

Supplemental Files

Development and validation of a new hierarchical composite endpoint for clinical trials of kidney disease progression

H.J.L. Heerspink et.al.

Table of contents:

Supplement Table 1: Definition of established primary kidney endpoints in each trial.

Supplement Table 2: Hierarchical comparison of the components of the kidney HCE

Supplement Table 3: Contribution of individual endpoint components to the primary composite kidney endpoint as defined in each trial.

Supplement Table 4: Number of wins on active and placebo treatment, number of ties, and win odds for each component of the HCE using a shared follow-up time for each pairwise comparison.

Supplement Table 5: Comparison of the win odds using a non-shared and shared follow-up time

Supplement Figure 1: Number (%) of individual components of the hierarchical composite kidney endpoint excluding all-cause mortality

Supplement Figure 2: Maraca plot of the hierarchical composite kidney endpoint excluding mortality. The Maraca plots show the contribution of the different time-to-event endpoint components; the treatment effect on the different time-to-event components of the composite; and the treatment effect on the continuous GFR slope component, for patients not experiencing any of the dichotomous outcomes.

Panel A: DAPA-CKD trial; Panel B: CREDENCE trial; Panel C; FIDELIO-DKD trial; Panel D: SONAR trial; Panel E: RENAAL trial; Panel F: IDNT trial; Panel G: ALTITUDE trial

Supplement Figure 3: Sample size curves for all trials showing the sample size and statistical power of the original kidney endpoint in each trial, GFR slope, and HCE excluding

all-cause mortality. Panel A: DAPA-CKD trial; Panel B: CREDENCE trial; Panel C; FIDELIO-DKD trial; Panel D: SONAR trial; Panel E: RENAAL trial; Panel F: IDNT trial

Supplemental Document: Technical Appendix

Supplemental Excel File 1: ADSL – Patient level demographic data

Supplemental Excel File 2: 2ADLB – eGFR measurements of patients over time

Supplemental Excel File 3: ADET – Time-to-event outcomes of patients

Supplement Table 1: Definition of established primary kidney endpoints in each trial.

Clinical trial	Endpoint definition
DAPA-CKD	Sustained 50% eGFR decline, kidney failure, renal or cardiovascular death
CREDESCENCE	Sustained 57% eGFR decline, kidney failure, renal or cardiovascular death
FIDELIO-DKD	Sustained 40% eGFR decline, kidney failure, renal death
SONAR	Sustained 57% eGFR decline, kidney failure, renal death
RENAAL	Sustained 57% eGFR decline, kidney failure**, all cause mortality
IDNT	Sustained 57% eGFR decline, kidney failure‡, all cause mortality
ALTITUDE	Sustained 57% eGFR decline, kidney failure, all cause mortality

*kidney failure defined as chronic dialysis, kidney transplantation or sustained GFR<15 mL/min/1.73m²

** kidney failure defined as chronic dialysis, kidney transplantation

‡ kidney failure defined as chronic dialysis, kidney transplantation or sustained serum creatinine >6 mg/dL

Supplement Table 2 Principle of hierarchical comparison of events for a pair of patients, one in the active arm and one in the control arm.

	Outcome
First comparison:	
Patient in active arm dies first	Control wins
Patient in control arm dies first	Active wins
Both patients die on the same day	Tie
Neither of the patients die	Go to second comparison
Second comparison:	
Patient in active arm has KFRT first	Control wins
Patient in control arm has KFRT first	Active wins
Both patients have KFRT on the same day	Tie
Neither of the patients have KFRT	Go to third comparison
Third comparison:	
Patient in active arm has GFR<15 first	Control wins
Patient in control arm has GFR<15 first	Active wins
Both patients have a GFR<15 on the same day	Tie
Neither of the patients have a GFR<15	Go to fourth comparison
Fourth comparison:	
Patient in active arm has 57% GFR decline first	Control wins
Patient in control arm has 57% GFR decline first	Active wins
Both patients have a 57% GFR decline on the same day	Tie
Neither of the patients have a 57% GFR decline	Go to fifth comparison
Fifth comparison:	
Patient in active arm has 50% GFR decline first	Control wins
Patient in control arm has 50% GFR decline first	Active wins
Both patients have a 50% GFR decline on the same day	Tie
Neither of the patients have a 50% GFR decline	Go to sixth comparison
Sixth comparison:	
Patient in active arm has 40% GFR decline first	Control wins
Patient in control arm has 40% GFR decline first	Active wins
Both patients have a 40% GFR decline on the same day	Tie
Neither of the patients have a 40% GFR decline	Go to last comparison
Seventh comparison:	
Patient in active arm has faster progressing eGFR slope	Control wins
Patient in control arm has faster progressing eGFR slope	Active wins
Patients have similar eGFR slope	Tie

Supplement Table 3: Contribution of individual endpoint components to the primary composite kidney endpoint as defined in each trial.

	DAPA-CKD	CREDESCENCE	FIDELIO-DKD	SONAR	RENAAL	IDNT	ALTITUDE*
Sample size	N=4304	N=4401	N=5674	N=3668	N=1513	N=1715	N=8561
Composite Endpoint	385 (100)	377 (100)	1104 (100)	344 (100)	686 (100)	411 (100)	966 (100)
All-cause mortality	N/A	N/A	N/A	N/A	197 (28.72)	128 (31.14)	459 (30.30)
Death due to Kidney Failure	2 (0.51)	1 (0.27)	1 (0.09)	N/A	N/A	N/A	5 (0.33)
Kidney Failure	191 (49.61)	180 (47.75)	171 (15.49)	223 (64.83)	129 (18.80)	90 (21.90)	87 (5.74)
57% GFR decline	N/A	196 (51.99)	N/A	121 (35.17)	360 (52.48)	193 (46.96)	415 (27.39)
50% GFR decline	192 (49.87)	N/A	N/A	N/A	N/A	N/A	N/A
40% GFR decline	N/A	N/A	1048 (94.93)	N/A	N/A	N/A	N/A

The different trials used different endpoint components in the primary kidney endpoint. Kidney failure includes GFR<15 mL/min/1.73m². *For each subject only the first occurrence of a component of the composite efficacy outcome event is considered. If a subject experienced two or more events on the same date, all these events are presented. Hence, the number of events of the individual components do not necessarily add up to the number of subjects with the composite efficacy outcome event. *The ALTITUDE trial used a composite cardio-kidney endpoint. For the purpose of this study we replaced the cardiovascular composite component of the primary endpoint for all cause mortality.

Supplement Table 4: Number of wins on active and placebo treatment, number of ties, and win odds for each component of the HCE using a shared follow-up approach for a pairwise comparison.

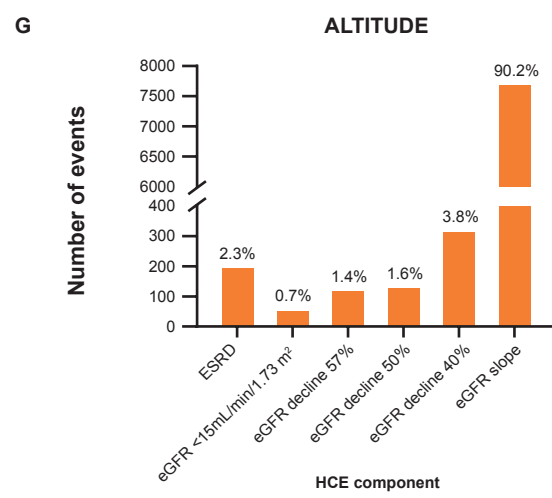
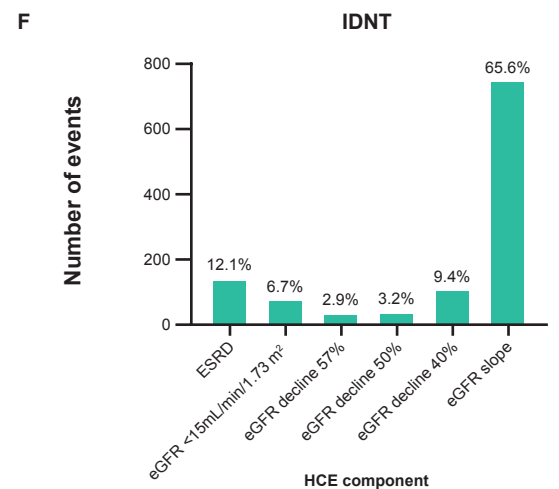
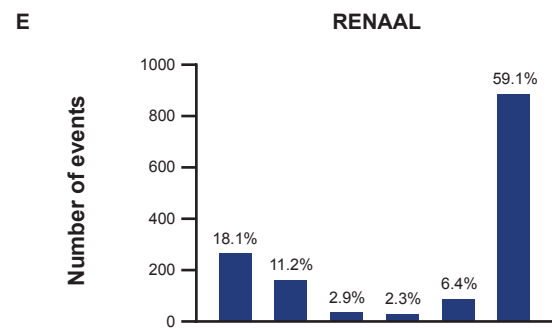
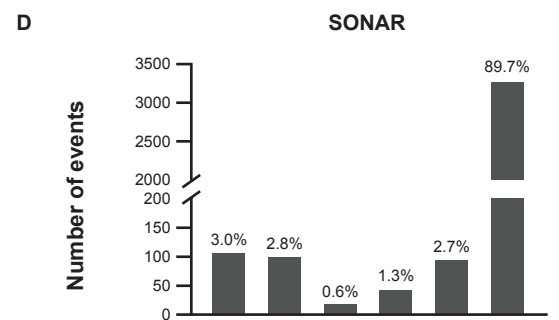
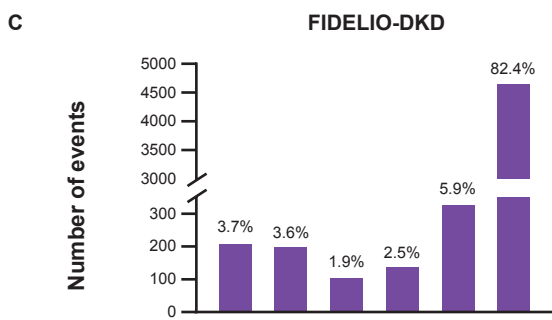
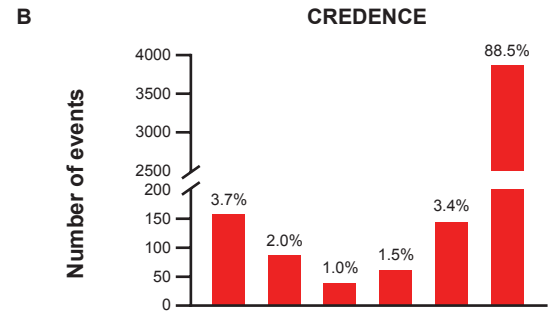
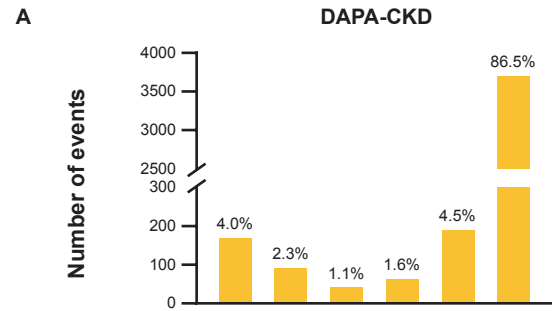
	DAPA-CKD			CREDESCENCE			FIDELIO-DKD			SONAR		
Treatment comparisons	Dapagliflozin vs placebo			Canagliflozin vs placebo			Finerenone vs placebo			Atrasentan vs placebo		
Median follow-up	2.4 years			2.6 years			2.6 years			2.2 years		
Hazard ratio (Cox)	0.61 (0.51, 0.73)			0.70 (0.59, 0.82)			0.82 (0.73, 0.93)			0.71 (0.58, 0.88)		
HCE	Win active	Win placebo	Tie	Win Active	Win placebo	Tie	Win Active	Win placebo	Tie	Win Active	Win placebo	Tie
All-cause Mortality	248185 (5.36 %)	171609 (3.71 %)	13 (<0.01 %)	357710 (7.39 %)	296126 (6.12 %)	22 (<0.01 %)	508892 (6.32 %)	453699 (5.64 %)	48 (<0.01 %)	97512 (2.91 %)	114951 (3.43 %)	5 (<0.01 %)
Kidney replacement	125860 (2.72 %)	86045 (1.86 %)	8 (<0.01 %)	118814 (2.46 %)	82520 (1.71 %)	5 (<0.01 %)	203794 (2.53 %)	172365 (2.14 %)	10 (<0.01 %)	71710 (2.14 %)	41017 (1.23 %)	8 (<0.01 %)
GFR<15 mL/min/1.73m ²	92599 (2 %)	67244 (1.45 %)	24 (<0.01 %)	89373 (1.85 %)	51941 (1.07 %)	9 (<0.01 %)	211663 (2.63 %)	170421 (2.12 %)	64 (<0.01 %)	69898 (2.09 %)	50045 (1.5 %)	5 (<0.01 %)
57% GFR decline	174153 (3.76 %)	129663 (2.8 %)	99 (<0.01 %)	55175 (1.14 %)	17924 (0.37 %)	6 (<0.01 %)	133912 (1.66 %)	72600 (0.9 %)	11 (<0.01 %)	13984 (0.42 %)	9657 (0.29 %)	2 (<0.01 %)
50% GFR decline	70102 (1.51 %)	28995 (0.63 %)	10 (<0.01 %)	66786 (1.38 %)	50458 (1.04 %)	20 (<0.01 %)	158844 (1.97 %)	107889 (1.34 %)	29 (<0.01 %)	35651 (1.07 %)	17604 (0.53 %)	1 (<0.01 %)
40% GFR decline	56646 (1.22 %)	20298 (0.44 %)	2 (<0.01 %)	166382 (3.44 %)	97957 (2.02 %)	95 (<0.01 %)	333894 (4.15 %)	325093 (4.04 %)	180 (<0.01 %)	55439 (1.66 %)	60264 (1.8 %)	11 (<0.01 %)
GFR slope	1948743 (42.08 %)	1410806 (30.46 %)	0	2039535 (42.16 %)	1346942 (27.84 %)	0	2972943 (36.94 %)	2222202 (27.61 %)	0	1452892 (43.41 %)	1256414 (37.54 %)	0
Total	2716288 (58.65 %)	1914660 (41.34 %)	156 (<0.01 %)	2893775 (59.82 %)	1943868 (40.18 %)	157 (<0.01 %)	4523942 (56.21 %)	3524269 (43.79 %)	342 (<0.01 %)	1797086 (53.69 %)	1549952 (46.31 %)	32 (<0.01 %)
Win odds	1.42 (1.32, 1.52)			1.49 (1.39, 1.59)			1.28 (1.21, 1.36)			1.16 (1.08, 1.25)		
	RENAAL			IDNT			ALTITUDE					
Treatment comparisons	Losartan vs placebo			Irbesartan vs placebo			Aliskiren vs placebo					
Median follow-up	3.4 years			2.6 years			2.7 years					
Hazard ratio (Cox)	0.79 (0.66, 0.94)			0.74 (0.59, 0.94)			1.08 (0.95, 1.23)					
HCE	Win active	Win placebo	Tie	Win Active	Win placebo	Tie	Win Active	Win placebo	Tie			
All-cause Mortality	84651 (14.79 %)	82460 (14.41 %)	14 (<0.01 %)	24889 (7.59 %)	25638 (7.82 %)	3 (<0.01 %)	1171659 (6.41 %)	1259160 (6.89 %)	100 (<0.01 %)			
Kidney replacement	70405 (12.3 %)	46052 (8.05 %)	16 (<0.01 %)	18602 (5.68 %)	15016 (4.58 %)	2 (<0.01 %)	240800 (1.32 %)	281136 (1.54 %)	12 (<0.01 %)			
GFR<15 mL/min/1.73m ²	39275 (6.86 %)	37227 (6.51 %)	42 (0.01 %)	26021 (7.94 %)	19728 (6.02 %)	5 (<0.01 %)	98818 (0.54 %)	87533 (0.48 %)	13 (<0.01 %)			
57% GFR decline	9931 (1.74 %)	8766 (1.53 %)	5 (<0.01 %)	13199 (4.03 %)	8653 (2.64 %)	7 (<0.01 %)	171538 (0.94 %)	189509 (1.04 %)	19 (<0.01 %)			
50% GFR decline	6404 (1.12 %)	7738 (1.35 %)	2 (<0.01 %)	9587 (2.93 %)	6258 (1.91 %)	1 (<0.01 %)	215180 (1.18 %)	209595 (1.15 %)	15 (<0.01 %)			
40% GFR decline	16325 (2.85 %)	18647 (3.26 %)	15 (<0.01 %)	17405 (5.31 %)	20755 (6.33 %)	7 (<0.01 %)	466196 (2.55 %)	577521 (3.16 %)	82 (<0.01 %)			

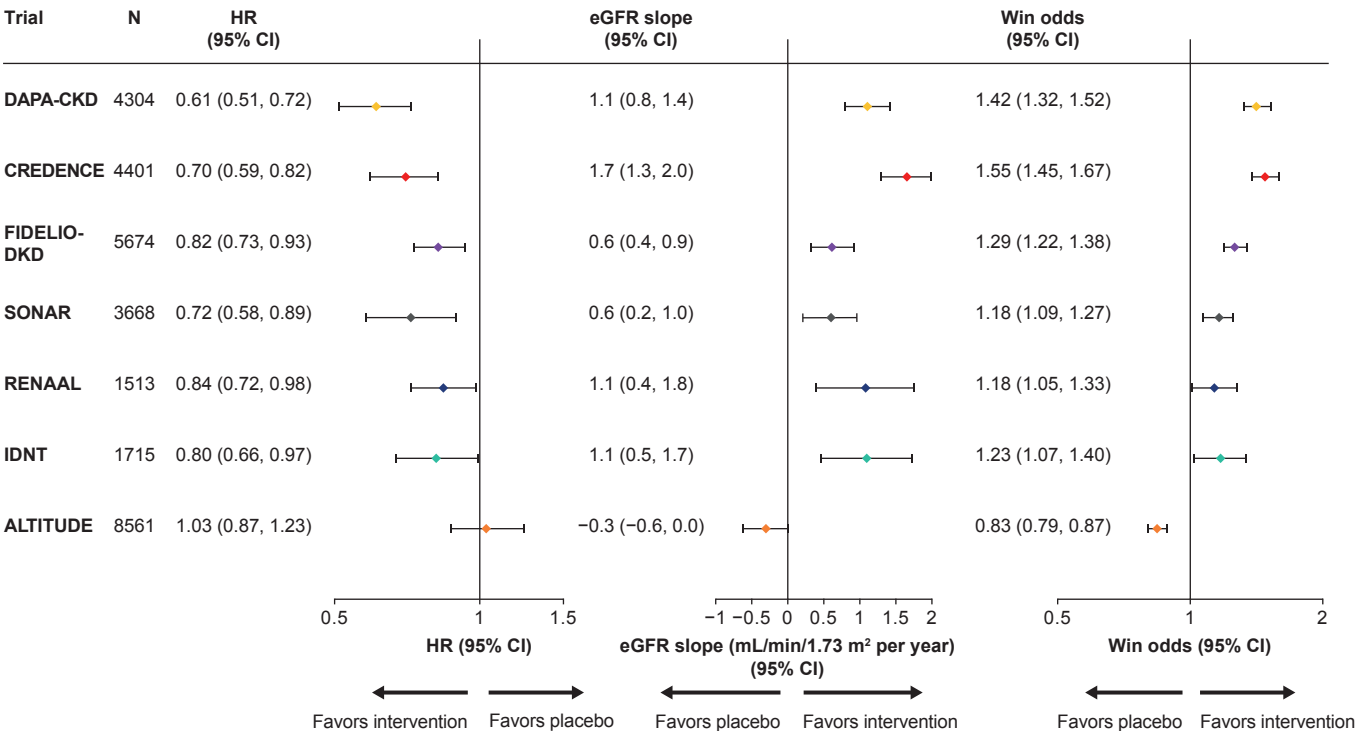
GFR slope	77067 (13.47 %)	67220 (11.75 %)	0	66591 (20.32 %)	55359 (16.89 %)	0	5977793 (32.72 %)	7324599 (40.09 %)	0
Total	304058 (53.13 %)	268110 (46.85 %)	94 (0.02 %)	176294 (53.79 %)	151407 (46.2 %)	25 (0.01 %)	8341984 (45.66 %)	9929053 (54.34 %)	241 (<0.01 %)
Win odds	1.13 (1.01, 1.27)			1.16 (1.02, 1.33)			0.84 (0.8, 0.88)		

Percentages are calculated using the overall number of comparisons. Because of low number of ties win odds is equal to win ratio. The confidence interval is calculated using the formula from Yu RX, Ganju J. Sample size formula for a win ratio endpoint. *Statistics in medicine*. 2022 Mar 15;41(6):950-63.

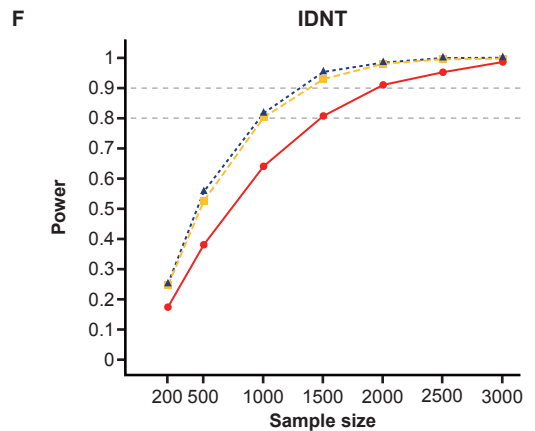
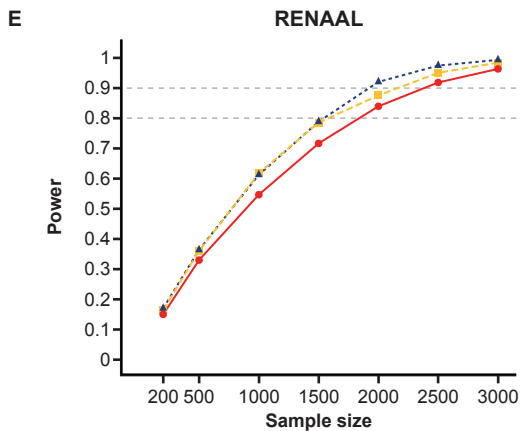
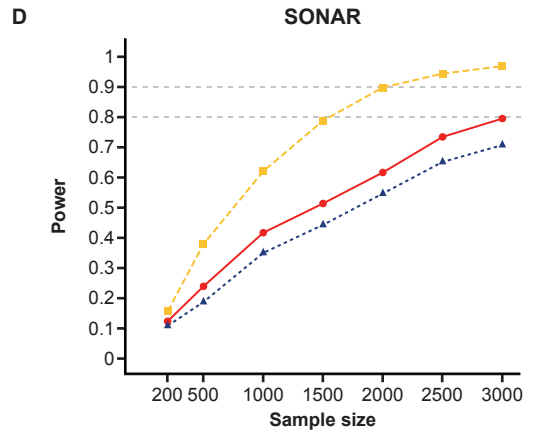
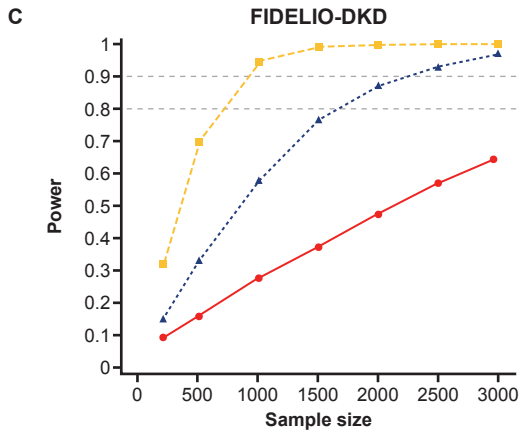
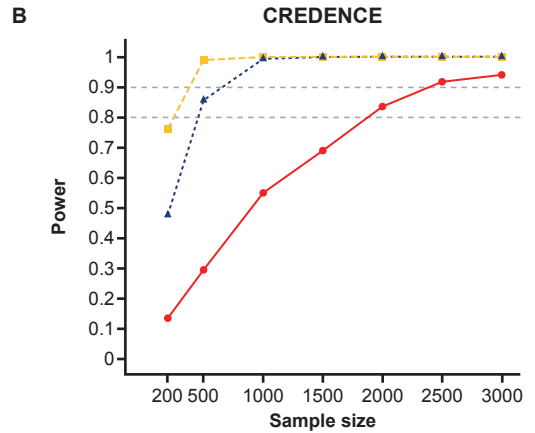
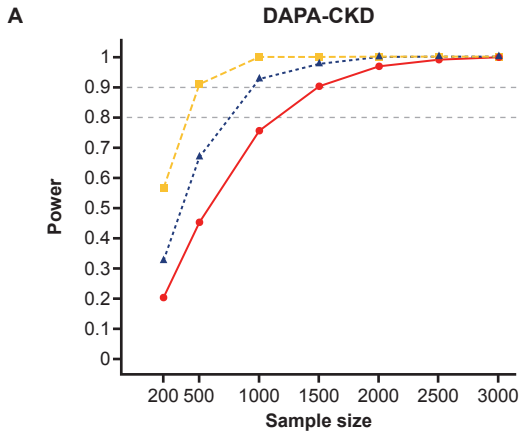
Supplement table 5: Comparison of win odds using a non-shared and shared follow-up time

Trial acronym	Win odds non-shared follow-up	Win odds shared follow-up
DAPA-CKD	1.41 (1.32, 1.52)	1.42 (1.32, 1.52)
CREDESCENCE	1.48 (1.38, 1.58)	1.49 (1.39, 1.59)
FIDELIO-DKD	1.26 (1.19, 1.34)	1.28 (1.21, 1.36)
SONAR	1.16 (1.07, 1.25)	1.16 (1.08, 1.25)
RENAAL	1.13 (1.00, 1.27)	1.13 (1.01, 1.27)
IDNT	1.17 (1.02, 1.34)	1.16 (1.02, 1.33)
ALTITUDE	0.84 (0.80, 0.88)	0.84 (0.80, 0.88)





Type: Primary endpoint (red solid line with circles), eGFR slope (blue dashed line with triangles), HCE (yellow dashed line with squares)



Derivation and analysis of kidney HCE

Niels Jongs, Samvel B. Gasparyan

25 July, 2023

Contents

Introduction	2
Input datasets	2
Total GFR slope difference calculation	5
Derivation of individual GFR slopes	6
Kidney HCE dataset	7
Win odds and maraca plots	8
Acknowledgment	8
Appendix	10
References	12


List of Tables


1	ADSL dataset structure	3
2	ADET dataset structure	3
3	Kidney HCE	3
4	ADLB dataset structure	4
5	Frequency of outcomes by treatment group	4
6	Total GFR slope difference between treatment groups	6
7	Total GFR slope by treatment groups	6
8	Predicted individual GFR slope values	7
9	Kidney HCE data structure	8
10	Win odds and its confidence interval	8

List of Figures

1	The maraca plot of kidney HCE	9
---	---	---

Introduction

In this supplementary material we will explain how to derive the analysis dataset for the kidney hierarchical composite endpoint (HCE), how to analyze it using win odds (Samvel B. Gasparyan et al. 2021), and how to visualize it using *maraca* plots (Karpefors, Lindholm, and Gasparyan 2023). Individual GFR slopes are calculated from the power-of-the-mean (POM) mixed-effects model (Vonesh et al. 2019). All calculations are done using the  software (R Core Team 2022).

- The win odds and its confidence interval are calculated using the `hce` package (Samvel B. Gasparyan 2022).
- The *maraca* plots are implemented in the package `maraca`  (Martin Karpefors, Samvel B. Gasparyan, and Monika Huhn 2023).

In addition, we will be using the following packages:

- The package `readxl` for reading the data (Wickham and Bryan 2023).
- The package `nlme` for fitting a power-of-the-mean model for GFR slopes (Pinheiro and Bates 2000).
- The package `multcomp` to derive the total GFR slope from the power-of-the-mean model (Hothorn, Bretz, and Westfall 2008).
- The package `dplyr` for data manipulation (Wickham et al. 2023).
- The package `ggplot2` for customizing the *maraca* plots (Wickham 2016).

```
library(readxl)
library(ggplot2)
library(nlme)
library(multcomp)
library(dplyr)
```

```
library(hce)
library(maraca)
```

```
packageVersion("hce")
```

```
## [1] '0.5.0'
```

```
packageVersion("maraca")
```

```
## [1] '0.5.0'
```

Input datasets

We will use the following three input synthetic datasets that are structured according to clinical data standards for analysis datasets. The datasets are `ADSL` containing patient level data, `ADLB` containing GFR measurements for all patients, and `ADET` containing events of interest (death, chronic dialysis, sustained $eGFR < 15$, sustained $eGFR$ declines of given threshold) and their study day of occurrence. Only one event per patient per type is kept in the `ADET` dataset.

```
ADSL <- read_excel("data/ADSL.xlsx")
ADET <- read_excel("data/ADET.xlsx")
ADLB <- read_excel("data/ADLB.xlsx")
```

```
head(ADSL)
```

Table 1: ADSL dataset structure

ID	TRTPN	EGFRBL	STRATAN
1	1	44	1
2	1	31	3
3	1	28	4
4	1	61	4
5	1	37	4
6	2	35	1

The dataset `ADSL` contains (see Table 1) the columns `ID` for patient ID, `TRTPN`, the planned treatment group, 1 for the active group, 2, for the control group. `STRATAN` contains randomization stratum for each patient (1-4), where a higher value means a higher risk of kidney progression.

```
head(ADET)
```

Table 2: ADET dataset structure

ID	AVAL	PARAM	PARAMCD	PARAMN	TRTPN
11	359	Sustained $\geq 50\%$ decline in eGFR	EGFR50	5	2
11	467	Sustained $\geq 57\%$ decline in eGFR	EGFR57	4	2
11	359	Sustained $\geq 40\%$ decline in eGFR	EGFR40	6	2
11	841	Sustained eGFR <15 (mL/min/1.73 m ²)	EGFR15	3	2
15	962	Death (adj)	DTHADJ	1	1
21	737	Sustained $\geq 40\%$ decline in eGFR	EGFR40	6	1

The dataset `ADET` contains subject IDs, `AVAL` column specifies the study day of the event (days from randomization), `PARAM` specifies the type of the event, `PARAMCD` the coded type of the event, while `PARAMN` specifies the priority of the event, the numbers 1 to 6, where a higher value means a better (less severe) outcome.

```
ADET |> select(PARAM, PARAMCD, PARAMN) |> arrange(PARAMN) |> unique()
```

Table 3: Kidney HCE

PARAM	PARAMCD	PARAMN
Death (adj)	DTHADJ	1
Chronic dialysis (adj) ≥ 90 days	DIAL90	2
Sustained eGFR <15 (mL/min/1.73 m ²)	EGFR15	3
Sustained $\geq 57\%$ decline in eGFR	EGFR57	4
Sustained $\geq 50\%$ decline in eGFR	EGFR50	5
Sustained $\geq 40\%$ decline in eGFR	EGFR40	6

We need to derive the category 7 of the hierarchical composite endpoint (measurements), which is for patients not having any of the dichotomous outcomes 1-6 and who contribute to the analysis with their individual GFR slope. For this we will use the dataset `ADLB` of laboratory measurements.

```
head(ADLB)
```

Table 4: ADLB dataset structure

ID	TRTPN	AVAL	ADAY	AVISITN	PARAM	PARAMCD	PARAMN
1	1	40	357	7	eGFR measurements	eGFR	7
1	1	49	119	5	eGFR measurements	eGFR	7
1	1	41	238	6	eGFR measurements	eGFR	7
1	1	36	469	8	eGFR measurements	eGFR	7
1	1	40	721	10	eGFR measurements	eGFR	7
1	1	34	602	9	eGFR measurements	eGFR	7

In this dataset the column `AVAL` is the GFR measurement of patients done at the visit specified in the column `AVISITN`, while `ADAY` is the analysis day of the visit.

We will restrict the analysis to events and GFR measurements up to given cut-off of 3 years since randomization and we will consider the acute effect happening at the first visit, which is Day 14 (we divide by 360 to convert to years). We will derive also the coefficient for the chronic phase, which is the proportion of the length of the chronic phase (total follow-up minus the acute phase) divided by the total follow-up for annualizing the results. This will be used in deriving the total GFR slope from the two slope power-of-the-mean model.

```
CUTOFF <- 3
ACUTE <- 14/360
CHRONIC_coef <- (CUTOFF - ACUTE)/CUTOFF
```

The following code annualizes the analysis days (by dividing them by 360) and restricts the GFR measurements and dichotomous outcomes to the cut-off timepoint, and only one event per patient, their most severe event, is selected (the event with the lowest `PARAMN` value). The resulting summary table shows the number of patients with a given most severe event.

```
ADLB$ADAY <- ADLB$ADAY/360
ADET$AVAL <- ADET$AVAL/360
ADLB <- ADLB[ADLB$ADAY <= CUTOFF, ]
EVNT <- ADET |> filter(AVAL <= CUTOFF) |> arrange(ID, PARAMN)|> group_by(ID) |>
  mutate (n = row_number()) |> ungroup() |> filter(n == 1) |> select(- c("n"))
```

```
table(EVNT$PARAM, EVNT$TRTPN)
```

Table 5: Frequency of outcomes by treatment group

	Active	Control
Chronic dialysis (adj) ≥ 90 days	17	29
Death (adj)	40	50
Sustained $\geq 40\%$ decline in eGFR	36	34
Sustained $\geq 50\%$ decline in eGFR	7	22

	Active	Control
Sustained $\geq 57\%$ decline in eGFR	2	9
Sustained eGFR < 15 (mL/min/1.73 m ²)	16	28

Total GFR slope difference calculation

We will merge the datasets ADLB and ADSL in order to have STRATAN and EGFRBL columns present with the GFR measurements in the same dataset. Then we will fit a *linear mixed effect model (LME)* using the function `lme()` in the package `nlme`.

The model will include only measurements up to the cut-off day, will be adjusted for baseline GFR values, and will include in the model covariates for the stratification, the treatment group (converted to a 0, 1 variable with 1 for the active group and 0 for the control), analysis day ADAY, the derived SPLINE variable which corresponds to time since the acute phase, treatment by analysis day, and treatment by spline interaction terms, without an intercept term.

```
ADLB0 <- merge(ADLB, ADSL[, c("ID", "STRATAN", "EGFRBL")], by = "ID", all.x = T)
ADLB0$SPLINE <- ifelse(ADLB0$ADAY <= ACUTE, 0, ADLB0$ADAY - ACUTE)

ADLB0$TRT <- ifelse(ADLB0$TRTP == 1, 1, 0)
fit <- lme(AVAL ~ EGFRBL + STRATAN + TRT + ADAY + SPLINE + ADAY*TRT + SPLINE*TRT - 1,
          random = list(ID = pdSymm(form = ~ 1 + ADAY)),
          weights = varComb(varIdent(form = ~ 1|TRT),
                            varPower(form = ~ 1 + ADAY)),
          na.action = na.omit,
          data = ADLB0,
          method = "REML",
          control = lmeControl(maxIter = 1e8,
                               msMaxIter = 1e8,
                               opt = "optim",
                               optimMethod = "L-BFGS-B",
                               msVerbose = T))
```

```
## Warning in optim(c(oldPars), function(lmePars) -logLik(lmeSt, lmePars), :
## method L-BFGS-B uses 'factr' (and 'pgtol') instead of 'reltol' and 'abstol'

## final value 92903.802147
## converged
```

Then, the function `glht()` from the package `multcomp` can be used to calculate the total GFR slope from a two-slope model accounting for the chronic phase through the `CHRONIC_coef` coefficient.

```
k <- ncol(coef(fit)) - 1
MTP <- glht(fit,
            linfct = rbind("Total: A - C" = c(rep(0, k - 2), 1, CHRONIC_coef)))
SLP <- confint(summary(MTP))
Slope <- SLP$confint[1:3]
```

Therefore, the total GFR slope difference between (and its 95% confidence interval) active and control group per year is retained in the variable `Slope` and presented in Table 6. GFR slope in the active and control groups is calculated in the variables `Slope0` and `Slope1` respectively (using the same model) and presented in Table 7.


```

MTP0 <- glht(fit,
             linfct = rbind("Total: A - C" = c(rep(0, k - 4), 1,
                                             CHRONIC_coef, 1, CHRONIC_coef)))
SLP0 <- confint(summary(MTP0))
Slope0 <- SLP0$confint[1:3]
MTP1 <- glht(fit,
             linfct = rbind("Total: A - C" = c(rep(0, k - 4), 1, CHRONIC_coef, 0, 0)))
SLP1 <- confint(summary(MTP1))
Slope1 <- SLP1$confint[1:3]

```

Table 6: Total GFR slope difference between treatment groups

GFR Slope difference	Lower 95% CI	Upper 95% CI
1.1132	0.5888	1.6377

Table 7: Total GFR slope by treatment groups

	GFR Slope	Lower 95% CI	Upper 95% CI
Active	-3.6464	-4.0162	-3.2765
Control	-2.5332	-2.9049	-2.1614

We will use this model for predicting individual GFR slope values for all patients at the cut-off timepoint.

Derivation of individual GFR slopes

For predicting the annualized change from baseline in GFR values for all patients at the end of follow-up, we need to create a dataset containing the necessary (the ones included in the model above) baseline covariates of all patients.

Analysis day is selected as the length of follow-up (ADAY), the SPLINE variable for the years since the acute phase will be length of follow-up minus the acute phase. Then we can predicted the GFR values at the end of follow-up using the `predict()` function to obtain predicted GFR values for all patients at the end of follow-up, and subtract the corresponding baseline GFR values for each patient (dividing by the length of follow-up will provide the annualized individual GFR slopes for all patients).

```

ADLB1 <- ADLBO[ADLBO$ADAY == 0, ]
ADLB1$ADAY <- CUTOFF
ADLB1$SPLINE <- CUTOFF - ACUTE
ADLB1$AVALP <- (predict(fit, newdata = ADLB1) - ADLB1$EGFRBL)/CUTOFF
Desc <- tapply(ADLB1$AVALP, ADLB1$TRT, function(x) list(mean = mean(x), sd = sd(x)))
do.call(rbind, Desc)

```

```

##   mean      sd
## 0 -3.690067355 4.37486335
## 1 -2.545862177 4.328511519

```

```
ADLB1$AVALP <- round(ADLB1$AVALP, 2)
ADLB1 <- ADLB1[, c("ID", "AVALP")]
ADLB2 <- merge(ADLB1, EVNT, by = "ID", all.x = T)
ADLB2$AVAL <- ifelse(is.na(ADLB2$AVAL), ADLB1$AVALP, ADLB2$AVAL)
ADLB2$PARAMCD <- ifelse(is.na(ADLB2$PARAMCD), "eGFR", ADLB2$PARAMCD)
ADLB2$PARAM <- ifelse(is.na(ADLB2$PARAM), "eGFR slope", ADLB2$PARAM)
ADLB2$PARAMN <- ifelse(is.na(ADLB2$PARAMN), 7, ADLB2$PARAMN)
```

Note that although these individual GFR slope values are calculated for all patients, but they will be used only for those patients who did not get one of the dichotomous events described in Table 3. Then, we derive the PARAM and PARAMCD values for those patients, and the priority is set to 7, PARAMN=7. We will obtain the following dataset.

```
head(ADLB2)
```

Table 8: Predicted individual GFR slope values

ID	AVALP	AVAL	PARAM	PARAMCD	PARAMN	TRTPN
1	-3.03	-3.03	eGFR slope	eGFR	7	NA
2	1.75	1.75	eGFR slope	eGFR	7	NA
3	-1.90	-1.90	eGFR slope	eGFR	7	NA
4	-4.27	-4.27	eGFR slope	eGFR	7	NA
5	-2.29	-2.29	eGFR slope	eGFR	7	NA
6	-1.18	-1.18	eGFR slope	eGFR	7	NA

Kidney HCE dataset

Lastly, we derive the dataset for the kidney HCE based on predicted individual GFR values (if the patient did not experience a dichotomous outcome during the follow-up) or the study day of the most severe dichotomous event, if the patient experienced one of the outcomes 1-6 in Table 3.

```
HCE <- ADLB2[, c("ID", "AVAL", "PARAM", "PARAMCD", "PARAMN")]
names(HCE)[names(HCE) == "AVAL"] <- "AVALO"
names(HCE)[names(HCE) == "PARAM"] <- "GROUP"
```

In the code below we derive the column GROUPN which is the prioritization of outcomes and is meant to introduce order between categories and within each category. Since within each category 1-6 the maximum value is the length of fixed follow-up (since within these categories the study day of the most severe events is used), then we can multiply the length of the follow-up with the priority number. We do this since in the next step we will be adding the study day of the events to the GROUPN values to introduce a ranking within each category. A particular care is needed for the category 7. The values in this category can be negative hence adding to the value of GROUPN we may get lower values than the values in the category 6. Hence we need to make sure that the value for GROUPN in this category is large enough.

```
M <- floor(abs(max(HCE$AVALO[HCE$PARAMCD == "eGFR"]))) + 1
CUTOFF0 <- max(c(M, CUTOFF))
HCE$GROUPN <- CUTOFF*HCE$PARAMN
HCE$GROUPN[HCE$PARAMCD == "eGFR"] <- CUTOFF0*HCE$PARAMN[HCE$PARAMCD == "eGFR"]
HCE$AVAL <- HCE$AVALO + HCE$GROUPN

ADHCE <- merge(HCE, ADSL, by = "ID")
ADHCE$TRTP <- ifelse(ADHCE$TRTPN == 1, "A", "P")
```

```
head(ADHCE)
```

Table 9: Kidney HCE data structure

ID	AVAL0	GROUP	PARAMCD	PARAMNG	GROUPN	AVAL	TRTPN	EGFRBL	STRATAN	TRTP
1	-3.03	eGFR slope	eGFR		7	217	213.97	1	44	1 A
2	1.75	eGFR slope	eGFR		7	217	218.75	1	31	3 A
3	-1.90	eGFR slope	eGFR		7	217	215.10	1	28	4 A
4	-4.27	eGFR slope	eGFR		7	217	212.73	1	61	4 A
5	-2.29	eGFR slope	eGFR		7	217	214.71	1	37	4 A
6	-1.18	eGFR slope	eGFR		7	217	215.82	2	35	1 P

Win odds and maraca plots

In this section we calculate the win odds and its confidence interval and plot the kidney HCE using the *maraca* plot (Karpefors, Lindholm, and Gasparyan 2023).

```
res <- calcWO(x = ADHCE, AVAL = "AVAL", TRTP = "TRTP", ref = "P")
res0 <- res[, c("WO", "LCL", "UCL", "Pvalue")]
```

```
res0
```

Table 10: Win odds and its confidence interval

WO	LCL	UCL	Pvalue
1.32	1.1733	1.485	0

And the maraca plot can be created as follows:

```
hce_test <- maraca(
  data = ADHCE,
  tte_outcomes = c("DTHADJ", "DIAL90", "EGFR15", "EGFR57", "EGFR50", "EGFR40"),
  continuous_outcome = "eGFR",
  fixed_followup_days = CUTOFF,
  column_names = c(outcome = "PARAMCD", arm = "TRTP", value = "AVAL0"),
  arm_levels = c(active = "A", control = "P"),
  compute_win_odds = FALSE
)
```

```
plot(hce_test) + theme_classic() + xlab("") + ylab("") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, size = 7),
        legend.position = "bottom")
```

Acknowledgment

The synthetic datasets were created and kindly provided by the Analytics Data Preparation Team (ADAPT), Data Office, Data Science and AI, AstraZeneca.

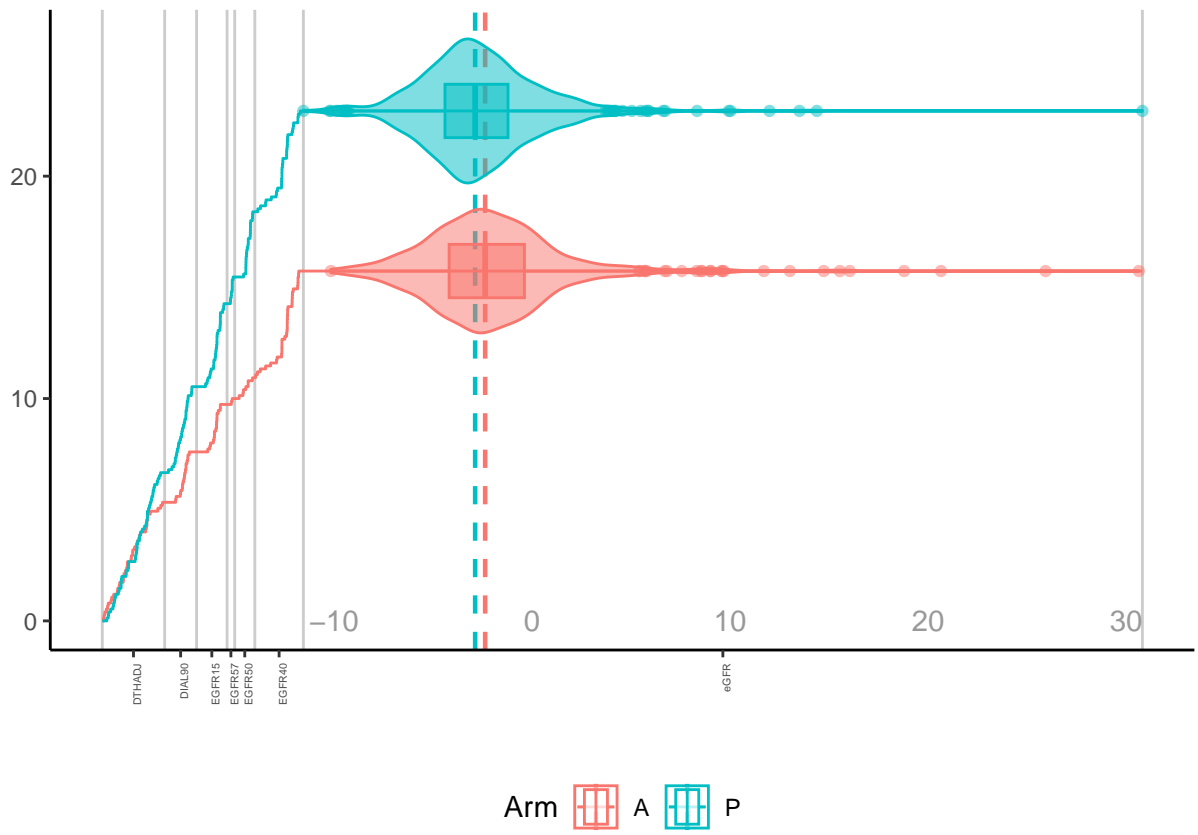


Figure 1: The maraca plot of kidney HCE

Appendix

```
library(readxl)
library(ggplot2)
library(nlme)
library(multcomp)
library(hce)
library(maraca)
library(dplyr)

ADSL <- read_excel("data/ADSL.xlsx")
ADET <- read_excel("data/ADET.xlsx")
ADLB <- read_excel("data/ADLB.xlsx")
CUTOFF <- 3
ACUTE <- 14/360
CHRONIC_coef <- (CUTOFF - ACUTE)/CUTOFF

ADLB$ADAY <- ADLB$ADAY/360
ADET$AVAL <- ADET$AVAL/360
ADLB <- ADLB[ADLB$ADAY <= CUTOFF, ]
EVNT <- ADET |> filter(AVAL <= CUTOFF) |> arrange(ID, PARAMN)|> group_by(ID) |>
  mutate(n = row_number()) |> ungroup() |> filter(n == 1) |> select(- c("n"))

table(EVNT$PARAM, EVNT$TRTPN)
ADLBO <- merge(ADLB, ADSL[, c("ID", "STRATAN", "EGFRBL")], by = "ID", all.x = T)
ADLBO$SPLINE <- ifelse(ADLBO$ADAY <= ACUTE, 0, ADLBO$ADAY - ACUTE)
ADLBO$TRT <- ifelse(ADLBO$TRTP == 1, 1, 0)

fit <- lme(AVAL ~ EGFRBL + STRATAN + TRT + ADAY + SPLINE + ADAY*TRT + SPLINE*TRT - 1,
  random = list(ID = pdSymm(form = ~ 1 + ADAY)),
  weights = varComb(varIdent(form = ~ 1|TRT),
    varPower(form = ~ 1 + ADAY)),
  na.action = na.omit,
  data = ADLBO,
  method = "REML",
  control = lmeControl(maxIter = 1e8,
    msMaxIter = 1e8,
    opt = "optim",
    optimMethod = "L-BFGS-B",
    msVerbose = T))

k <- ncol(coef(fit)) - 1
MTP <- glht(fit,
  linfct = rbind("Total: A - C" = c(rep(0, k - 2), 1, CHRONIC_coef)))
SLP <- confint(summary(MTP))
Slope <- SLP$confint[1:3]
Slope

MTP0 <- glht(fit,
  linfct = rbind("Total: A - C" = c(rep(0, k - 4), 1,
    CHRONIC_coef, 1, CHRONIC_coef)))
SLP0 <- confint(summary(MTP0))
Slope0 <- SLP0$confint[1:3]
Slope0
```

```

MTP1 <- glht(fit,
             linfct = rbind("Total: A - C" = c(rep(0, k - 4), 1, CHRONIC_coef, 0, 0)))
SLP1 <- confint(summary(MTP1))
Slope1 <- SLP1$confint[1:3]
Slope1

ADLB1 <- ADLBO[ADLBO$ADAY == 0, ]
ADLB1$ADAY <- CUTOFF
ADLB1$SPLINE <- CUTOFF - ACUTE
ADLB1$AVALP <- (predict(fit, newdata = ADLB1) - ADLB1$EGFRBL)/CUTOFF
Desc <- tapply(ADLB1$AVALP, ADLB1$TRT,
              function(x) list(mean = mean(x), sd = sd(x)))
do.call(rbind, Desc)

ADLB1$AVALP <- round(ADLB1$AVALP, 2)
ADLB1 <- ADLB1[, c("ID", "AVALP")]
ADLB2 <- merge(ADLB1, EVNT, by = "ID", all.x = T)
ADLB2$AVAL <- ifelse(is.na(ADLB2$AVAL), ADLB1$AVALP, ADLB2$AVAL)
ADLB2$PARAMCD <- ifelse(is.na(ADLB2$PARAMCD), "eGFR", ADLB2$PARAMCD)
ADLB2$PARAM <- ifelse(is.na(ADLB2$PARAM), "eGFR slope", ADLB2$PARAM)
ADLB2$PARAMN <- ifelse(is.na(ADLB2$PARAMN), 7, ADLB2$PARAMN)

HCE <- ADLB2[, c("ID", "AVAL", "PARAM", "PARAMCD", "PARAMN")]
names(HCE)[names(HCE) == "AVAL"] <- "AVALO"
names(HCE)[names(HCE) == "PARAM"] <- "GROUP"
M <- floor(abs(max(HCE$AVALO[HCE$PARAMCD == "eGFR"]))) + 1
CUTOFF0 <- max(c(M, CUTOFF))

HCE$GROUPN <- CUTOFF*HCE$PARAMN
HCE$GROUPN[HCE$PARAMCD == "eGFR"] <- CUTOFF0*HCE$PARAMN[HCE$PARAMCD == "eGFR"]
HCE$AVAL <- HCE$AVALO + HCE$GROUPN
ADHCE <- merge(HCE, ADSL, by = "ID")
ADHCE$TRTP <- ifelse(ADHCE$TRTPN == 1, "A", "P")
head(ADHCE)

res <- calcWO(x = ADHCE, AVAL = "AVAL", TRTP = "TRTP", ref = "P")
res0 <- res[, c("WO", "LCL", "UCL", "Pvalue")]
res0

hce_test <- maraca(data = ADHCE,
                  tte_outcomes = c("DTHADJ", "DIAL90", "EGFR15",
                                   "EGFR57", "EGFR50", "EGFR40"),
                  continuous_outcome = "eGFR",
                  fixed_followup_days = CUTOFF,
                  column_names = c(outcome = "PARAMCD", arm = "TRTP", value = "AVALO"),
                  arm_levels = c(active = "A", control = "P"),
                  compute_win_odds = FALSE)

plot(hce_test) + theme_classic() + xlab("") + ylab("") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, size = 7),
        legend.position = "bottom")

```

References

- Gasparian, Samvel B. 2022. *hce: Design and Analysis of Hierarchical Composite Endpoints*. <https://CRAN.R-project.org/package=hce>.
- Gasparian, Samvel B, Elaine K Kowalewski, Folke Folkvaljon, Olof Bengtsson, Joan Buenconsejo, John Adler, and Gary G Koch. 2021. “Power and Sample Size Calculation for the Win Odds Test: Application to an Ordinal Endpoint in COVID-19 Trials.” *Journal of Biopharmaceutical Statistics* 31 (6): 765–87.
- Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. “Simultaneous Inference in General Parametric Models.” *Biometrical Journal* 50 (3): 346–63. <https://CRAN.R-project.org/package=multcomp>.
- Karpefors, Martin, Daniel Lindholm, and Samvel B Gasparian. 2023. “The Maraca Plot: A Novel Visualization of Hierarchical Composite Endpoints.” *Clinical Trials* 20 (1): 84–88.
- Martin Karpefors, Samvel B. Gasparian, and Monika Huhn. 2023. *Maraca: The Maraca Plot: Visualization of Hierarchical Composite Endpoints in Clinical Trials*. <https://CRAN.R-project.org/package=maraca>.
- Pinheiro, José C, and Douglas M Bates. 2000. “Fitting Nonlinear Mixed-Effects Models.” *Mixed-Effects Models in S and S-PLUS*, 337–421. <https://CRAN.R-project.org/package=nlme>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing, Version 4.2.2*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Vonesh, Edward, Hocine Tighiouart, Jian Ying, Hiddo L Heerspink, Julia Lewis, Natalie Staplin, Lesley Inker, and Tom Greene. 2019. “Mixed-Effects Models for Slope-Based Endpoints in Clinical Trials of Chronic Kidney Disease.” *Statistics in Medicine* 38 (22): 4218–39.
- Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, and Jennifer Bryan. 2023. *readxl: Read Excel Files*. <https://CRAN.R-project.org/package=readxl>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.