Dataset Integrity Check for Effects of Nicotinamide and Lanthanum Carbonate on Serum Phosphate and Fibroblast Growth Factor-23 in CKD: The COMBINE Trial

Prepared by Dominick Parisi
IMS Inc.
3901 Calverton Blvd, Suite 200 Calverton, MD 20705
19 August 2020

Contents

1 Standard Disclaimer

The intent of this DSIC is to provide confidence that the data distributed by the NIDDK repository is a true copy of the study data. Our intent is not to assess the integrity of the statistical analyses reported by study investigators. As with all statistical analyses of complex datasets, complete replication of a set of statistical results should not be expected in secondary analysis. This occurs for a number of reasons including differences in the handling of missing data, restrictions on cases included in samples for a particular analysis, software coding used to define complex variables, etc. Experience suggests that most discrepancies can ordinarily be resolved by consultation with the study data coordinating center (DCC), however this process is labor-intensive for both DCC and Repository staff. It is thus not our policy to resolve every discrepancy that is observed in an integrity check. Specifically, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses, unless NIDDK Repository staff suspect that the observed discrepancy suggests that the dataset may have been corrupted in storage, transmission, or processing by repository staff. We do, however, document in footnotes to the integrity check those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

2 Study Background

COMBINE is a study to determine new methods to lower serum phosphorus levels and levels of FGF23 in chronic kidney disease patients.

3 Archived Datasets

All the SAS data files, as provided by the Data Coordinating Center (DCC), are located in the following datasets:

1. COMBINE/private_created_data/rand.sas7bdat
2. COMBINE/private_created_data/form107.sas7bdat
3. COMBINE/private_created_data/form115.sas7bdat
4. COMBINE/private_created_data/form122.sas7bdat
5. COMBINE/private_created_data/form203.sas7bdat
6. COMBINE/private_created_data/form216.sas7bdat
7. COMBINE/private_created_data/form364.sas7bdat
8. COMBINE/private_created_data/serum_data_2019.sas7bdat
9. COMBINE/private_created_data/pth_data_2019.sas7bdat
10. COMBINE/private_created_data/hr24_urine_data_2019.sas7bdat
11. COMBINE/private_created_data/spot_urine_data_2019.sas7bdat

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Ix et al [1] in the Journal of the American Society of Nephrology in June 2019. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Baseline characteristics by treatment arm in the COMBINE trial, Table A lists the variables that were used in the replication and Table B compares the results calculated from the archived data files to the results published in Table 1.

6 Conclusions

The NIDDK repository is confident that the COMBINE data files to be distributed are a true copy of the study data.

Some of the results do not match perfectly because there were multiple measurements from screening/baseline.

7 References

[1] Joachim H. Ix, Tamara Isakova, Brett Larive, Kalani L. Raphael, Dominic S. Raj, Alfred K. Cheung, Stuart M. Sprague, Linda F. Fried, Jennifer J. Gassman, John P. Middleton, Michael F. Flessner, Geoffrey A. Block, and Myles Wolf. J Am Soc nephrol. 2019 Jun; 30(6): 1096-1108.

Table A: Variables used to replicate Table 1: Baseline characteristics by treatment arm in the COMBINE trial

Table Variable	dataset.variable
Treatment arm	rand.treatment
Age (yr)	form107.age
Female	form107.gender
Race	form107.race
Diabetes	form122.diabetes
Systolic BP	form203.avg_sys
Diastolic BP	form203.avg_dia
BMI	form203.ht_cm,form203.wt_kg
eGFR (ml/min per 1.73 m^2)	form115.egfr
Urine ACR, g/g	spot_urine_data_2019.ualb, spot_urine_data_2019.ucr
Calcitrol use, n (%)	form216.vitd
Serum phosphate, mg/dL	form115.phos
Serum FGF23, pg/mL	form364.fgf23
Serum calcium, mg/dl	serum_data_2019.cal
Intact PTH, pg/mL	pth_data_2019.pth
24 hr urine phosphate, mg/d	hr24_urine_data.newuphos
24 hr urine phosphate-to-creatinine, mg/g	hr24_urine_data.newuphos,hr24_urine_data.newucr

Table B: Comparison of values computed in integrity check to reference article Table 1 values for entire cohort and N-L group.

	Entire Cohort (n=205) Manuscript	Entire Cohort (n=205) DSIC	Diff	N-L (n=53) Manuscript	N-L (n=53) DSIC	Diff
Age in years, mean ± SD	69 ± 12	66 ± 12	3 ± 0	69 ± 12	66 ± 12	3 ± 0
Female, n (%)	77 (38)	77 (38)	0 (0)	20 (38)	20 (38)	0 (0)
Race, n (%)						
• White	136 (66)	136 (66)	0 (0)	35 (66)	35 (66)	0 (0)
• Black	52 (25)	52 (25)	0 (0)	13 (25)	13 (25)	0 (0)
• Other	17 (8)	17 (8)	0 (0)	5 (9)	5 (9)	0 (0)
Diabetes, n (%)	112 (55)	112 (55)	0 (0)	31 (58)	31 (58)	0 (0)
SBP in mmHg, mean ± SD	129 ± 17	129 ± 18.9	0 ± 1.9	126 ± 15	130 ± 19.0	4 ± 4
DBP in mmHg, mean ± SD	71 ± 12	71 ± 12	0 ± 0	70 ± 11	72 ± 11	2 ± 0
Body Mass Index (kg/m^2), mean ± SD	31.9 ± 7.3	32.0 ± 7.4	0.1 ± 0.1	33.2 ± 8.5	33.2 ± 8.7	0 ± 0.2
eGFR (ml/min per 1.73 m^2), mean ± SD	32 ± 7	32 ± 7	0 ± 0	32 ± 7	32 ± 7.6	0 ± 0.6
Urine ACR, g/g, median (10%, 90%)	0.14 (0.01, 1.35)	0.104 (0.01, 1.30)	0.036 (0, 0.05)	0.14 (0.01, 1.09)	0.06 (0.01, 0.93)	0.08 (0, 0.16)
Calcitrol use, n (%)	41 (20)	41 (20)	0 (0)	8 (15)	8 (15)	0 (0)
Serum phosphate (mg/dL), mean ± SD	3.7 ± 0.6	3.7 ± 0.6	0 ± 0	3.6 ± 0.5	3.7 ± 0.6	0 ± 0
Serum FGF23 (pg/mL), median (10%, 90%)	99 (59, 205)	99 (56, 220)	0 (3, 15)	104 (57, 188)	96 (57, 215)[2]	8 (0, 27)
Serum calcium (mg/dl), mean ± SD	9.5 ± 0.5	9.5 ± 0.5	0 ± 0	9.5 ± 0.5	9.5 ± 0.5	0 ± 0
Intact PTH (pg/mL), mean ± SD	121 ± 85	121 ± 86.6	0 ± 1.6	121 ± 89	121 ± 82	0 ± 7
24 hr urine phosphate (mg/d), mean ± SD	668 ± 253	668 ± 253	0 ± 0	657 ± 265	657 ± 265	0 ± 0
24 hr urine phosphate-to-creatinine (mg/g), mean ± SD	462 ± 132	459 ± 127	3 ± 5	447 ± 146	450 ± 146	3 ± 0

Table B-1: Comparison of values computed in integrity check to reference article Table 1 values for N-p and p-L groups.

	N-p (n=51) Manuscript	N-p (n=51) DSIC	Diff	p-L (n=50) Manuscript	p-L (n=50) DSIC	Diff
Age in years, mean ± SD	70 ± 12	67 ± 12	3 ± 0	67 ± 13	64 ± 13	3 ± 0
Female, n (%)	17 (33)	17 (33)	0 (0)	22 (44)	22 (44)	0 (0)
Race, n (%)						
• White	37 (73)	37 (73)	0 (0)	32 (64)	32 (64)	0 (0)
• Black	9 (18)	9 (18)	0 (0)	14 (28)	14 (28)	0 (0)
• Other	5 (10)	5 (10)	0 (0)	4 (8)	4 (8)	0 (0)
Diabetes, n (%)	23 (45)	23 (45)	0 (0)	33 (66)	33 (66)	0 (0)
SBP in mmHg, mean ± SD	130 ± 15	130 ± 16.7	0 ± 1.7	129 ± 17	127 ± 18	4 ± 4
DBP in mmHg, mean ± SD	72 ± 13	70 ± 13	2 ± 0	71 ± 12	70 ± 11.4	2 ± 0
Body Mass Index (kg/m^2), mean ± SD	32.5 ± 8.1	32.4 ± 8.1	0.1 ± 0	30.9 ± 6.3	31.1 ± 6.4	0 ± 0.2
eGFR (ml/min per 1.73 m^2), mean ± SD	32 ± 7	32 ± 7	0 ± 0	33 ± 8	33 ± 8	0 ± 0.6
Urine ACR, g/g, median (10%, 90%)	0.06 (0.01, 1.42)	0.04 (0.01, 1.26)	0.02 (0, 0.16)	0.17 (0.01, 1.65)	0.17 (0.01,1.83)	0.08 (0, 0.16)
Calcitrol use, n (%)	13 (26)	13 (26)	0 (0)	10 (20)	10 (20)	0 (0)
Serum phosphate (mg/dL), mean ± SD	3.8 ± 0.5	3.7 ± 0.6	0.1 ± 0.1	3.7 ± 0.6	3.9 ± 0.7	0 ± 0
Serum FGF23 (pg/mL), median (10%, 90%)	104 (61, 192)	103 (61, 220)	1 (0, 28)	95 (49, 188)	95.6 (48.2, 195)	8 (0, 27)
Serum calcium (mg/dl), mean ± SD	9.5 ± 0.5	9.4 ± 0.4	0.1 ± 0.1	9.5 ± 0.5	9.5 ± 0.5	0 ± 0
Intact PTH (pg/mL), mean ± SD	116 ± 80	116 ± 84.6	0 ± 4.6	123 ± 76	124 ± 85	0 ± 7
24 hr urine phosphate (mg/d), mean ± SD	713 ± 235	713 ± 235	0 ± 0	659 ± 268	659 ± 268	0 ± 0
24 hr urine phosphate-to-creatinine (mg/g), mean ± SD	487 ± 126	479 ± 116	8 ± 10	452 ± 121	457 ± 119	5 ± 2

Table B-2: Comparison of values computed in integrity check to reference article Table 1 values for p-p group.

	p-p (n=51) Manuscript	p-p (n=51) DSIC	Diff
Age in years, mean ± SD	69 ± 10	66 ± 10	3 ± 0
Female, n (%)	18 (35)	18 (35)	0 (0)
Race, n (%)			
• White	32 (63)	32 (63)	0 (0)
• Black	16 (31)	16 (31)	0 (0)
• Other	3 (6)	3 (6)	0 (0)
Diabetes, n (%)	25 (49)	25 (49)	0 (0)
SBP in mmHg, mean ± SD	129 ± 21	130.5 ± 21.6	1.5 ± 0.6
DBP in mmHg, mean ± SD	72 ± 13	72 ± 11	0 ± 2
Body Mass Index (kg/m^2), mean ± SD	31.0 ± 6.1	31.1 ± 6.1	0.1 ± 0
eGFR (ml/min per 1.73 m^2), mean ± SD	32 ± 7	32 ± 7	0 ± 0
Urine ACR, g/g, median (10%, 90%)	0.17 (0.01, 1.29)	0.19 (0.01,1.42)	0.02 (0, 0.13)
Calcitrol use, n (%)	10 (20)	10 (20)	0 (0)
Serum phosphate (mg/dL), mean ± SD	3.6 ± 0.5	3.7 ± 0.5	0.1 ± 0
Serum FGF23 (pg/mL), median (10%, 90%)	96 (63, 225)	98 (62, 226)	2 (1, 1)
Serum calcium (mg/dl), mean ± SD	9.5 ± 0.5	**9.5** ± 0.5	0 ± 0
Intact PTH (pg/mL), mean ± SD	125 ± 95	123 ± 96.8	2 ± 1.8
24 hr urine phosphate (mg/d), mean ± SD	644 ± 245	644 ± 245	0 ± 0
24 hr urine phosphate-to-creatinine (mg/g), mean ± SD	463 ± 131	451 ± 126	12 ± 5

Attachment A: SAS Code

```
/***********************************************************************************************************************
   Program: /prj/niddk/ims_analysis/COMBINE/prog_initial_analysis/COMBINE.dsic.20200528.sas
***********************************************************************************************************************/
/***********/
/* Formats */
/***********/
PROC FORMAT;
  VALUE TREAT
    1 = 'N-L'
    2 = 'N-p'
    3 = 'p-L'
    4 = 'p-p';
  VALUE RACE
    5 = 'White'
    4 = 'Black'
    1,2,6,9 = 'Other';
  VALUE DIAB
    0 = 'No'
    1,2,3 = 'Yes';
RUN;


/*******************/
/* Import datasets */
/*******************/
LIBNAME SASDATA '/prj/niddk/ims_analysis/COMBINE/private_created_data/COMBINE_sas_data';

PROC SORT DATA=SASDATA.RAND OUT=RAND;
  BY PID;
RUN;

PROC SORT DATA=SASDATA.FORM107 OUT=FORM107;
  BY PID;
RUN;

PROC SORT DATA=SASDATA.FORM122 OUT=FORM122;
  BY PID;
RUN;

PROC SORT DATA=SASDATA.FORM115 OUT=FORM115;
  BY PID;
```

```
RUN;


PROC SORT DATA=SASDATA.FORM203 OUT=FORM203;
  BY PID;
RUN;

PROC SORT DATA=SASDATA.FORM216 OUT=FORM216_S0;
  BY PID;
  WHERE VIST='S';
RUN;

PROC SORT DATA=SASDATA.FORM362 OUT=FORM362;
  BY PID;
RUN;

PROC SORT DATA=SASDATA.FORM364 OUT=FORM364_B;
  BY PID COLL_DT;
  WHERE VISIT='B';
RUN;

PROC SORT DATA=SASDATA.SERUM_DATA_2019 OUT=SERUM_DATA_2019_BASE;
  BY PID VISIT_DT;
  WHERE VIST='B';
RUN;

PROC SORT DATA=SASDATA.PTH_DATA_2019 OUT=PTH_DATA_2019_BASE;
  BY PID VISIT_DT;
  WHERE VIST='B';
RUN;

PROC SORT DATA=SASDATA.HR24_URINE_DATA_2019 OUT=HR24_URINE_DATA_2019_BASE;
  BY PID VISIT_DT;
RUN;

PROC SORT DATA=SASDATA.SPOT_URINE_DATA_2019 OUT=SPOT_URINE_DATA_2019_BASE;
  BY PID VISIT_DT;
  WHERE VIST='B';

/*****************************/
/* Prepare datasets for merge */
/*****************************/
DATA FORM364_B;
  SET FORM364_B (RENAME=FGF23=FGF23_CHAR);
  BY PID COLL_DT;
  IF LAST.PID;
```

9

```
     FGF23=INPUT(FGF23_CHAR,8.);
  RUN;

  DATA SERUM_DATA_2019_BASE;
    SET SERUM_DATA_2019_BASE;
    BY PID VISIT_DT;
    IF FIRST.PID;
  RUN;

  DATA PTH_DATA_2019_BASE;
    SET PTH_DATA_2019_BASE;
    BY PID VISIT_DT;
    IF FIRST.PID;
  RUN;

  DATA HR24_URINE_DATA_2019_BASE;
    MERGE HR24_URINE_DATA_2019_BASE (IN=INHR24)
          RAND                      (IN=INRAND KEEP=PID RAND_DT);
    BY PID;
    IF INHR24 AND INRAND;

    IF DATEPART(RAND_DT)>=VISIT_DT;

    PHOS_PER_CREA=UPHOS/(UCR/1000);
  RUN;

  PROC SORT DATA=HR24_URINE_DATA_2019_BASE;
    BY PID DESCENDING CER DESCENDING VISIT_DT;
  RUN;

  DATA HR24_URINE_DATA_2019_BASE_BEST;
    SET HR24_URINE_DATA_2019_BASE;
    BY PID DESCENDING CER DESCENDING VISIT_DT;
    IF FIRST.PID;

    IF CER=0 THEN NEWUPHOS=.;
  RUN;

  DATA SPOT_URINE_DATA_2019_BASE;
    SET SPOT_URINE_DATA_2019_BASE;
    BY PID VISIT_DT;
    IF FIRST.PID;
    ACR=UALB/UCR;
  RUN;
```

```
PROC SORT DATA=FORM203 OUT=FORM203_S0;
  BY PID;
  WHERE INTENDED_VISIT='S0';
RUN;

PROC SORT DATA=FORM115 OUT=FORM115_EGFR;
  BY PID VISIT_DT;
RUN;

DATA FORM115_EGFR;
  SET FORM115_EGFR (WHERE=(EGFR^=.));
  BY PID VISIT_DT;
  IF FIRST.PID;
RUN;

/************************************************/
/* Create dataset to replicate manuscript numbers */
/************************************************/
DATA POPULATION;
  MERGE RAND     (IN=INRAND KEEP=PID TREATMENT)
         FORM107 (IN=INFORM107 KEEP=PID AGE GENDER RACE)
         FORM122 (IN=INFORM122 KEEP=PID DIABETES)
         FORM203_S0 (KEEP=PID HT_CM WT_KG AVG_SYS AVG_DIA)
         FORM203     (WHERE=(INTENDED_VISIT='B1') KEEP=PID INTENDED_VISIT AVG_DIA AVG_SYS)
         FORM115_EGFR (KEEP=PID EGFR PHOS)
         FORM216_S0   (KEEP=PID VITD)
         FORM364_B    (KEEP=PID FGF23)
         SERUM_DATA_2019_BASE (KEEP=PID CAL)
         PTH_DATA_2019_BASE    (KEEP=PID PTH)
         HR24_URINE_DATA_2019_BASE_BEST (KEEP=PID UPHOS UCR NEWUPHOS PHOS_PER_CREA)
         SPOT_URINE_DATA_2019_BASE (KEEP=PID UALB ACR UCR RENAME=UCR=SPOT_UCR);

  BY PID;
  IF INRAND AND INFORM107;

  BMI=WT_KG/((HT_CM/100)**2);

  IF TREATMENT='LANA' THEN TREATMENT_NUM=1;
  ELSE IF TREATMENT='LPNA' THEN TREATMENT_NUM=2;
  ELSE IF TREATMENT='LANP' THEN TREATMENT_NUM=3;
  ELSE IF TREATMENT='LPNP' THEN TREATMENT_NUM=4;
  FORMAT TREATMENT_NUM TREAT.;
RUN;

TITLE2 'Frequency tables and means to check table 1';
```

11

```
PROC MEANS DATA=POPULATION;
   VAR AGE;
RUN;

PROC MEANS DATA=POPULATION;
   VAR AGE;
   CLASS TREATMENT_NUM;
RUN;

PROC FREQ DATA=POPULATION;
   TABLE (GENDER RACE DIABETES) * TREATMENT_NUM;
   FORMAT RACE RACE. DIABETES DIAB.;
RUN;

PROC MEANS DATA=POPULATION;
   VAR AVG_SYS AVG_DIA BMI EGFR;
RUN;

PROC MEANS DATA=POPULATION;
   VAR AVG_SYS AVG_DIA BMI EGFR;
   CLASS TREATMENT_NUM;
RUN;

PROC MEANS DATA=POPULATION MEDIAN P10 P90;
   var ACR;
RUN;

PROC MEANS DATA=POPULATION MEDIAN P10 P90;
   var ACR;
   CLASS TREATMENT_NUM;
RUN;

PROC FREQ DATA=POPULATION;
   TABLE VITD * TREATMENT_NUM /MISSING;
RUN;

PROC MEANS DATA=POPULATION;
   VAR PHOS;
RUN;

PROC MEANS DATA=POPULATION;
   VAR PHOS;
   CLASS TREATMENT_NUM;
RUN;
PROC MEANS DATA=POPULATION MEDIAN P10 P90;
```

```
    var FGF23;
RUN;

PROC MEANS DATA=POPULATION MEDIAN P10 P90;
  var FGF23;
  CLASS TREATMENT_NUM;
RUN;

PROC MEANS DATA=POPULATION;
  VAR CAL PTH UPHOS NEWUPHOS PHOS_PER_CREA;
RUN;

PROC MEANS DATA=POPULATION;
  VAR CAL PTH UPHOS NEWUPHOS PHOS_PER_CREA;
  CLASS TREATMENT_NUM;
RUN;
```