Norma Pugh
May 11, 2010

Integrity Check for the Acute Renal Failure Trial Network Study (ATN) Data Files

As a partial check of the integrity of the ATN datasets archived in the NIDDK data repository, a set of tabulations was performed to verify that published results can be reproduced using the archived datasets. Analyses were performed to duplicate results for the data published by Palevsky et al [1] in *The New England Journal of Medicine* in July 2008. The results of this dataset integrity check (DSIC) are described below. The full text of the *New England Journal of Medicine* article can be found in Attachment 1, and the SAS code for the tabulations is included in Attachment 2. Attachment 3 includes the calculation of select study variables, as provided by the study data coordinating center. Attachment 4 contains clarifications provided by the DCC for minor discrepancies noted in the DSIC.

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 on a first (or second) exercise 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.

Background. The Veterans Affairs/National Institutes of Health (VA/NIH) Acute Renal Failure Trial Network (ATN) study was a multi-center, prospective, randomized, parallel-group trial of two strategies for renal-replacement therapy in critically ill patients with acute kidney injury [1].

The ATN consists of twenty-seven VA and university-affiliated medical centers and was designed to test the hypothesis that an intensive renal-replacement therapy would decrease mortality in critically ill patients as compared to conventional renal-replacement therapy [2].

Baseline Characteristics. Table 1 in the publication [1] reports on baseline characteristics of patients enrolled in the ATN Trial. Table A lists the variables we used in our replication. All variables were taken from the SAS datasets, d_analysisdemo and d_form01.

Table A: Variables Used to Replicate Table 1

Table Variable	Variables Used in Replication
Sample size	TreatmentGp
Age	Age
Sex	Gender
Race or ethnic group	Race
Renal function before onset of acute kidney injury: Serum creatinine	Form01PreMorbidCreat, Form01BaselineCreat, Gender
Renal function before onset of acute kidney injury: Estimated GFR	Form01PreMorbidCreat, Form01BaselineCreat, Gender, Race, Age
Cause of acute kidney injury: Ischemia	Ischemic
Cause of acute kidney injury: Nephrotoxins	Nephrotoxic
Cause of acute kidney injury: Sepsis	Sepsis
Cause of acute kidney injury: Multifactorial causes	Multifactorial
Primary treating service	PrimaryTreat
Weight before acute illness	Weight
Length of stay before randomization: Hospital	Dayshospbrand
Length of stay before randomization: ICU	DaysICUbrand
Charlson comorbidity index	CharlsonAgeScore
Mechanical ventilation	MechVentilation
Sepsis	d_form01: Sepsis
APACHE II score	bTotApacheII
SOFA organ system score: Respiratory	bRespiratory
SOFA organ system score: Coagulation	bCoagulation
SOFA organ system score: Liver	bLiver
SOFA organ system score: Cardiovascular	bCardiovascular
SOFA organ system score: Central nervous system	bCentralNerve
SOFA organ system score: Total	bTotalSOFA
Cleveland Clinic ICU Acute Renal Failure score	BTotCCFARF
BUN at initiation of RRT	PriorRandBUN
Cardiovascular SOFA score, categorical	CardioSOFA
Oliguria	Olig
One session of IHD or SLED or <24 hr of continuous RRT before randomization	HemoCRRT

In Table B, we compare the results calculated from the archived datasets to the results published in Table 1, Baseline Characteristics of the Study Patients. As Table B shows, the results are similar.

Table B: Comparison of Values Computed in Integrity Check to Reference Article Table 1 Values, Treatment = Intensive Strategy

Characteristic	Palevsky et al	Integrity Check	Diff
Sample size	563	563	0
Age, years	59.6 ± 15.3	59.6 ± 15.3	0
Sex			
Male	409/563 (72.6)	409/563 (72.6)	0
Female	154/563 (27.4)	154/563 (27.4)	0
Race or ethnic group			
White	415/563 (73.7)	415/563 (73.7)	0
Black	91/563 (16.2)	91/563 (16.2)	0
Hispanic	44/563 (7.8)	44/563 (7.8)	0
Other	13/563 (2.3)	13/563 (2.3)	0
Renal function before onset of acute kidney injury			
Serum creatinine, mg/dl	1.1 ± 0.4	1.1 ± 0.4	0
Estimated GFR			
≥60 ml/min/1.73 m^2	342/524 (65.3)	342/524 (65.3)	0
45-59 ml/min/1.73 m^2	122/524 (23.3)	122/524 (23.3)	0
30-44 ml/min/1.73 m^2	60/524 (11.5)	60/524 (11.5)	0
Cause of acute kidney injury			
Ischemia	440/536 (82.1)	440/536 (82.1)	0
Nephrotoxins	143/514 (27.8)	143/514 (27.8)	0
Sepsis	300/531 (56.5)	300/531 (56.5)	0
Multifactorial causes	317/534 (59.4)	317/534 (59.4)	0
Primary treating service			
Medical	272/563 (48.3)	272/563 (48.3)	0
Surgical	229/563 (40.7)	229/563 (40.7)	0
Other	62/563 (11.0)	62/563 (11.0)	0
Weight before acute illness, kg	84.1 ± 19.6	84.1 ± 19.6	0
Length of stay before randomization, days			
Hospital	11.1 ± 13.6	11.1 ± 13.6	0
ICU	6.9 ± 10.1	6.9 ± 10.1	0
Charlson comorbidity index	4.3 ± 3.0	4.3 ± 3.0	0
Mechanical ventilation	453/563 (80.5)	453/563 (80.5)	0
Sepsis	358 (63.6)	358 (63.6)	0
APACHE II score	26.6 ± 7.2	26.6 ± 7.2	0
SOFA organ-system score			
Respiratory	2.4 ± 1.1	2.4 ± 1.1	0
Coagulation	1.4 ± 1.2	1.4 ± 1.2	0
Liver	1.5 ± 1.3	1.5 ± 1.3	0
Cardiovascular	2.3 ± 1.7	2.3 ± 1.7	0
Central nervous system	2.5 ± 1.4	2.5 ± 1.4	0
Total	14.7 ± 3.7	14.7 ± 3.7	0

Table B (continued): Comparison of Values Computed in Integrity Check to Reference Article Table 1 Values, Treatment = Intensive Strategy

Characteristic	Palevsky et al	Integrity Check	Diff
Cleveland Clinic ICU Acute Renal Failure score	12.3 ± 3.3	12.3 ± 3.3	0
BUN at initiation of RRT	65.9 ± 30.2	65.9 ± 30.2	0
Cardiovascular SOFA score			
0-2	255 (45.3)	251 (44.7)	4, 0.6
3-4	308 (54.7)	311 (55.3)	3, 0.6
Oliguria			
No	124 (22.0)	124 (22.0)	0
Yes	439 (78.0)	439 (78.0)	0
One session of IHD or SLED of <24 hr of continuous RRT before randomization	358/563 (63.6)	358/563 (63.6)	0

Notes:
(1) GFR glomerular filtration rate, IHD intermittent hemodialysis, RRT renal-replacement therapy, SLED sustained low-efficiency dialysis

Table B (continued): Comparison of Values Computed in Integrity Check to Reference Article Table 1 Values, Treatment = Less-Intensive Strategy

Characteristic	Palevsky et al	Integrity Check	Diff
Sample size	561	561	0
Age, years	59.7 ± 15.2	59.7 ± 15.2	0
Sex			
Male	384/560 (68.6)	384/560 (68.6)	0
Female	176/560 (31.4)	176/560 (31.4)	0
Race or ethnic group			
White	420/560 (75.0)	420/560 (75.0)	0
Black	88/560 (15.7)	88/560 (15.7)	0
Hispanic	33/560 (5.9)	33/560 (5.9)	0
Other	19/560 (3.4)	19/560 (3.4)	0
Renal function before onset of acute kidney injury			
Serum creatinine, mg/dl	1.1 ± 0.3	1.1 ± 0.3	0
Estimated GFR			
≥60 ml/min/1.73 m^2	344/523 (65.8)	344/523 (65.8)	0
45-59 ml/min/1.73 m^2	115/523 (22.0)	115/523 (22.0)	0
30-44 ml/min/1.73 m^2	64/523 (12.2)	64/523 (12.2)	0
Cause of acute kidney injury			
Ischemia	431/541 (79.7)	431/541 (79.7)	0
Nephrotoxins	143/509 (28.1)	143/509 (28.1)	0
Sepsis	279/524 (53.2)	279/524 (53.2)	0
Multifactorial causes	309/527 (58.6)	309/527 (58.6)	0
Primary treating service			
Medical	259/560 (46.2)	259/560 (46.2)	0
Surgical	234/560 (41.8)	234/560 (41.8)	0
Other	67/560 (12.0)	67/560 (12.0)	0
Weight before acute illness, kg	84.1 ± 18.9	84.1 ± 18.9	0
Length of stay before randomization, days			
Hospital	10.3 ± 14.7	10.3 ± 14.7	0
ICU	6.4 ± 7.8	6.4 ± 7.8	0
Charlson comorbidity index	4.2 ± 2.8	4.2 ± 2.8	0
Mechanical ventilation	452/560 (80.7)	452/560 (80.7)	0
Sepsis	350 (62.4)	350 (62.4)	0
APACHE II score	26.1 ± 7.5	26.1 ± 7.5	0
SOFA organ-system score			
Respiratory	2.3 ± 1.1	2.3 ± 1.1	0
Coagulation	1.3 ± 1.2	1.3 ± 1.2	0
Liver	1.4 ± 1.3	1.4 ± 1.3	0
Cardiovascular	2.2 ± 1.7	2.2 ± 1.7	0
Central nervous system	2.5 ± 1.4	2.5 ± 1.4	0
Total	14.4 ± 3.7	14.4 ± 3.7	0

Table B (continued): Comparison of Values Computed in Integrity Check to Reference Article Table 1 Values, Treatment = Less-Intensive Strategy

Characteristic	Palevsky et al	Integrity Check	Diff
Cleveland Clinic ICU Acute Renal Failure score	12.0 ± 3.4	12.0 ± 3.4	0
BUN at initiation of RRT	66.7 ± 35.2	66.7 ± 35.2	0
Cardiovascular SOFA score			
0-2	254 (45.3)	256 (45.9)	2, 0.6
3-4	307 (54.7)	302 (54.1)	5, 0.6
Oliguria			
No	123 (21.9)	123 (21.9)	0
Yes	438 (78.1)	438 (78.1)	0
One session of IHD or SLED of <24 hr of continuous RRT before randomization	366/560 (65.4)	366/560 (65.4)	0
Notes: (1) GFR glomerular filtration rate, IHD intermittent hemodialysis, RRT renal-replacement therapy, SLED sustained low-efficiency dialysis			

Table B (continued): Comparison of Values Computed in Integrity Check to Reference Article Table 1 Values, p-values

Characteristic	Palevsky et al	Integrity Check	Diff
Age, years	0.97	0.97	0
Sex	0.13	0.13	0
Race or ethnic group	0.43	0.43	0
Renal function before onset of acute kidney injury			
Serum creatinine, mg/dl	0.71	0.80	0.9
Estimated GFR	0.84	0.84	0
Cause of acute kidney injury			
Ischemia	0.31	0.31	0
Nephrotoxins	0.92	0.92	0
Sepsis	0.29	0.29	0
Multifactorial causes	0.81	0.81	0
Primary treating service	0.76	0.76	0
Weight before acute illness, kg	>0.99	>0.99	0
Length of stay before randomization, days			
Hospital	0.36	0.36	0
ICU	0.38	0.38	0
Charlson comorbidity index	0.66	0.66	0
Mechanical ventilation	0.91	0.91	0
Sepsis	0.68	0.68	0
APACHE II score	0.29	0.29	0
SOFA organ-system score			
Respiratory	0.10	0.10	0
Coagulation	0.49	0.49	0
Liver	0.29	0.29	0
Cardiovascular	0.23	0.23	0
Central nervous system	0.69	0.69	0
Total	0.21	0.21	0
Cleveland Clinic ICU Acute Renal Failure score	0.11	0.11	0
BUN at initiation of RRT	0.68	0.68	0
Cardiovascular SOFA score	>0.99	0.68	0.31
Oliguria	0.97	0.97	0
One session of IHD or SLED of <24 hr of continuous RRT before randomization	0.54	0.54	0

Notes:
(1) GFR glomerular filtration rate, IHD intermittent hemodialysis, RRT renal-replacement therapy, SLED sustained low-efficiency dialysis

Primary and Secondary Outcomes. Table 3 in the publication [1] reports on primary and secondary study outcomes of patients enrolled in the ATN Trial. Table C lists the variables we used in our replication. All variables were taken from the SAS datasets, d_analysismortality and d_analysisdemo.

Table C: Variables Used to Replicate Selected Table 3 Results

Table Variable	Variables Used in Replication
Sample size	d_analysismortality: TreatmentGp
Death from any cause by day 60	d_analysismortality: EndPtDeath60
In-hospital death	d_analysismortality: EndPtDeathHosp
Discharged to home, off dialysis, by day 60	d_analysismortality: Discharge_Home
Recovery of kidney function by day 28: Complete, Partial, None	d_analysismortality: Rrfstage, Age, Gender; d_analysisdemo: bCardiovascular, Olig
RRT-free days through day 28	d_analysismortality: DaysFreeRRT28; d_analysisdemo: bCardiovascular, Olig
Hospital-free days through day 60	d_analysismortality: HospFreeDays60, Age, Gender; d_analysisdemo: bCardiovascular, Olig
ICU-free days through day 60	d_analysismortality: ICUFreeDays60, Age, Gender; d_analysisdemo: bCardiovascular, Olig

In Table D, we compare the results calculated from the archived datasets to the results published in Table 3, Primary and Secondary Outcomes. As Table D shows, the results are similar.

Table D: Comparison of Values Computed in Integrity Check to Reference Article Table 3 Values, Treatment = Intensive Strategy

Characteristic	Palevsky et al	Integrity Check	Diff
Sample size	563	563	0
Death from any cause by day 60	302 (53.6)	302 (53.6)	0
In-hospital death	288 (51.2)	288 (51.2)	0
Discharged to home, off dialysis, by day 60	88/560 (15.7)	88/560 (15.7)	0
Recovery of kidney function by day 28			
Complete	85/553 (15.4)	85/553 (15.4)	0
Partial	49/553 (8.9)	49/553 (8.9)	0
None	419/553 (75.8)	419/553 (75.8)	0
RRT-free days through day 28	6.0 ± 0.4	6.0 ± 0.4	0
Hospital-free days through day 60	11.0 ± 0.7	11.0 ± 0.7	0
ICU-free days through day 60	18.7 ± 0.9	18.4 ± 0.9	0.3, 0

Treatment = Less-Intensive Strategy

Characteristic	Palevsky et al	Integrity Check	Diff
Sample size	561	561	0
Death from any cause by day 60	289 (51.5)	289 (51.5)	0
In-hospital death	269 (48.0)	269 (48.0)	0
Discharged to home, off dialysis, by day 60	92/561 (16.4)	92/561 (16.4)	0
Recovery of kidney function by day 28			
Complete	102/555 (18.4)	102/555 (18.4)	0
Partial	50/555 (9.0)	50/555 (9.0)	0
None	403/555 (72.6)	403/555 (72.6)	0
RRT-free days through day 28	7.0 ± 0.4	7.0 ± 0.4	0
Hospital-free days through day 60	13.0 ± 0.7	13.0 ± 0.7	0
ICU-free days through day 60	20.1 ± 0.9	20.3 ± 1.0	0.2, 0.1

Table D (continued): Comparison of Values Computed in Integrity Check to Reference Article Table 3 Values, Odds Ratio or Mean Difference (95% CI) and p-values

Characteristic	Palevsky et al	Integrity Check	Diff
Death from any cause by day 60	1.09 (0.86 to 1.40), 0.47	1.09 (0.86 to 1.40), 0.48	0 (0,0), 0.01
In-hospital death	1.15 (0.90 to 1.47), 0.27	1.14 (0.90 to 1.44), 0.28	0.01 (0,0.03), 0.01
Discharged to home, off dialysis, by day 60	0.95 (0.68 to 1.32), 0.75	0.95 (0.69 to 1.31), 0.75	0 (0.01,0.01), 0
Recovery of kidney function by day 28	0.03 (0.02 to 0.07), 0.24	0.03 (-0.02 to 0.07), 0.23	0 (0.04,0), 0.01
RRT-free days through day 28	-0.9 (-1.9 to 0.1), 0.07	-0.9 (-1.9 to 0.1), 0.08	0 (0,0), 0.01
Hospital-free days through day 60	-1.9 (-3.9 to 0.0), 0.053	-1.9 (-3.9 to 0.0), 0.056	0 (0,0), 0.003
ICU-free days through day 60	-1.5 (-4.0 to 1.0), 0.25	-1.5 (-4.0 to 1.1), 0.26	0 (0,0.1), 0.01

References

1. Paul M. Palevsky, M.D., et al, **Intensity of Renal Support in Critically Ill Patients with Acute Kidney Injury**, The New England Journal of Medicine; 2008 July 3; 359(1)7-20.

2. ClinicalTrials.gov Website: A service of the U.S. National Institutes of Health. Acute Renal Failure Trial Network (ATN) Study page.

ATTACHMENT 1

Full Text of Article

Paul M. Palevsky, M.D., et al, **Intensity of Renal Support in Critically Ill Patients with Acute Kidney Injury**, The New England Journal of Medicine; 2008 July 3; 359(1)7-20.

ATTACHMENT 2

SAS Code for Tabulations from the Acute Renal Failure Trial Network Study (ATN) Datasets in the NIDDK Repository

```
/*******************************************************************************/
/*
/* Program: R:\05_Users\Norma\ATN\PalevskyPaper\table1.sas
/* Author:  Norma Pugh
/* Date:    May 2010
/* Purpose: Replicate table 1 results.
/*
/*******************************************************************************/
/* DATA SOURCE */
libname data 'R:\03_Data_And_Tools\Studies\ATN\delivery_from_DCC\20100205';

/* OPTIONS */
options nofmterr;

/* GET DATASET */
data table1;
 merge data.d_analysisdemo data.d_form01(keep=PatientKey Sepsis rename=(Sepsis=F01Sepsis)); by
PatientKey;;
 /* Include randomized patients, not those from observational cohort. */
 if TreatmentGp in(1,2);
 /* Re-categorize race */
 if race in(4,5,6,7) then newrace=9; else newrace=race;
 /* Categorize SOFA: Cardiovascular scores */
 if bCardiovascular in(0,1,2) then CardioSOFA=1;
  if bCardiovascular in(3,4) then CardioSOFA=2;
 /* BUN at initiation of RRT, defn. provided by DCC */
 if BUNPr^=. then PriorRandBUN=BUNPr;
  else if (BUNPr=. & RecentBUN^=.) then PriorRandBUN=RecentBUN;
 /* ESTIMATED GFR */
 /* Step A */
 if Form01PreMorbidCreat>. then TEMPeGFRSCreat=Form01PreMorbidCreat;
  else if (Form01PreMorbidCreat=. & Form01BaselineCreat>.) then
TEMPeGFRSCreat=Form01BaselineCreat;
 if gender=0 & TEMPeGFRSCreat>1.5 then eGFRSCreat=.;
  else if gender=1 & TEMPeGFRSCreat>2 then eGFRSCreat=.;
  else eGFRSCreat=TEMPeGFRSCreat;

 /* Step B */
 if gender>. then do;
  if gender=1 then GenderFact=1;
  if gender=0 then GenderFact=0.742;
 end;

 /* Step C */
 if race>. then do;
  if race^=2 then EthFact=1;
  if race=2 then EthFact=1.21;
 end;

 /* Step D */
 if age>=18 & eGFRSCreat>0 then
  eGFR = 186 * (eGFRSCreat**(-1.154)) * (Age**(-0.203)) * GenderFact * EthFact;
 else eGFR=.;
```

```
/* Categorize */
 if eGFR>=60 then eGFR_CAT=1;
 if 45<=eGFR<60 then eGFR_CAT=2;
 if 30<=eGFR<45 then eGFR_CAT=3;

 /* Serum Creatinine, defn. provided by DCC */
 if eGFRSCreat^=. then eGFRSCreatUsed=eGFRSCreat;
run;


/* REPLICATE ANALYSIS RESULTS */
proc freq data=table1; tables TreatmentGp / list nopct nocum; title'Treatment Group Counts:
Overall'; run;

%macro frq(var);
proc freq data=table1(where=(&var>.)) noprint; tables TreatmentGp / out=denom(keep=TreatmentGp
count rename=(count=denom)); run;
proc freq data=table1(where=(&var>.)) noprint; tables TreatmentGp*&var /
out=frqstats(drop=percent) chisq; output out=pstats chisq; run;
data frqstats; merge frqstats denom; by TreatmentGp; pct=(count/denom)*100; run;
proc print data=frqstats; title"Frequency Counts: &var"; run;
proc print data=pstats; var p_pchi; title"P-value: &var"; run;
%mend frq;

proc sort data=table1; by TreatmentGp; run;

%macro mean_(var);
proc means data=table1 n mean std; by TreatmentGp; var &var; run;
proc glm data=table1; class TreatmentGp; model TreatmentGp=&var; title"P-value: &var"; run;
%mend mean_;

%mean_(age);
%frq(gender);
%frq(newrace);
%mean_(eGFRSCreatUsed);
%frq(eGFR_CAT);
%frq(ischemic);
%frq(nephrotoxic);
%frq(sepsis);
%frq(multifactorial);
%frq(PrimaryTreat);
%mean_(weight);
%mean_(dayshospbrand);
%mean_(daysICUbrand);
%mean_(CharlsonAgeScore);
%frq(MechVentilation);
%frq(F01Sepsis);
%mean_(bTotApacheII);
%mean_(bRespiratory);
%mean_(bCoagulation);
%mean_(bLiver);
%mean_(bCardiovascular);
%mean_(bCentralNerve);
%mean_(bTotalSOFA);
%mean_(BTotCCFARF);
%mean_(PriorRandBUN);
%frq(CardioSOFA);
```

```
%frq(olig);
%frq(HemoCRRT);
```

```
/*******************************************************************************/
/*
/* Program: R:\05_Users\Norma\ATN\PalevskyPaper\table3.sas
/* Author:  Norma Pugh
/* Date:    May 2010
/* Purpose: Replicate table 3 results.
/*
/*******************************************************************************/
/* DATA SOURCE */
libname data 'R:\03_Data_And_Tools\Studies\ATN\delivery_from_DCC\20100205';

/* OPTIONS */
options nofmterr;

/* GET DATASET */
data table3;
 merge data.d_analysismortality data.d_analysisdemo(keep=patientkey bCardiovascular olig);
 by patientkey;
 if TreatmentGp in(1,2);
                                          /* Include randomized patients, not those from
observational cohort. */
run;

/* REPLICATE ANALYSIS RESULTS */
proc freq data=table3; tables TreatmentGp / list nopct nocum; title'Treatment Group Counts:
Overall'; run;

%macro frq(var);
proc freq data=table3(where=(&var>.)) noprint; tables TreatmentGp / out=denom(keep=TreatmentGp
count rename=(count=denom)); run;
proc freq data=table3(where=(&var>.)) noprint; tables TreatmentGp*&var /
out=frqstats(drop=percent) chisq; run;
data frqstats; merge frqstats denom; by TreatmentGp; pct=(count/denom)*100; run;
proc print data=frqstats; title"Frequency Counts: &var"; run;
%mend frq;

%macro mean_(indata,var);
proc sort data=&indata; by TreatmentGp; run;
proc means data=&indata n mean stderr; by TreatmentGp; var &var; run;
%mend mean_;

%macro odds(indata,var);
proc logistic data=&indata; class TreatmentGp; model &var=TreatmentGp; title"Odds Ratio: &var";
run;
%mend odds;

%macro meandiff_frq(indata,var,adj);
proc glm data=&indata;
 class TreatmentGP gender bCardiovascular olig;
 model &var=TreatmentGP &adj;
 lsmeans TreatmentGP / adjust=nelson cl pdiff;
 title"Mean Difference: &var";
run;
%mend meandiff_frq;
```

```
%macro meandiff(indata,var,adj);
proc glm data=&indata;
 class TreatmentGP gender bCardiovascular olig;
 model &var=TreatmentGP &adj;
 lsmeans TreatmentGP / cl pdiff=control;
 title"Mean Difference: &var";
run;
%mend meandiff;

%frq(EndPtDeath60);
%frq(EndPtDeathHosp);
%frq(Discharge_Home);
%frq(rrfstage);

%mean_(table3,%str(DaysFreeRRT28 HospFreeDays60 ICUFreeDays60));

%odds(table3,EndPtDeath60);
%odds(table3,EndPtDeathHosp);
%odds(table3,Discharge_Home);
%meandiff_frq(table3,rrfstage,%str(bCardiovascular olig));
%meandiff(table3,DaysFreeRRT28,%str(age gender bCardiovascular olig));
%meandiff(table3,HospFreeDays60,%str(age gender bCardiovascular olig));
%meandiff(table3,ICUFreeDays60,%str(age gender bCardiovascular olig));
```

ATTACHMENT 3

Calculation of Select Study Variables

Provided by the Study Data Coordinating Center

Calculations for Serum Creatinine Level and Estimated GFR
as Provided by Study Data Coordinating Center

Serum creatinine, mg/dl:
Variable eGFRSCreatUsed.

In the formula for Estimated GFR, below, the first term (derived to be eGFRSCreat) is reported as eGFRSCreatUsed when the formula for Estimated GFR is calculated. Otherwise it is set to missing .

Estimated GFR, ml/min/1.73m-squared:
Variable eGFR.

The Estimated GFR is calculated according to an algorithm (provided by Paul Palevsky on March 14, 2008) from other variables as follows:

eGFR = 186 * (eGFRSCreat ** (-1.154)) * (Age ** (-0.203)) * GenderFact *EthFact,
where the following logic is used to obtain the four factors in the equation:

(A)
eGFRSCreat is obtained from Form01PreMorbidCreat if Form01PreMorbidCreat is not missing
else if Form01PreMorbidCreat is missing and Form01BaselineCreat is not missing
then eGFRSCreat is obtained from Form01BaselineCreat
then
if Gender is Female (0) and the obtained eGFRSCreat>1.5 (from (A)) the eGFRSCreat is set to missing
else if Gender is Male(1) and the obtained eGFRSCreat>2 the eGFRSCreat is set to missing

(B)
if Gender is specified (>.) then GenderFact is set equal to 1 if the Gender is Male(1)
else if the Gender is Female (0) then GenderFact is set equal to 0.742

(C)
if Race is specified (>.) then EthFact is set equal to 1 if the Race is not Black
else if the Race is Black (2) then EthFact is set equal to 1.21

(D)
If Age greater than or equal 18 (>=18) and the eGFRSCreat is greater than 0 (after steps (A), (B) and (C) above) then the quantity eGFR is calculated by the formula above.
else the eGFR is set to missing (.)

ATTACHMENT 4

Appendix to Integrity Check for the Acute Renal Failure Trial Network Study (ATN) Data Files

After reviewing the DSIC, the DCC provided clarifications for the minor discrepancies that were found. As stated in the introduction, the NIDDK data repository does 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, review all comments from the DCC and have included these clarifications for completeness.

June 2, 2010 – Prepared by Terry O'Connor

DSIC Table/Page#	Characteristic	Reason for Difference
1.Table B, pg. 4	Cardiovascular (CV) SOFA Score	ATN used CV Score called "SOFA" on the AnalysisDemo file per design of the trial as described in the study protocol. This variable was a randomization stratification variable. Integrity Check (IC) used baseline CV Score, called "bCardiovascular" on the AnalysisDemo file.
2.Table B, pg. 6	Cardiovascular (CV) SOFA Score	Same as item 1 above
3.Table B, pg. 7	Serum creatinine (SC), mg/dl	There were two zero (0) creatinine values that were not included in the ATN calculation which were included in the IC calculation– see attached document: **EGFR Difference with NIDDK.doc**
4.Table B, pg. 7	Cardiovascular (CV) SOFA Score	Same as items 1 and 2 above
5.Table D, pg. 9	ICU-free days through day 60	ATN used Proc MIXED and CV SOFA score called "SOFA" on the AnalysisDemo file.
6.Table D, pg. 10	Death from any cause by day 60	ATN used conditional logistic with site as strata (Proc PHREG) and CV SOFA score called "SOFA" on the AnalysisDemo file.
7.Table D, pg. 10	In-hospital death	Same as item 6 above
8.Table D, pg. 10	Discharged to home off dialysis by day 60	Same as item 6 above
9.Table D, pg. 10	Recovery of kidney function by day 28	ATN used Proc CATMOD. There is a typo or misprint in the lower 95%CL for ATN results. The 0.02 shown as the lower CL should be -0.02.
10.Table D, pg. 10	RRT-free days through day 28	ATN used Proc MIXED and CV SOFA score called "SOFA" on the AnalysisDemo file.
11.Table D, pg. 10	Hospital-free days through day 60	Same as item 10 above
12.Table D, pg. 10	ICU-free days through day 60	Same as item 10 above

CSP530 (6/1/10):
JNV
Difference between NIDDK Calculation of eGFR Serum Creatinine and CSP530 Calculation:

CSP530 reported the difference between Treatment Groups with respect to the calculation of Serum Creatinine as follows:

eGFRCreatUsed:	Treatment Group	N	Mean	StdDev		
	Conventional	523	1.1207	0.3467		
	Intensive	524	1.1124	0.3653		
	Difference		0.0082	0.3561		
	CSP530: --------	Pooled Difference:		t-Value 0.37	Pr>|t| 0.71	
	NIDDK indicated: ----------			Pr>|t| 0.80		

A calculation shows that NIDDK included two more Conventional cases when reporting the difference between Treatment Groups with respect to the calculation of Serum Creatinine as follows:

eGFRNIDDKUsed:	Treatment Group	N	Mean	StdDev		
	Conventional	525(*)	1.1179	0.3497		
	Intensive	524	1.1124	0.3653		
	Difference		0.0055	0.3576		
	NIDDK: -------	Pooled Difference:		t-Value 0.25	Pr>|t| 0.80	

(*) The two extra cases included in the Conventional Group by NIDDK were cases which allowed the inclusion of the value ‘0’ in that Conventional group. In the CSP 530 version the value of ‘0’ was intentionally eliminated from the calculation of the overall eGFR Creatinine (as prescribed by PP). Those two ‘0’ cases included in the NIDDK version explain the difference between the p-values obtained between the two calculations for Reference Article Table#1.