	1	597	3.8571	10.7768	0.906
	H4a: No Total Slope Effect (2 Yrs)	1	597	3.8571	0.9320	0.161
	H4b: No Total Slope Effect (3 Yrs)	1	597	3.8571	2.7954	0.385
	H4c: No Total Slope Effect (4 Yrs)	1	597	3.8571	4.3944	0.552
400	H1: No Acute Slope Effect	1	797	3.8532	1.7232	0.258
	H2: No Delta Slope Effect	1	797	3.8532	3.4761	0.461
	H3: No Chronic Slope Effect	1	797	3.8532	14.3691	0.966
	H4a: No Total Slope Effect (2 Yrs)	1	797	3.8532	1.2426	0.199
	H4b: No Total Slope Effect (3 Yrs)	1	797	3.8532	3.7272	0.487
	H4c: No Total Slope Effect (4 Yrs)	1	797	3.8532	5.8592	0.676
500	H1: No Acute Slope Effect	1	997	3.8508	2.1540	0.311
	H2: No Delta Slope Effect	1	997	3.8508	4.3451	0.548
	H3: No Chronic Slope Effect	1	997	3.8508	17.9614	0.988
	H4a: No Total Slope Effect (2 Yrs)	1	997	3.8508	1.5533	0.238
	H4b: No Total Slope Effect (3 Yrs)	1	997	3.8508	4.6590	0.577
	H4c: No Total Slope Effect (4 Yrs)	1	997	3.8508	7.3240	0.771
600	H1: No Acute Slope Effect	1	1197	3.8492	2.5848	0.362
	H2: No Delta Slope Effect	1	1197	3.8492	5.2141	0.626
	H3: No Chronic Slope Effect	1	1197	3.8492	21.5536	>0.99
	H4a: No Total Slope Effect (2 Yrs)	1	1197	3.8492	1.8639	0.276
	H4b: No Total Slope Effect (3 Yrs)	1	1197	3.8492	5.5908	0.656
	H4c: No Total Slope Effect (4 Yrs)	1	1197	3.8492	8.7888	0.841

Power based on a PA POM variance and a dropout rate of 0% per year.

Table A.4

Approximate power of t-tests for comparing eGFR slope parameters (mL/min/1.73m2/year) between treatment groups based on a linear spline mixed-effects model with specified model parameters, 0 month accrual period, 60 month follow-up, and designated sample sizes.

Power calculation results for jointly testing the following null hypothesis at a 2-sided type I error=0.05

H0: No Differences in Slopes Between Two Treatment Groups (t-test)

			Sample Size per Group							
			300		*400*		*500*		*600*	
Slope Parameter	*Treatment Comparison*	*Assumed Value*	*Standard Error*	*Power*	*Standard Error*	*Power*	*Standard Error*	*Power*	*Standard Error*	*Power*
Acute Slope	Control	-12.144	1.5087	.	1.3066	.	1.1686	.	1.0668	.
	Treated	-14.556	1.4917	.	1.2919	.	1.1555	.	1.0548	.
	Treated-Control	-2.412	2.1217	0.205	1.8374	0.258	1.6434	0.311	1.5002	0.362
Chronic Slope	Control	-5.028	0.2227	.	0.1929	.	0.1725	.	0.1575	.
	Treated	-4.020	0.2114	.	0.1830	.	0.1637	.	0.1495	.
	Treated-Control	1.008	0.3071	0.906	0.2659	0.966	0.2378	0.988	0.2171	>0.99
Delta Slope	Control	7.116	1.5035	.	1.3021	.	1.1646	.	1.0632	.
	Treated	10.536	1.4919	.	1.2920	.	1.1556	.	1.0549	.
	Treated-Control	3.420	2.1181	0.364	1.8343	0.461	1.6407	0.548	1.4977	0.626
Total Slope (2 Yrs)	Control	-6.214	0.3267	.	0.2830	.	0.2531	.	0.2310	.
	Treated	-5.776	0.3146	.	0.2725	.	0.2437	.	0.2225	.
	Treated-Control	0.438	0.4536	0.161	0.3928	0.199	0.3514	0.238	0.3207	0.276
Total Slope (3 Yrs)	Control	-5.819	0.2716	.	0.2352	.	0.2104	.	0.1921	.
	Treated	-5.192	0.2592	.	0.2244	.	0.2008	.	0.1833	.
	Treated-Control	0.628	0.3754	0.385	0.3251	0.487	0.2908	0.577	0.2655	0.656
Total Slope (4 Yrs)	Control	-5.621	0.2500	.	0.2165	.	0.1936	.	0.1768	.
	Treated	-4.899	0.2375	.	0.2056	.	0.1839	.	0.1679	.
	Treated-Control	0.723	0.3448	0.552	0.2986	0.676	0.2671	0.771	0.2438	0.841

Power based on a PA POM variance and a dropout rate of 0% per year.

An extended time to ESKD due to treatment is predicted to be, on average, 1.344 years where ESKD is defined to have occurred once eGFR reaches 15 mL/min/1.73m2.

Table A.5

Approximate power of t-tests for comparing eGFR slope parameters (mL/min/1.73m2/year) between treatment groups based on a linear spline mixed-effects model with specified model parameters, 36 month accrual period, 24 month follow-up, and designated sample sizes.

Power calculation results for jointly testing the following null hypothesis at a 2-sided type I error=0.05

H0: No Differences in Slopes Between Two Treatment Groups (t-test)

			Sample Size per Group							
			300		*400*		*500*		*600*	
Slope Parameter	*Treatment Comparison*	*Assumed Value*	*Standard Error*	*Power*	*Standard Error*	*Power*	*Standard Error*	*Power*	*Standard Error*	*Power*
Acute Slope	Control	-12.144	1.6694	.	1.4563	.	1.3044	.	1.1859	.
	Treated	-14.556	1.6561	.	1.4457	.	1.2891	.	1.1739	.
	Treated-Control	-2.412	2.3515	0.176	2.0520	0.216	1.8339	0.259	1.6686	0.303
Chronic Slope	Control	-5.028	0.2905	.	0.2504	.	0.2305	.	0.2052	.
	Treated	-4.020	0.2729	.	0.2434	.	0.2113	.	0.1961	.
	Treated-Control	1.008	0.3985	0.714	0.3492	0.822	0.3127	0.896	0.2838	0.943
Delta Slope	Control	7.116	1.7258	.	1.4998	.	1.3520	.	1.2237	.
	Treated	10.536	1.7115	.	1.4967	.	1.3295	.	1.2136	.
	Treated-Control	3.420	2.4306	0.289	2.1188	0.364	1.8962	0.437	1.7234	0.509
Total Slope (2 Yrs)	Control	-6.214	0.3481	.	0.3054	.	0.2720	.	0.2479	.
	Treated	-5.776	0.3352	.	0.2943	.	0.2621	.	0.2391	.
	Treated-Control	0.438	0.4832	0.147	0.4241	0.177	0.3777	0.212	0.3445	0.245
Total Slope (3 Yrs)	Control	-5.819	0.3009	.	0.2631	.	0.2361	.	0.2140	.
	Treated	-5.192	0.2857	.	0.2524	.	0.2231	.	0.2046	.
	Treated-Control	0.628	0.4149	0.326	0.3645	0.405	0.3249	0.488	0.2960	0.563
Total Slope (4 Yrs)	Control	-5.621	0.2865	.	0.2497	.	0.2255	.	0.2034	.
	Treated	-4.899	0.2703	.	0.2395	.	0.2107	.	0.1939	.
	Treated-Control	0.723	0.3939	0.449	0.3460	0.550	0.3087	0.647	0.2810	0.729

Power based on a EM-POM variance, a dropout rate of 5% per year, and censoring due to ESKD. An extended time to ESKD due to treatment is predicted to be, on average, 1.344 years where ESKD is defined to have occurred once eGFR reaches 15 mL/min/1.73m2.

Table A.6

Approximate power of t-tests for comparing eGFR slope parameters (mL/min/1.73m2/year) between treatment groups based on a linear spline mixed-effects model with specified model parameters, 36 month accrual period, 24 month follow-up, a sample size of 400 subjects per group and 100 replicated 'dummy' datasets to assess how robust estimates from a single 'dummy' dataset are to variation in simulated screening and simulated dropout to ESKD.

Power calculation results for testing the following null hypotheses at a 2-sided type I error=0.05.

H0: No Differences in Slopes Between Two Treatment Groups (t-test)

Summary of Power Calculations

N = 400 Subjects per Group

Slope Parameter	*Treatment Comparison*	*Assumed Value*	*Mean Power*	*SD*	*Median Power*	*Minimum Power*	*Maximum Power*
Acute Slope	Control	-12.144					
	Treated	-14.556					
	Treated-Control	-2.412	0.215	0.003	0.215	0.208	0.223
Chronic Slope	Control	-5.028					
	Treated	-4.020					
	Treated-Control	1.008	0.827	0.007	0.828	0.809	0.845
Delta Slope	Control	7.116					
	Treated	10.536					
	Treated-Control	3.420	0.361	0.005	0.361	0.348	0.376
Total Slope (2 Yrs)	Control	-6.214					
	Treated	-5.776					
	Treated-Control	0.438	0.178	0.001	0.178	0.174	0.182
Total Slope (3 Yrs)	Control	-5.819					
	Treated	-5.192					
	Treated-Control	0.628	0.408	0.004	0.408	0.399	0.419
Total Slope (4 Yrs)	Control	-5.621					
	Treated	-4.899					
	Treated-Control	0.723	0.556	0.005	0.555	0.541	0.570

Power based on an EM-POM variance, a dropout rate of 5% per year, and censoring due to ESKD. An extended time to ESKD due to treatment is predicted to be, on average, 1.344 years where ESKD is defined to have occurred once eGFR reaches 15 mL/min/1.73m2.

Finally, to demonstrate the impact that an acute treatment effect can have on power, we ran two additional power calculations using, as input, the parameters from SP model 2 (Table 5). In Case 3 (see the code below), we held the all the parameters fixed except for the treatment group chronic slope (and hence delta slope) which we varied by specifying clinically relevant differences between the treated and control group chronic slopes through specification of the SAS macro variable DIFF.

```
/*=================================================================================================
Case 3: Assumes administrative censoring due to staggered entry, random dropout and
censoring due to ESKD. We use slope and variance-covariance estimates from SP Model 2.
Here we summarize the power required to detect specified differences in the chronic slope
as well as for delta and total slopes holding the observed difference in the control and
treated acute slopes fixed. Results assume a within-subject EM-POM variance structure.
We kept the same seed as for Case 1 and Case 2 above to maintain the same "dummy" dataset.
We set the value of DIFF = .0625 to .125 by .02083 in units of mL/min/1.73m2/month which
is equivalent to DIFF = 0.75 to 1.50 by 0.25 in units of mL/min/1.73m2/year which
correspond to incremental differences (Treated minus Control) in the chronic slopes.
=================================================================================================*/
%GFR_Slope_Power(Data=CKD, knot=4, beta0=50.00, beta1c=-1.038, beta1t=-1.227, beta2c=0.579,
 beta2t=0.834, Var_e=3.355, Var_u0=327.977, Var_u1=1.034, Var_u2=0.794, Cov_u0u1=1.831,
 Cov_u0u2=-2.265, Cov_u1u2=-0.863, kappa=0.028, theta=0.922, DIFF=.0625 to .125 by .02083,
 Months=%str(0,3,6,12,18,24,30,36,42,48,54), POM=EM, Reps=500, LowerGFR=15, UpperGFR=120,
 Screen=100, Accrual=36, Followup=24, Dropout=0.05, InfCensor=YES, Alpha=0.05, DDF=NLMIXED,
 N=300 to 600 by 100, Seed=95738, Close=close, Check=NO, Print=YES, Table_F=., Table_t=A.7,
 Directory=d:\Stat in Med, Filename=GFR Slope Power Estimates - Case 3);
```

In Case 4 (see code below), we set the acute treatment group slope at -1.133 mL/min/1.73m^2/month corresponding to a 50% reduction in the acute treatment effect at 4 months with a mean difference in eGFR going from -0.756 mL/min/1.73m^2 (treated – control) to a mean difference of -0.38 mL/min/1.73m^2. This corresponds to an approximate 50% reduction in the acute slope differences (from -2.268 mL/min/1.73m^2/year to –1.14 mL/min/1.73m^2/year or, equivalently, from -0.189 mL/min/1.73m^2/month to -0.095 mL/min/1.73m^2/month). All other specifications remained the same.

```
/*=================================================================================================
Case 4: Assumes administrative censoring due to staggered entry, random dropout and
censoring due to ESKD. We use slope and variance-covariance estimates from SP Model 2 but
we set the treatment group acute slope at -1.133 mL/min/1.73m2/month corresponding to an
approximate 50% reduction in the acute treatment effect at 4 months with the mean
difference in eGFR going from -0.756 mL/min/1.73m2 (treated ╥ control) to a mean difference
of -0.38 mL/min/1.73m2. Here we summarize the power required to detect the same specified
differences in the chronic slopes but with alternative differences in the delta and total
slopes resulting from a change to the treatment group acute slope. Results assume a
within-subject EM-POM variance structure. We kept the same seed as in Cases 1-3 above.
=================================================================================================*/
%GFR_Slope_Power(Data=CKD, knot=4, beta0=50.00, beta1c=-1.038, beta1t=-1.133, beta2c=0.579,
 beta2t=0.834, Var_e=3.355, Var_u0=327.977, Var_u1=1.034, Var_u2=0.794, Cov_u0u1=1.831,
 Cov_u0u2=-2.265, Cov_u1u2=-0.863, kappa=0.028, theta=0.922, DIFF=.0625 to .125 by .02083,
 Months=%str(0,3,6,12,18,24,30,36,42,48,54), POM=EM, Reps=500, LowerGFR=15, UpperGFR=120,
 Screen=100, Accrual=36, Followup=24, Dropout=0.05, InfCensor=YES, Alpha=0.05, DDF=NLMIXED,
 N=300 to 600 by 100, Seed=95738, Close=close, Check=NO, Print=YES, Table_F=., Table_t=A.8,
 Directory=d:\Stat in Med, Filename=GFR Slope Power Estimates - Case 4);
```

The results from these two calls to *%GFR_Slope_Power* are summarized in Tables A.7-A.8, respectively. As noted in the Section 6, by reducing the acute treatment effect by 50%, the power to detect differences in the total slopes improves dramatically. To illustrate, for a chronic slope treatment effect of 1.25 mL/min/1.73m^2/year, 500 patients per group is required to detect a 4 year total slope treatment effect of 0.957 mL/min/1.73m^2/year with power of 83% (Table A.7) compared

to 400 patients per group being required to detect a 4 year total slope treatment effect of 1.051 mL/min/1.73m^2/year with power of 81% when the acute treatment effect is reduced by 50% (Table A.8).

A full description of the SAS macro *%GFR_Slope_Power* along with the macro *%Replicate* is included in the SAS program "SIM Supplement - SAS MACRO GFR_Slope_Power.sas" which is being made available through the repository Figshare.

Table A.7

Approximate power of t-tests for comparing eGFR slope parameters (mL/min/1.73m2/year) between treatment groups based on a linear spline mixed-effects model with specified model parameters, 36 month accrual period, 24 month follow-up, and designated sample sizes. The treated and control group acute slopes are fixed at -14.724 and -12.456 (mL/min/1.73m2/yr) corresponding to an acute treatment effect of -0.756 (mL/min/1.73m2) at 4 months. The control group chronic slope is fixed at -5.508 (mL/min/1.73m2/yr).

Power calculation results for testing the following null vs alternative hypotheses at a 2-sided type I error=0.05

H0: Treatment Slope - Control Slope = 0 (no difference)

HA: Treatment Slope - Control Slope = DIFF (difference)

				Sample Size (N) per Group			
				300	*400*	*500*	*600*
				Treated-Control	*Treated-Control*	*Treated-Control*	*Treated-Control*
Chronic Slope Treatment Difference mL/min/1.73m2 per year	*Extended Years to ESKD Due to Treatment*	*Parameter*	*DIFF*	*Power*	*Power*	*Power*	*Power*
0.75	0.724	Chronic Slope	0.750	0.420	0.515	0.614	0.689
		Delta Slope	3.018	0.234	0.294	0.355	0.412
		Total Slope (2 Yrs)	0.247	0.077	0.085	0.095	0.104
		Total Slope (3 Yrs)	0.414	0.153	0.186	0.224	0.256
		Total Slope (4 Yrs)	0.498	0.217	0.267	0.327	0.375
1.00	1.075	Chronic Slope	1.000	0.650	0.758	0.852	0.905
		Delta Slope	3.268	0.266	0.335	0.405	0.469
		Total Slope (2 Yrs)	0.455	0.144	0.175	0.210	0.239
		Total Slope (3 Yrs)	0.637	0.299	0.373	0.454	0.517
		Total Slope (4 Yrs)	0.727	0.403	0.498	0.598	0.669
1.25	1.466	Chronic Slope	1.250	0.835	0.914	0.963	0.983
		Delta Slope	3.518	0.301	0.379	0.457	0.526
		Total Slope (2 Yrs)	0.663	0.255	0.319	0.389	0.445
		Total Slope (3 Yrs)	0.859	0.489	0.597	0.702	0.771
		Total Slope (4 Yrs)	0.957	0.617	0.729	0.827	0.883
1.50	1.907	Chronic Slope	1.500	0.940	0.979	>0.99	>0.99
		Delta Slope	3.768	0.337	0.424	0.509	0.583
		Total Slope (2 Yrs)	0.872	0.400	0.498	0.596	0.668
		Total Slope (3 Yrs)	1.081	0.682	0.793	0.880	0.925
		Total Slope (4 Yrs)	1.186	0.799	0.890	0.949	0.974

Power based on the EM-POM variance, a dropout rate of 5% per year, and censoring due to ESKD.

Table A.8

Approximate power of t-tests for comparing eGFR slope parameters (mL/min/1.73m2/year) between treatment groups based on a linear spline mixed-effects model with specified model parameters, 36 month accrual period, 24 month follow-up, and designated sample sizes. The treated and control group acute slopes are fixed at -13.596 and -12.456 (mL/min/1.73m2/yr) corresponding to an acute treatment effect of -0.38 (mL/min/1.73m2) at 4 months. The control group chronic slope is fixed at -5.508 (mL/min/1.73m2/yr).

Power calculation results for testing the following null vs alternative hypotheses at a 2-sided type I error=0.05

H0: Treatment Slope - Control Slope = 0 (no difference)

HA: Treatment Slope - Control Slope = DIFF (difference)

				Sample Size (N) per Group			
				300	*400*	*500*	*600*
				Treated-Control	*Treated-Control*	*Treated-Control*	*Treated-Control*
Chronic Slope Treatment Difference mL/min/1.73m2 per year	*Extended Years to ESKD Due to Treatment*	*Parameter*	*DIFF*	*Power*	*Power*	*Power*	*Power*
0.75	0.803	Chronic Slope	0.750	0.420	0.514	0.615	0.689
		Delta Slope	1.890	0.120	0.143	0.168	0.193
		Total Slope (2 Yrs)	0.435	0.136	0.164	0.195	0.222
		Total Slope (3 Yrs)	0.540	0.228	0.283	0.346	0.397
		Total Slope (4 Yrs)	0.592	0.287	0.356	0.435	0.497
1.00	1.158	Chronic Slope	1.000	0.650	0.758	0.852	0.905
		Delta Slope	2.140	0.141	0.171	0.203	0.234
		Total Slope (2 Yrs)	0.643	0.242	0.303	0.369	0.423
		Total Slope (3 Yrs)	0.762	0.403	0.499	0.598	0.669
		Total Slope (4 Yrs)	0.821	0.491	0.598	0.703	0.772
1.25	1.555	Chronic Slope	1.250	0.834	0.914	0.964	0.983
		Delta Slope	2.390	0.164	0.202	0.242	0.280
		Total Slope (2 Yrs)	0.852	0.385	0.480	0.576	0.648
		Total Slope (3 Yrs)	0.984	0.600	0.715	0.814	0.873
		Total Slope (4 Yrs)	1.051	0.698	0.806	0.891	0.934
1.50	2.001	Chronic Slope	1.500	0.940	0.979	>0.99	>0.99
		Delta Slope	2.640	0.189	0.236	0.284	0.330
		Total Slope (2 Yrs)	1.060	0.546	0.662	0.764	0.829
		Total Slope (3 Yrs)	1.206	0.774	0.873	0.938	0.967
		Total Slope (4 Yrs)	1.280	0.855	0.930	0.973	0.988

Power based on the EM-POM variance, a dropout rate of 5% per year, and censoring due to ESKD.

WEB APPENDIX B: Study acronyms/abbreviations associated with the CKD-EPI investigators/collaborators listed in the Acknowledgements.

Appendix B: Abbreviations, Units, and Terms

AASK	African American Study of Kidney Disease and Hypertension
ABCD	Appropriate Blood Pressure Control in Diabetes trial
ACE	angiotensin converting enzyme
ADVANCE	Action in Diabetes and Vascular Disease: Preterax and Diamicron MR Controlled Evaluation trial
AGQ	adaptive Gaussian quadrature
AIC	Akaike information criterion
ALTITUDE	Aliskiren Trial in Type 2 Diabetes Using Cardiorenal Endpoints
ARB	angiotensin II receptor blocker
CanPREVENT	Canadian Prevention of Renal and Cardiovascular Endpoints Trial
CCB	Calcium channel blocker
CI	confidence interval
CKD	chronic kidney disease
CSG	Collaborative Study Group
eGFR	estimated glomerular filtration rate
EMPA-REG OUTCOME	Empagliflozin Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients (referred to as EMPA-REG here on in)
EPV	events per predictor variable
GFR	glomerular filtration rate(mL/min/1.73 m2)
HALT-PKD	Halt Progression of Polycystic Kidney Disease study
HKVIN	Hong Kong study using Valsartan in IgA Nephropathy
HR	hazard ratio
IDNT	Irbesartan Diabetic Nephropathy Trial
IgA	immunoglobulin A nephropathy
IPF	idiopathic pulmonary fibrosis
MAR	missing at random
MASTERPLAN	Multifactorial Approach and Superior Treatment Efficacy in Renal Patients with the Aid of Nurse Practitioners study
MDRD Study	Modification of Diet in Renal Disease study
N	sample size
NKF	National Kidney Foundation
ORIENT	Olmesartan Reducing Incidence of Endstage Renal Disease in Diabetic Nephropathy Trial
PKD	polycystic kidney disease
RCT	randomized control trial
REIN	Ramipril Efficacy In Nephropathy study
RENAAL	Reduction of Endpoints in NIDDM with the Angiotensin II Antagonist Losartan study
ROAD	Renoprotection of Optimal Antiproteinuric Doses study
SCr	serum creatinine (mg/dL)
SE	standard error
SHARP	Study of Heart and Renal Protection
SP	shared parameter
STOP-IgAN	Supportive Versus Immunosuppressive Therapy for the Treatment of Progressive IgA Nephropathy trial
SUN-MACRO	Sulodexide Macroalbuminuria trial