

Dataset Integrity Check for A
Randomized Trial Comparing the Safety,
Adherence, and Pharmacodynamics
Profiles of Two Doses of Sodium
Bicarbonate in CKD: the BASE Pilot Trial

Prepared by Dominick Parisi

IMS Inc.

3901 Calverton Blvd, Suite 200 Calverton, MD 20705

14 August 2020

Contents

1 Standard Disclaimer	2
2 Study Background	2
3 Archived Datasets	2
4 Statistical Methods	2
5 Results	2
6 Conclusions	3
7 References	3
Table A: Variables used to replicate Table 1: Baseline clinical and demographic characteristics of randomized participants overall and by treatment assignment	4
Table B: Comparison of values computed in integrity check to reference article Table 1 values.....	5
Attachment A: SAS Code	9

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

Sodium bicarbonate may help prevent kidney failure in people with chronic kidney disease. However, the dose to prescribe in order to test this possibility in a clinical trial is uncertain. The Bicarbonate Administration to Stabilize Estimated Glomerular Filtration Rate (BASE) clinical trial assessed the best dose of sodium bicarbonate to prescribe in a future study that will test the long-term safety and efficacy of sodium bicarbonate as a treatment to preserve kidney function in individuals with chronic kidney disease.

3 Archived Datasets

All the SAS data files, as provided by the Data Coordinating Center (DCC), are located in the datasets listed in Table A below.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Raphael et al [1] in the Journal of the American Society of Nephrology in January 2020. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Baseline clinical and demographic characteristics of randomized participants overall and by treatment assignment, 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

Results showed minor discrepancies due to multiple measures from screening/baseline and found to be within expected limits. The NIDDK repository is confident that the BASE data files to be distributed are a true copy of the study data.

7 References

[1] Kalani L Raphael, Tamara Isakova, Joachim H Ix, Dominic S Raj, Myles Wolf, Linda F Fried, Jennifer J Gassman, Cynthia Kendrick, Brett Larive, Michael F Flessner, Susan R Mendley, Thomas H Hostetter, Geoffrey A Block, Ping Li, John P Middleton, Stuart M Sprague, Donald E Wesson, Alfred K Cheung. J Am Soc Nephrol, 2020 Jan;31(1):161-174. doi: 10.1681/ASN.2019030287. Epub 2019 Dec 17.

Table A: Variables used to replicate Table 1: Baseline clinical and demographic characteristics of randomized participants overall and by treatment assignment

Table Variable	dataset.variable
Age (yr)	form106_base.age
Female	form106.gender
Race	form106.race
Ethnicity	form106.ethnicity
CKD stage	form351.egfr
Diabetes mellitus	form122.diabetes
Heart disease	form122.mi, form122.chf, form122.angina, form122.revascularization, form122.ischemia, form122.stroke
CHF	form122.chf
COPD	form122.copd
Smoking status	form123.smoke
Weight (kg)	form202.wt_kg
Body Mass Index (kg/m ²)	form202.wt_kg, form202.ht_cm
Lean body weight (kg)	form202.lbw_kg
SBP (mmHg)	form202.avg_sys
ACE-i/ARB use (no. [%])	form202.acei_arb_optimized
eGFR (ml/min per 1.73 m ²)	form351.egfr
Urinary albumin/creatinine (mg/g)	form356.alb_creat_ratio
Serum bicarbonate (meq/L)	form351.bicarbonate
Serum potassium (meq/L)	form351.potassium
Urinary ammonium (meq/L)	form355.ammonium, form355.est_vol, form326.est_vol, form326.est_vol_added
Urinary pH	form355.ph
Estimated dietary protein intake (g/d)	form355.urea_nitrogen, form355.est_vol, form326.est_vol, form326.est_vol_added, form326.total_hrs

Table B: Comparison of values computed in integrity check to reference article Table 1 values.

	Entire Cohort (n=194) Manuscript	Entire Cohort (n=194) DSIC	Diff (n=0)	HD-NaHCO3 (n=90) Manuscript	HD-NaHCO3 (n=90) DSIC	Diff (n=0)
Age (yr)	67 ± 12	66 ± 12	1 ± 0	67 ± 11	67 ± 11	0 ± 0
Female (no. [%])	63 (33)	63 (33)	0 (0)	29 (32)	29 (32)	0 (0)
Race (no. [%])						
• Black	67 (35)	67 (35)	0 (0)	27 (30)	27 (30)	0 (0)
• White	109 (56)	109 (56)	0 (0)	55 (61)	55 (61)	0 (0)
• Other	15 (8)	15 (8)	0 (0)	7 (8)	7 (8)	0 (0)
• Unknown/not reported	3 (2)	3 (2)	0 (0)	1 (1)	1 (1)	0 (0)
Ethnicity (no. [%])						
• Hispanic/Latino	17 (9)	17 (9)	0 (0)	8 (9)	8 (9)	0 (0)
CKD stage (no. [%])*						
• 3A	30 (15.5)	30 (15.6)	0 (0.1)	17 (18.9)	17 (19.3)	0 (0.4)
• 3B	107 (55.2)	107 (55.7)	0 (0.5)	46 (51.1)	46 (52.3)	0 (1.2)
• 4	57 (29.4)	55 (28.7)	2 (0.7)	27 (30.0)	25 (28.4)	2 (1.6)
Diabetes mellitus (no. [%])	105 (54)	105 (54)	0 (0)	53 (59)	53 (59)	0 (0)
Heart disease (no. [%])	66 (34)	66 (34)	0 (0)	28 (31)	28 (31)	0 (0)
CHF (no. [%])	20 (10)	20 (10)	0 (0)	8 (9)	8 (9)	0 (0)
COPD (no. [%])	10 (5)	10 (5)	0 (0)	6 (7)	6 (7)	0 (0)
Smoking Status (no. [%])						
• Never	98 (51)	98 (51)	0 (0)	47 (52)	47 (52)	0 (0)
• Former	79 (41)	79 (41)	0 (0)	39 (43)	39 (43)	0 (0)
• Current	17 (9)	17 (9)	0 (0)	4 (4)	4 (4)	0 (0)
Weight (kg)	95.1 ± 22.7	95.1 ± 22.8	0 ± 0.1	95.4 ± 22.4	95.5 ± 22.5	0.1 ± 0.1
Body Mass Index (kg/m ²)	32.5 ± 7.5	32.3 ± 7.0	0.2 ± 0.5	32.8 ± 7.9	32.5 ± 6.7	0.3 ± 1.2
Lean body weight (kg)	58.5 ± 9.9	60.4 ± 11.9	1.9 ± 2.0	58.4 ± 9.8	60.3 ± 11.8	1.9 ± 2.0

	Entire Cohort (n=194) Manuscript	Entire Cohort (n=194) DSIC	Diff (n=0)	HD-NaHCO3 (n=90) Manuscript	HD-NaHCO3 (n=90) DSIC	Diff (n=0)
SBP (mmHg)	127 ± 13	127 ± 13.9	0 ± 0.9	127 ± 12	127 ± 13.3	0 ± 1.3
ACE-i/ARB use (no. [%])	135 (70)	134 (69)	1 (1)	60 (67)	59 (66)	1 (1)
eGFR (ml/min per 1.73 m ²)	36 ± 9	36 ± 9	0 ± 0	36 ± 10	36.8 ± 9.4	0.8 ± 0.6
Urinary albumin/creatinine (mg/g)	181 (25, 745)	181 (25, 745)	0 (0, 0)	176 (23, 769)	176 (23, 769)	0 (0, 0)
Serum bicarbonate (meq/L)	24 ± 2	24 ± 2	0 ± 0	24 ± 2	24 ± 2	0 ± 0
Serum potassium (meq/L)	4.5 ± 0.4	4.4 ± 0.46	0.1 ± 0.06	4.4 ± 0.4	4.4 ± 0.46	0.1 ± 0.06
Urinary ammonium (meq/L)	21 ± 12	21 ± 12	0 ± 0	22 ± 12	22 ± 12	0 ± 0
Urinary pH	5.8 ± 0.5	5.8 ± 0.5	0 ± 0	5.8 ± 0.5	5.8 ± 0.5	0 ± 0
Estimated dietary protein intake (g/d)	79 ± 29	79 ± 29	0 ± 0	82 ± 31	82 ± 30.44	0 ± 0.56

*CKD Stage for HD-NaHCO3 group for the integrity check is based on n= 88 total count

Table B (Continued): Comparison of values computed in integrity check to reference article Table 1 values.

	LD-NaHCO3 (n=52) Manuscript	LD-NaHCO3 (n=52) DSIC	Diff (n=0)	Placebo (n=52) Manuscript	Placebo (n=52) DSIC	Diff (n=0)
Age (yr)	66 ± 14	65.25 ± 14	0.75 ± 0	66 ± 11	66 ± 11	0 ± 0
Female (no. [%])	21 (40)	21 (40)	0 (0)	13 (25)	13 (25)	0 (0)
Race/ethnicity (no. [%])						
• Black	17 (33)	17 (33)	0 (0)	23 (44)	23 (44)	0 (0)
• White	28 (54)	28 (54)	0 (0)	26 (50)	26 (50)	0 (0)
• Other	5 (10)	5 (10)	0 (0)	3 (6)	3 (6)	0 (0)
• Unknown/not reported	2 (4)	2 (4)	0 (0)	0 (0)	0 (0)	0 (0)
Ethnicity (no. [%])						
• Hispanic/Latino	6 (12)	6 (12)	0 (0)	3 (6)	3 (6)	0 (0)
CKD stage (no. [%])						
• 3A	9 (17.3)	9 (17.3)	0 (0.0)	4 (7.7)	4 (7.7)	0 (0.0)
• 3B	30 (57.7)	30 (57.7)	0 (0.0)	31 (59.6)	31 (59.6)	0 (0.0)
• 4	13 (25.0)	13 (25.0)	0 (0.0)	17 (32.7)	17 (32.7)	0 (0.0)
Diabetes mellitus (no. [%])	27 (52)	27 (52)	0 (0)	25 (48)	25 (48)	0 (0)
Heart disease (no. [%])	18 (35)	18 (35)	0 (0)	20 (39)	20 (39)	0 (0)
CHF (no. [%])	5 (10)	5 (10)	0 (0)	7 (14)	7 (14)	0 (0)
COPD (no. [%])	4 (8)	4 (8)	0 (0)	0 (0)	0 (0)	0 (0)
Smoking Status (no. [%])						
• Never	26 (50)	26 (50)	0 (0)	25 (48)	25 (48)	0 (0)
• Former	15 (29)	15 (29)	0 (0)	25 (48)	25 (48)	0 (0)
• Current	11 (21)	11 (21)	0 (0)	2 (4)	2 (4)	0 (0)
Weight (kg)	96.1 ± 25.8	96.1 ± 26.0	0 ± 0.2	93.6 ± 20.3	93.6 ± 20.3	0 ± 0
Body Mass Index (kg/m ²)	32.8 ± 8.4	32.9 ± 8.5	0.1 ± 0.1	31.6 ± 5.9	31.5 ± 6.0	0.1 ± 0.1
Lean body weight (kg)	58.4 ± 10.7	59.7 ± 12.7	1.3 ± 2.0	58.7 ± 9.4	61.2 ± 11.6	2.5 ± 2.2

	LD-NaHCO3 (n=52) Manuscript	LD-NaHCO3 (n=52) DSIC	Diff (n=0)	Placebo (n=52) Manuscript	Placebo (n=52) DSIC	Diff (n=0)
SBP (mmHg)	127 ± 14	126 ± 14.8	1 ± 0.8	127 ± 13	126 ± 14	1 ± 1.0
ACE-i/ARB use (no. [%])	40 (77)	40 (77)	0 (0)	35 (67)	35 (67)	0 (0)
eGFR (ml/min per 1.73 m ²)	37 ± 10	37 ± 10	0.0 ± 0.0	35 ± 9	35 ± 9	0.0 ± 0.0
Urinary albumin/creatinine (mg/g)	222 (58, 788)	222 (58, 788)	0 (0, 0)	147 (12, 593)	147 (12, 593)	0 (0, 0)
Serum bicarbonate (meq/L)	24 ± 2	24 ± 2	0 ± 0	24 ± 2	24 ± 2	0 ± 0
Serum potassium (meq/L)	4.5 ± 0.4	4.5 ± 0.4	0 ± 0	4.5 ± 0.5	4.5 ± 0.5	0 ± 0
Urinary ammonium (meq/L)	19 ± 11	19 ± 11	0 ± 0	21 ± 13	21 ± 13	0 ± 0
Urinary pH	5.9 ± 0.5	5.9 ± 0.5	0 ± 0	5.8 ± 0.6	5.8 ± 0.6	0 ± 0
Estimated dietary protein intake (g/d)	73 ± 25	73 ± 25	0 ± 0	79 ± 30	79 ± 30	0 ± 0

Attachment A: SAS Code

```

/*****
Program: /prj/niddk/ims_analysis/BASE/prog_initial_analysis/BASE.dsic.20200527.sas
*****/

/*****/
/* Formats */
/*****/
PROC FORMAT;
  VALUE $TREAT
    'HA' = 'HD-NaHCO3'
    'LA' = 'LD-NaHCO3'
    'HP','LP' = 'Placebo';
  VALUE SEX
    1 = 'Male'
    2 = 'Female';
  VALUE RACE
    1,2,3,6 = 'Other'
    4       = 'Black'
    5       = 'Caucasian'
    9       = 'Unknown/Not Reported';
  VALUE DIABETES
    0 = 'No'
    1,2,3 = 'Yes';
  VALUE SMOKE
    0 = '0: Never'
    1 = '1: Former'
    2 = '2: Current';
  VALUE ACEI
    1,2 = 'ACE/ARB use'
    4,5 = 'No ACE/ARB use';
RUN;

/*****/
/* Filename statements */
/*****/
LIBNAME SASDATA '/prj/niddk/ims_analysis/BASE/private_created_data/BASE_sas_data';

```

```

/*****/
/* Import datasets */
/*****/
PROC SORT DATA=SASDATA.RAND_BASE OUT=RAND_BASE;
  BY PID;
RUN;

PROC SORT DATA=SASDATA.FORM106_BASE OUT=FORM106_BASE;
  BY PID;
RUN;

PROC SORT DATA=SASDATA.FORM122_BASE OUT=FORM122_BASE;
  BY PID;
RUN;

PROC SORT DATA=SASDATA.FORM123_BASE OUT=FORM123_BASE;
  BY PID;
RUN;

PROC SORT DATA=SASDATA.FORM202_BASE OUT=FORM202_BASE;
  BY PID VIST INTENDED_VISIT VISIT_DT;
RUN;

PROC SORT DATA=SASDATA.FORM326_BASE OUT=FORM326_BASE;
  BY PID INTENDED_VISIT;
RUN;

PROC SORT DATA=SASDATA.FORM351_BASE OUT=FORM351_BASE;
  BY PID INTENDED_VISIT;
RUN;

PROC SORT DATA=SASDATA.FORM355_BASE OUT=FORM355_BASE_B;
  BY PID VISIT_DT;
  where vist='B';
RUN;

PROC SORT DATA=SASDATA.FORM356_BASE OUT=FORM356_BASE;
  BY PID VISIT_DT;
RUN;

```

```

/*****
/* Subset data from certain visits for forms with multiple records */
*****/
DATA FORM351_BASE_LAST_S;
  SET FORM351_BASE (WHERE=(VIST='S'));
  BY PID INTENDED_VISIT;
  IF LAST.PID;
RUN;

DATA FORM355_BASE_LAST_B;
  SET FORM355_BASE_B;
  BY PID VISIT_DT;
  IF LAST.PID;
RUN;

DATA FORM356_BASE_LAST_BASE;
  SET FORM356_BASE (WHERE=(VIST IN('S','B')));
  BY PID VISIT_DT;
  IF LAST.PID;
RUN;

DATA FORM202_BASE_LAST_BASE;
  SET FORM202_BASE (WHERE=(VIST='B'));
  BY PID VIST INTENDED_VISIT;
  IF LAST.PID;
RUN;

DATA FORM202_BASE_S0;
  SET FORM202_BASE (WHERE=(VIST='S'));
  BY PID VIST INTENDED_VISIT VISIT_DT;
  IF LAST.PID;
RUN;

DATA FORM326_BASE_LAST_B;
  SET FORM326_BASE (WHERE=(INTENDED_VISIT='B1'));
  BY PID INTENDED_VISIT;
  IF LAST.PID;

  VOLUME_SUM=SUM(EST_VOL,EST_VOL_ADD);
RUN;

```

```

/*****
/* Create dataset to use for paper checking */
*****/
DATA ANALYSIS;
  MERGE RAND_BASE              (IN=INRAND KEEP=PID TREATMENT)
        FORM106_BASE           (IN=IN106 KEEP=PID AGE GENDER RACE ETHNICITY)
        FORM122_BASE           (IN=IN122 KEEP=PID DIABETES CHF COPD MI ANGINA REVASCULARIZATION ISCHEMIA STROKE)
        FORM123_BASE           (IN=IN123 KEEP=PID SMOKE)
        FORM202_BASE_LAST_BASE (KEEP=PID WT_KG LBW_KG AVG_SYS ACEI_ARB_OPTIMIZED)
        FORM202_BASE_S0        (KEEP=PID HT_CM RENAME=HT_CM=HT_CM_S0)
        FORM326_BASE_LAST_B    (KEEP=PID VOLUME_SUM TOTAL_HRS)
        FORM351_BASE_LAST_S    (KEEP=PID EGFR BICARBONATE POTASSIUM)
        FORM355_BASE_LAST_B    (KEEP=PID EST_VOL AMMONIUM PH UREA_NITROGEN)
        FORM356_BASE_LAST_BASE (KEEP=PID ALB_CREAT_RATIO);
  BY PID;
  IF INRAND AND IN106 AND IN122 AND IN123;

  BMI=WT_KG/(HT_CM_S0/100)**2;

  *** Heart disease (code provided by Cynthia Kendrick);
  IF MI=1 OR CHF=1 OR ANGINA=1 OR REVASCULARIZATION=1 OR ISCHEMIA=1 OR STROKE=1 THEN HX_HEART=1;
  ELSE IF MI=0 AND CHF=0 AND ANGINA=1 AND REVASCULARIZATION=0 AND ISCHEMIA=0 AND STROKE=0 THEN HX_HEART=0;
  ELSE HX_HEART=.;

  *** CKD stage (code provided by Cynthia Kendrick);
  IF 15<=EGFR<30 THEN CKD_STAGE=40;
  ELSE IF 30<=EGFR<45 THEN CKD_STAGE=32;
  ELSE IF 45<=EGFR<60 THEN CKD_STAGE=31;

  *** Volume of urine for 24 hour test (code provided by Cynthia Kendrick);
  IF EST_VOL^=. AND VOLUME_SUM^=. AND ((EST_VOL<(0.667*VOLUME_SUM)) OR (EST_VOL>(1.333*VOLUME_SUM))) THEN VOLUME=VOLUME_SUM;
  ELSE IF EST_VOL=. AND VOLUME_SUM^=. THEN VOLUME=VOLUME_SUM;
  ELSE VOLUME=EST_VOL;

  *** Ammonium meq/day value (code provided by Cynthia Kendrick);
  AMMONIUM_24=AMMONIUM*(VOLUME/1000);

  *** Dietary protein intake (code provided by Cynthia Kendrick);
  NITROGEN_24=(UREA_NITROGEN/100)*(VOLUME);
  IF TOTAL_HRS<23 OR TOTAL_HRS>25 THEN NITROGEN=(NITROGEN_24*24)/TOTAL_HRS;
  ELSE NITROGEN=NITROGEN_24;
  NITROGEN_G_24=NITROGEN/1000;
  PROTEIN=6.25*(NITROGEN_G_24+(WT_KG*.031));
RUN;

```

```

TITLE2 'Frequency tables and means to check numbers in Table 1';
PROC FREQ DATA=ANALYSIS ORDER=FORMATTED;
  TABLE TREATMENT;
  FORMAT TREATMENT $TREAT.;
RUN;

PROC MEANS DATA=ANALYSIS;
  VAR AGE;
RUN;

PROC MEANS DATA=ANALYSIS ORDER=FORMATTED;
  VAR AGE;
  CLASS TREATMENT;
  FORMAT TREATMENT $TREAT.;
RUN;

PROC FREQ DATA=ANALYSIS ORDER=FORMATTED;
  TABLE (GENDER RACE ETHNICITY) * TREATMENT /MISSING;
  TABLE CKD_STAGE * TREATMENT;
  TABLE (DIABETES HX_HEART CHF COPD SMOKE) * TREATMENT /MISSING;
  FORMAT TREATMENT $TREAT. GENDER SEX. RACE RACE. DIABETES DIABETES. SMOKE SMOKE.;
RUN;

PROC MEANS DATA=ANALYSIS;
  VAR WT_KG BMI LBW_KG AVG_SYS;
RUN;

PROC MEANS DATA=ANALYSIS ORDER=FORMATTED;
  VAR WT_KG BMI LBW_KG AVG_SYS;
  CLASS TREATMENT;
  FORMAT TREATMENT $TREAT.;
RUN;

PROC FREQ DATA=ANALYSIS ORDER=FORMATTED;
  TABLE ACEI_ARB_OPTIMIZED * TREATMENT;
  FORMAT ACEI_ARB_OPTIMIZED ACEI. TREATMENT $TREAT.;
RUN;

PROC MEANS DATA=ANALYSIS;
  VAR EGFR BICARBONATE POTASSIUM AMMONIUM_24 PH PROTEIN;
RUN;

```

```
PROC MEANS DATA=ANALYSIS ORDER=FORMATTED;  
  VAR EGFR BICARBONATE POTASSIUM AMMONIUM_24 PH PROTEIN;  
  CLASS TREATMENT;  
  FORMAT TREATMENT $TREAT. ;  
RUN;
```

```
PROC MEANS DATA=ANALYSIS MEDIAN P25 P75;  
  VAR ALB_CREAT_RATIO;  
RUN;
```

```
PROC MEANS DATA=ANALYSIS MEDIAN P25 P75 ORDER=FORMATTED;  
  VAR ALB_CREAT_RATIO;  
  CLASS TREATMENT;  
  FORMAT TREATMENT $TREAT. ;  
RUN;
```