# Dataset Integrity Check for the TrialNet(14) Data Files

**Prepared by Michael Spriggs**
**IMS Inc.**
3901 Calverton Blvd, Suite 200 Calverton MD 20705
**April 13, 2015**

# Table of 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

Canakinumab is a fully human anti-interleukin-1β (anti-IL-1β) monoclonal antibody, designed to bind to human IL-1β and to functionally neutralize the bioactivity of this pro-inflammatory cytokine. In this phase II clinical trial, subjects were randomized to receive either monthly subcutaneous injections of 2.0 mg/kg canakinumab or placebo for 12 months. All subjects received standard intensive diabetes treatment with insulin and dietary management. Diabetes control was evaluated by measuring glycosylated hemoglobin (HbA1c) every three months. Insulin production was measured by a series of mixed meal glucose tolerance tests (MMTT) conducted regularly during the study. Subjects were followed for one year of treatment, plus one to three years of follow-up.

# 3 Archived Datasets

All SAS data files, as provided by the Data Coordinating Center (DCC), are located in the “data” folder in the data package. For this replication, variables were taken from the different analysis datasets created by the DCC.

# 4 Statistical Methods

Analyses were performed to duplicate results for the data published by Moran A, et al [1] in Lancet on June 1 2013. To verify the integrity of the three datasets, descriptive statistics of baseline characteristics were computed, by treatment group (Table B, Table D and Table F).

# 5 Results

Table 1 in the publication [1], Baseline Demographic and Laboratory Characteristics of Subjects at Entry. Table A lists the variables that were used in the replication and Table B compares the results calculated from the archived data file to the results published in Table 1. The results of the replication are similar to the published results.

Table 2 in the publication [1]: Adverse events by worst grade experienced in the canakinumab trial. Table C lists the variables that were used in the replication and Table D compares the results calculated from the archived data file to the results published in Table 2. The results of the replication are very similar to the published results.
Table 3 in the publication [1]: The number of events and participants by adverse event type in the canakinumab trial. Table E lists the variables that were used in the replication and Table F compares the results calculated from the archived data file to the results published in Table 3. The results of the replication are similar to the published results.

# 6 Conclusions

The NIDDK repository is confident that the TrialNet(14) data files to be distributed are a copy of the manuscript data.

# 7 References

Moran A, Bundy B, Becker DJ et al., Group AS. Interleukin-1 antagonism in type 1 diabetes of recent onset: two multicenter, randomized double-masked, placebo-controlled trials. Lancet 2013 June 1; 381(9881)

**Table A:** Variables used to replicate Table 1: Baseline Demographic and Laboratory Characteristics of Subjects at Entry.

| Table Variable | Variables Used in Replication from the Trialnet(14) Datasets |
|---|---|
| Age- yr | Demographics.age |
| Male sex- number of patients (%) | Demographics.sex |
| Race* - number of patients White (%) | Demographics.Race_White |
| Ethnicity – number of patients (%)Non-Hispanic | Demographics.Ethnicity |
| Number of autoantibodies | Research_labs.test_name<br>Research_labs.spec_name<br>Research_labs.result<br>Research_labs.visit |
| Number of days from diagnosis to first infusion | study_drug_administration.DATE_OF_<br>VISIT study_drug_administration. Visit<br>screening_medical_history.<br>DateOfT1DdiagnosisMonth<br>screening_medical_history.<br>DateOfT1DdiagnosisYear<br>screening_medical_history.<br>DateOfT1DdiagnosisDay<br>screening_medical_history.visit |
| Weight (kg) | Physical.weightkg |
| Body Mass Index (kg/m2) | Physical.weightkg Physical.HeightCM |
| Mean AUC for C-peptide (pmol/ml) | Research_labs.test_name<br>Research_labs.spec_name<br>Research_labs.result<br>Research_labs.visit |
| Glycated hemoglobin at baseline* (HbA1c -%) | Research_labs.test_name<br>Research_labs.spec_name<br>Research_labs.result<br>Research_labs.visit |
| Total daily insulin dose at baseline* – (U/kg) | Physical.weightkg<br>diabetes_management.<br>AverageUnitsOfItermediateInsul<br>Diabetes_management.<br>AvgUnitsShortActingInsulin |
| Diabetes associated HLA alleles present | Research_labs.test_name<br>Research_labs.outcome<br>Research_labs.spec_name<br>Research_labs.visit |

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

| Characteristic | Canakinumab [Manuscript n=47] | Canakinumab [DSIC n=49] | Canakinumab [Diff n=-2] |
|---|---|---|---|
| Age- yr: Mean | 11・7 (4・0) | 11.6 ± 3.9 | .1+-0.1 |
| Age-yr: Median(Range) | 11 (6–25) | 11 6-25 | 0 |
| Male sex- number of patients (%) | 24 (51%) | 24 (49.0) | 0(2) |
| Race* - number of patients White (%) | 42 (93%) | 45 (91.8) | 3(1.2) |
| Ethnicity – number of patients (%)Non-Hispanic | 47 (100%) | 44 (89.8) | 3(10.2) |
| Number of autoantibodies** - no. of patients (%) 1 | 4 (9%) | 3 (6.1) | 1(2.9) |
| Number of autoantibodies** - no. of patients (%) 2 | 8 (17%) | 7 (14.3) | 1(2.7) |
| Number of autoantibodies** - no. of patients (%) 3 | 17 (36%) | 17 (34.7) | 0(1.3) |
| Number of autoantibodies** - no. of patients (%) 4 | 18 (38%) | 22 (44.9) | -4 (-6.9) |
| Number of days from diagnosis to first infusion | 75・8 (17・9) | 75.4 ± 17.8 | .4+-(0.1) |
| Range | 76 (36–104) | 76 36-104 | 0 |
| Weight (kg) | 49・1 (21・7) | 49.4 ± 20.9 | -0.3+-0.8 |
| Body Mass Index (kg/m2) | 20・7 (5・53) | 20.88 ± 5.22 | -0.18+-0.31 |
| Mean AUC for C-peptide (pmol/ml) | 0・65 (0・36) | 0.662 ± 0.357 | -0.012± 0.003 |
| Glycated hemoglobin at baseline* (HbA1c -%) | 7・09 (1・16) | 7.08 ± 1.14 | 0.01± 0.02 |
| Total daily insulin dose at baseline* – (U/kg) | 0・37 (0・26) | 0.369 ± 0.268 | 0.001+--0.008 |
| Diabetes associated HLA alleles present* - number of patients (%) DR3 and DR4 | 13 (28%) | 13 (27.1) | 0(.9) |
| Diabetes associated HLA alleles present* - number of patients (%) DR3 only | 12 (26%) | 13 (27.1) | -1(-1.1) |
| Diabetes associated HLA alleles present* - number of patients (%) DR4 Only | 13 (28%) | 13 (27.1) | 0(.9) |
| Diabetes associated HLA alleles present* - number of patients (%) Neither | 9 (19%) | 9 (18.8) | 0(0) |

| Characteristic | Placebo [Manuscript n=22] | Placebo [DSIC n=22] | Placebo [Diff n=0] |
|---|---|---|---|
| Age- yr: Mean | 12・5 (6・4) | 12.5 ± 6.4 | 0 |
| Age-yr: Median(Range) | 10・5 (6–31) | 11 6-31 | 0 |
| Male sex- number of patients (%) | 14 (64%) | 14 (63.6) | 0 |
| Race* - number of patients White (%) | 22 (100%) | 22 (100.0) | 0 |
| Ethnicity – number of patients (%)Non-Hispanic | 22 (100%) | 20 (90.9) | 2(9.1) |
| Number of autoantibodies** - no. of patients (%) 1 | 0 (0%) | 0 (0%) | 0 |
| Number of autoantibodies** - no. of patients (%) 2 | 4 (18%) | 4 (18.2) | 0 |
| Number of autoantibodies** - no. of patients (%) 3 | 8 (36%) | 4 (18.2) | 4 (17.8) |
| Number of autoantibodies** - no. of patients (%) 4 | 10 (45%) | 14 (63.6) | -4(18.6) |
| Number of days from diagnosis to first infusion | 75・6 (21・8) | 75.6 ± 21.8 | 0 |
| Range | 80 (21–99) | 80 21-99 | 0 |
| Weight (kg) | 46・3 (18・5) | 46.6 ± 18.7 | -0.3+--0.2 |
| Body Mass Index (kg/m2) | 19・8 (3・73) | 20.02 ± 3.82 | -0.22+--0.09 |
| Mean AUC for C-peptide (pmol/ml) | 0・62 (0・29) | 0.615 ± 0.296 | 0(-0.006) |
| Glycated hemoglobin at baseline* (HbA1c -%) | 6・81 (0・95) | 6.81 ± 0.95 | 0 |
| Total daily insulin dose at baseline* – (U/kg) | 0・35 (0・15) | 0.363 ± 0.147 | -0.013+--0 |
| Diabetes associated HLA alleles present* - number of patients (%) DR3 and DR4 | 6 (29%) | 6 (28.6) | 0 |
| Diabetes associated HLA alleles present* - number of patients (%) DR3 only | 4 (19%) | 4 (19.0) | 0 |
| Diabetes associated HLA alleles present* - number of patients (%) DR4 Only | 7 (33%) | 7 (33.3) | 0 |
| Diabetes associated HLA alleles present* - number of patients (%) Neither | 4 (19%) | 4 (19.0) | 0 |

**Table C:** Variables used to replicate Table 2: Adverse events by worst grade experienced in the canakinumab trial.

| Table Variable | Variables Used in Replication |
|---|---|
| Adverse Event Grade | Adverseevent.severity1 |
| Treatment Group | treatment_table.treatmentname |

**Table D:** Comparison of values computed in integrity check to reference article Table 2 values

| Adverse Effect Grade | Canakinumab [Manuscript n=47] | Canakinumab [DSIC n=49] | Canakinumab [Diff n=-2] | Placebo [Manuscript n=22] | Placebo [DSIC n=22] | Placebo [Diff n=0] |
|---|---|---|---|---|---|---|
| 0 | 13 (28%) | 13 (27%) | 0(1%) | 6 (27%) | 6 (27%) | 0 (0%) |
| 1 | 2 (4%) | 3 (6%) | -1(-2%) | 0 (0%) | 0 (0%) | 0 (0%) |
| 2 | 28 (60%) | 27 (55%) | 1(5%) | 12 (55%) | 12 (55%) | 0 (0%) |
| 3 | 3 (6%) | 5 (10%) | -2(-4%) | 3 (14%) | 3 (14%) | 0 (0%) |
| 4 | 1 (2%) | 1 (2%) | 0 (0%) | 1 (5%) | 1 (5%) | 0 (0%) |
| 5 | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |

**Table E:** Variables used to replicate Table 3: The number of events and participants by adverse event type in the canakinumab trial.

| Table Variable | Variables Used in Replication |
|---|---|
| Adverse Effect Category | Adverseevent.Category |
| Treatment Group | treatment_table.treatmentname |

**Table F:** Comparison of values computed in integrity check to reference article Table 3 values

| Adverse Effect Category | Canakinumab [Manuscript n=47] | Canakinumab [DSIC n=49] | Canakinumab [Diff n=-2] | Placebo [Manuscript n=22] | Placebo [DSIC n=22] | Placebo [Diff n=0] |
|---|---|---|---|---|---|---|
| Pain | 12 6 (13%) | 12 6 | 0 0 | 2 1 (5%) | 2 1 | 0 0 |
| Flu-like symptoms | 1 1 (2%) | 0 0 | 1 1 | 1 1 (5%) | 0 0 | 1 1 |
| Infection | 22 13 (28%) | 24 14 | -2 -2 | 16 8 (36%) | 18 9 | -2 -1 |
| Gastrointestinal | 7 5 (11%) | 8 6 | -1 -1 | 6 5 (23%) | 7 5 | -1 0 |
| Pulmonary or upper respiratory | 2 2 (4%) | 2 2 | 0 0 | 2 2 (9%) | 2 2 | 0 0 |
| Constitutional symptoms | 5 4 (9%) | 5 4 | 0 0 | 1 1 (5%) | 1 1 | 0 0 |
| Blood or bone: ANC or AGC | 12 8 (17%) | 15 9 | -3 -1 | 6 3 (14%) | 7 3 | -1 0 |
| Blood or bone: other | 0 0 (0%) | 0 0 | 0 0 | 0 0 (0%) | 0 0 | 0 0 |
| Surgery or intraoperative injury | 0 0 (0%) | 0 0 | 0 0 | 1 1 (5%) | 1 1 | 0 0 |
| Neurological | 4 4 (9%) | 4 4 | 0 0 | 1 1 (5%) | 1 1 | 0 0 |
| Dermatological or skin | 5 5 (11%) | 6 6 | -1 -1 | 3 3 (14%) | 3 3 | 0 0 |
| Musculoskeletal or soft tissue | 4 4 (9%) | 4 4 | 0 0 | 0 0 (0%) | 0 0 | 0 0 |
| Auditory or ear | 1 1 (2%) | 1 1 | 0 0 | 0 0 (0%) | 2 1 | -2 -1 |
| Endocrine | 3 3 (6%) | 4 3 | -1 0 | 0 0 (0%) | 0 0 | 0 0 |
| Blood or bone marrow | 3 1 (2%) | 0 0 | 3 1 | 0 0 (0%) | 0 0 | 0 0 |
| Ocular or visual | 0 0 (0%) | 1 1 | -1 -1 | 1 1 (5%) | 1 1 | 0 0 |
| Total | 81 · · | 88 | -7 | 40 · · | 46 | -6 |

# Attachment A: SAS Code

```
title1 "%sysfunc(getoption(sysin))";
title2 " ";

options nofmterr source2 mprint symbolgen spool;
libname sas_data "/prj/niddk/ims_analysis/TrialNet/private_orig_data/new_data_3_6_15/Data.Extraction.14.Version150228PW/14/sasv9/";
filename sascsv "/prj/niddk/ims_analysis/TrialNet/private_orig_data/Data.Extraction.14.Version120531/14/csv/adverseeventedit.csv";

proc import datafile=sascsv
     out=adversecsv
     dbms=csv
     replace;
     getnames=yes;
run;

data adversecsv;
  set adversecsv;
  length severitynum 8.;
  if (substr(severity1,1,1))="1" then severitynum=1;
  else if (substr(severity1,1,1))="2" then severitynum=2;
  else if (substr(severity1,1,1))="3" then severitynum=3;
  else if (substr(severity1,1,1))="4" then severitynum=4;
  else if (substr(severity1,1,1))="5" then severitynum=5;
  else delete;

proc sort data=adversecsv;
  by MASKID;


proc format;
  value yesnof
        1="Yes"
        2="No"
        ;
        ;

*** File containing macro for examining each dataset ***;
%include '/prj/niddk/ims_analysis/sas_macros/redaction_data_summary.sas';

%macro freqdata1(order=, invar=, level=, popvar=, totalvl=);

%if &totalvl.=null %then %do;
  proc freq data=table1 noprint;
       tables &invar*treatmentname/out=data1 outpct;
       format _all_;
       run;

  data data1(keep=LEVEL treatmentname name CHARALL ORDERER);
    set data1(rename=(&invar=LEVEL));
    length name $100 CHARALL $100;
    name=upcase("&invar");
    PCT_DISP=round(PCT_COL,.1);
    CHARALL=compress(put(COUNT,8.))||" ("||compress(put(PCT_DISP,8.1))||")";
    ORDERER=&order;
    if level in &level then output;
%end;
%else %do;
```

```
  proc freq data=table1 noprint;
       tables &invar*treatmentname/out=data1 outpct;
       format _all_;
       where &popvar. in &totalvl.;
       run;

  data data1(keep=LEVEL treatmentname name CHARALL ORDERER);
    set data1(rename=(&invar=LEVEL));
    length name $100 CHARALL $100;
    name=upcase("&invar");
    PCT_DISP=round(PCT_COL,.1);
    CHARALL=compress(put(COUNT,8.))||" ("||compress(put(PCT_DISP,8.1))||")";
    ORDERER=&order;
    if level in &level then output;
%end;
data accumfreq1;
  set accumfreq1 data1;

%mend freqdata1;

%macro meandata1(order=, invar=, roundvar=, digit=);
proc means data=table1 mean stddev noprint;
       var &invar;
       class treatmentname;
       output out=data1 mean=mean stddev=stddev;
       run;

data data1(drop=_TYPE_ _FREQ_ mean stddev);
  set data1;
  length name CHARALL $100;
  name=upcase("&invar");
  mean=round(mean,&roundvar);
  stddev=round(stddev,&roundvar);
  CHARALL=compress(put(mean,8.&digit))||" ± "||compress(put(stddev,8.&digit));
  ORDERER=&order;

data accummean1;
  set accummean1 data1;

%mend meandata1;

%macro mediandata1(order=, invar=, roundvar=, digit=);
proc means data=table1 median p25 p75 min max noprint;
       var &invar;
       class treatmentname;
       output out=data1 median=median p25=p25 p75=p75 min=min max=max;
       run;

data data1(drop=_TYPE_ _FREQ_ median p25 p75 min max);
  set data1;
  length name CHARALL $100;
  name=upcase("&invar");
  median=round(median,&roundvar);
  min=round(min,&roundvar);
  max=round(max,&roundvar);
  ORDERER=&order;
  CHARALL=compress(put(median,8.&digit))||" "||compress(put(min,8.&digit)||"-"||put(max,8.&digit));
  output;
```

```
data accummedian1;
  set accummedian1 data1;

%mend mediandata1;

%macro inertdata1(order=);

  proc freq data=table1 noprint;
        tables treatmentname/out=data1;
        format _all_;
        run;

  data data1(keep=treatmentname name CHARALL ORDERER);
    set data1;
    length name $100 CHARALL $100;
    name=" ";
    CHARALL=" ";
    ORDERER=&order;

data accuminert1;
  set accuminert1 data1;

%mend inertdata1;

data accumfreq1 accummean1 accummedian1 accuminert1;
  set _null_;

data research_labs;
  set sas_data.research_labs;

proc sort data=research_labs;
  by maskid;

data research_aa(drop=test_name result rename=(s_test_name=test_name s_result=result));
  set research_labs;
  length s_test_name s_result $10. result_n 8.;
  s_test_name=strip(test_name);
  s_result=strip(result);
  if test_name in("GAD65H" "IA-2H" "ICA" "MIAA") and strip(visit)="Screening" then do;
    if s_result='<.05' then result_n=0.025;
    else if s_result ne " " then result_n=put(s_result,8.2);
    output;
  end;

data research_peptide(drop=test_name result rename=(s_test_name=test_name s_result=result));
  set research_labs;
  length s_test_name s_result $10. result_n 8.;
  s_test_name=strip(test_name);
  s_result=strip(result);
  if index(spec_name,"peptide")>0 and strip(visit)="Screening" then do;
    if s_result not in (" " "<.05") then result_n=put(s_result,8.2);
    output;
  end;

data research_hemo(drop=test_name result rename=(s_test_name=test_name s_result=result));
  set research_labs;
  length s_test_name s_result $10. Glyc_HEMO 8.;
```

```
  s_test_name=strip(test_name);
  s_result=strip(result);
  if index(spec_name,"HbA1c")>0 and Visit="Baseline" then do;
    if s_result ne " " then Glyc_HEMO =put(s_result,8.2);
    output;
  end;

proc freq data=research_labs;
  tables test_name/missing list;

data research_DR(drop=test_name result rename=(s_test_name=test_name s_result=result));
  set research_labs;
  length s_test_name s_result $10.;
  s_test_name=strip(test_name);
  s_result=strip(result);
  if index(test_name,"DR")>0 then  output;

data research_DR_short(keep=MaskID DR_NET);
  set research_DR;
  by MaskID;
  length DR_N 8. DR_NET $10;
  retain DR_N;
  if first.MASKID then DR_N=0;
  if strip(test_name)="DR3" and strip(result)="ABSENT" then DR_N=DR_N+2;
  if strip(test_name)="DR4" and strip(result)="ABSENT" then DR_N=DR_N+1;
  if last.MASKID then do;
    DR_NET=strip(put(DR_N,8.));
    output;
  end;

data research_short(keep=MASKID ANTIBODY_CHAR);
  set research_aa;
  by MASKID;
  length ANTIBODY_COUNT 8. ANTIBODY_CHAR $8.;
  retain ANTIBODY_COUNT;
  if first.MASKID then ANTIBODY_COUNT=0;
  if test_name="GAD65H" and result_n>.032 then ANTIBODY_COUNT=ANTIBODY_COUNT+1;
  if test_name="IA-2H" and result_n>5 then ANTIBODY_COUNT=ANTIBODY_COUNT+1;
  if test_name="ICA" and result_n>=10 then ANTIBODY_COUNT=ANTIBODY_COUNT+1;
  if test_name="MIAA" and result_n>.01 then ANTIBODY_COUNT=ANTIBODY_COUNT+1;
  if last.MASKID then do;
    ANTIBODY_CHAR=strip(put(ANTIBODY_COUNT,8.));
    output;
  end;

data screen;
  set sas_data.screening_informed_consent;
  screen_date=date_of_visit;

data diab;
  set SAS_DATA.diabetes_management;
  where visit="Baseline";


data treatment_table;
  set sas_data.treatment_table;
  where MASKID ne .;
```

```
data cscrmedhist;
  set sas_data.screening_medical_history;
  if strip(visit)="Screening" then do;
       if DateOfT1DdiagnosisMonth="Jan" then DateOfT1DdiagnosisMonthn=1;
  else if DateOfT1DdiagnosisMonth="Feb" then DateOfT1DdiagnosisMonthn=2;
  else if DateOfT1DdiagnosisMonth="Mar" then DateOfT1DdiagnosisMonthn=3;
  else if DateOfT1DdiagnosisMonth="Apr" then DateOfT1DdiagnosisMonthn=4;
  else if DateOfT1DdiagnosisMonth="May" then DateOfT1DdiagnosisMonthn=5;
  else if DateOfT1DdiagnosisMonth="Jun" then DateOfT1DdiagnosisMonthn=6;
  else if DateOfT1DdiagnosisMonth="Jul" then DateOfT1DdiagnosisMonthn=7;
  else if DateOfT1DdiagnosisMonth="Aug" then DateOfT1DdiagnosisMonthn=8;
  else if DateOfT1DdiagnosisMonth="Sep" then DateOfT1DdiagnosisMonthn=9;
  else if DateOfT1DdiagnosisMonth="Oct" then DateOfT1DdiagnosisMonthn=10;
  else if DateOfT1DdiagnosisMonth="Nov" then DateOfT1DdiagnosisMonthn=11;
  else if DateOfT1DdiagnosisMonth="Dec" then DateOfT1DdiagnosisMonthn=12;
  DXDATE=MDY(DateOfT1DdiagnosisMonthn,DateOfT1DdiagnosisDay,DateOfT1DdiagnosisYear);
  output;
  end;

data demographics;
  set sas_data.demographics;

data celig;
  set sas_data.eligibility;

data cdrugadmin;
  set sas_data.study_drug_administration;
  DA_DATE=DATE_OF_VISIT;
  if visit="Baseline" then output;

data adverseevent;
  set sas_data.adverseevent;

data physical;
  set sas_data.physical_exam;
  if visit="Baseline" then output;

data research_auc(drop=test_name result rename=(s_test_name=test_name s_result=result));
  set research_labs;
  length s_test_name s_result $10.;
  s_test_name=strip(test_name);
  s_result=strip(result);
  if visit="Screening" and index(test_name,"PEP")>0 then output;

data research_auc_proc(keep=MaskID result_:);
  set research_auc;
      by MaskID;
  length resultn
         result_m10
         result_0
         result_15
         result_30
         result_60
         result_90
         result_120
         result_150
         result_180
         result_210
```

```
        result_240
8.;
  retain result_m10
         result_0
         result_15
         result_30
         result_60
         result_90
         result_120
         result_150
         result_180
         result_210
         result_240
         ;

  if result="<.05" then resultn=.025;
  else resultn=input(result,8.);
  resultn=resultn*.33;

  if first.maskid then do;
    result_m10=.;
    result_0  =.;
    result_15 =.;
    result_30 =.;
    result_60 =.;
    result_90 =.;
    result_120=.;
    result_150=.;
    result_180=.;
    result_210=.;
    result_240=.;
  end;

  if test_name="PEP-10" then result_m10  =resultn;
  if test_name="PEP0"   then result_0    =resultn;
  if test_name="PEP120" then result_120  =resultn;
  if test_name="PEP15"  then result_15   =resultn;
  if test_name="PEP150" then result_150  =resultn;
  if test_name="PEP180" then result_180  =resultn;
  if test_name="PEP210" then result_210  =resultn;
  if test_name="PEP240" then result_240  =resultn;
  if test_name="PEP30"  then result_30   =resultn;
  if test_name="PEP60"  then result_60   =resultn;
  if test_name="PEP90"  then result_90   =resultn;

  if last.maskid then output;

data research_auc_proc_b;
  set research_auc_proc;
  length result_string $50;
  if result_m10>. then result_string=strip(result_string)||"-10,";
  if result_0  >. then result_string=strip(result_string)||"0,"  ;
  if result_15 >. then result_string=strip(result_string)||"15," ;
  if result_30 >. then result_string=strip(result_string)||"30," ;
  if result_60 >. then result_string=strip(result_string)||"60," ;
  if result_90 >. then result_string=strip(result_string)||"90," ;
  if result_120>. then result_string=strip(result_string)||"120,";
  if result_150>. then result_string=strip(result_string)||"150,";
```

```
  if result_180>. then result_string=strip(result_string)||"180,";
  if result_210>. then result_string=strip(result_string)||"210,";
  if result_240>. then result_string=strip(result_string)||"240,";

  if index(result_string,'-10,15,30,60,90,120,')>0 then do;
    AUC_RESULT=(7.5*result_m10+15*result_15+22.5*result_30+30*result_60+30*result_90+15*result_120)/120;
  end;
  *** Has to be looked at after the first because it is a substring of it ***;
  else if index(result_string,'0,15,30,60,90,120,')>0 then do;
    AUC_RESULT=(7.5*result_0+15*result_15+22.5*result_30+30*result_60+30*result_90+15*result_120)/120;
  end;
  else if index(result_string,'0,15,30,60,90,180,')>0 then do;
    AUC_RESULT=(7.5*result_0+15*result_15+22.5*result_30+30*result_60+37.5*result_90+7.5*result_180)/120;
  end;

proc freq data=research_auc_proc_b;
  tables result_string/missing list;

proc sort data=screen;
  by MASKID;

proc sort data=diab;
  by MASKID;

proc sort data=treatment_table;
  by MASKID;

proc sort data=cscrmedhist;
  by MASKID;

proc sort data=celig;
  by MASKID;

proc sort data=cdrugadmin;
  by MASKID;

proc sort data=adverseevent;
  by MASKID;

proc sort data=demographics;
  by MASKID;

proc sort data=physical;
  by MASKID;

data table1;
  merge screen(in=in_screen)
        demographics
        diab
        physical
        treatment_table
        cscrmedhist
        celig
        cdrugadmin
        research_hemo
        research_short
        research_dr_short
        research_auc_proc_b
```

```
      ;
  by maskid;
  length c_race_white $8.;
  c_race_white=strip(put(race_white,8.));
  total_insulin_kg=(AverageUnitsOfItermediateInsul+AvgUnitsShortActingInsulin)/weightkg;
  bmi=WEIGHTKG/(HeightCM*HeightCM/10000);
  INFUSE_DAYS=DA_DATE-DXDATE;
  if in_screen then output;

%meandata1(order=1,  invar=age             , roundvar=.1, digit=1);
%mediandata1(order=1.5,  invar=age            , roundvar=1, digit=0);
%freqdata1(order=2,  invar=sex             , level=("Male"),popvar=, totalvl=null);
%freqdata1(order=3,  invar=c_race_white    , level=("1"),popvar=ethnicity, totalvl=("Not Hispanic,Latino or Spanish origin" "Hispanic,Latino,or Spanish
origin"));
%freqdata1(order=4,  invar=ethnicity       , level=("Not Hispanic,Latino or Spanish origin"),popvar=ethnicity, totalvl=("Not Hispanic,Latino or Spanish
origin" "Hispanic,Latino,or Spanish origin"));
%freqdata1(order=7,  invar=ANTIBODY_CHAR  , level=("1"),popvar=, totalvl=null);
%freqdata1(order=8,  invar=ANTIBODY_CHAR  , level=("2"),popvar=, totalvl=null);
%freqdata1(order=9,  invar=ANTIBODY_CHAR  , level=("3"),popvar=, totalvl=null);
%freqdata1(order=10,  invar=ANTIBODY_CHAR  , level=("4"),popvar=, totalvl=null);
%meandata1(order=11,  invar=INFUSE_DAYS    , roundvar=.1, digit=1);
%mediandata1(order=11.5, invar=INFUSE_DAYS, roundvar=1, digit=0);
%meandata1(order=12,  invar=WEIGHTKG       , roundvar=.1, digit=1);
%meandata1(order=13,  invar=bmi            , roundvar=.01, digit=2);
%meandata1(order=13.5,  invar=AUC_RESULT   , roundvar=.001, digit=3);
%meandata1(order=14,  invar=GLYC_HEMO , roundvar=.01, digit=2);
%meandata1(order=15,  invar=total_insulin_kg     , roundvar=.001, digit=3);
%freqdata1(order=18,  invar=DR_NET  , level=("0"),popvar=, totalvl=null);
%freqdata1(order=19,  invar=DR_NET  , level=("1"),popvar=, totalvl=null);
%freqdata1(order=20,  invar=DR_NET  , level=("2"),popvar=, totalvl=null);
%freqdata1(order=21,  invar=DR_NET  , level=("3"),popvar=, totalvl=null);

proc freq data=treatment_table;
  tables treatmentname/missing list;
  title3 'N pre merge';

proc freq data=table1;
  tables TREATMENTNAME/missing list;
  title3 'N?';

data accumtab1;
  set accumfreq1 accummean1 accummedian1 accuminert1;
  if treatmentname=" " then delete;

proc sort data=accumtab1;
  by treatmentname orderer;

proc print data=accumtab1 noobs;
  by treatmentname;
  title3 'Table 1 stats (list)';

data adverseevent_a;
  set adversecsv;
  if ReportType="Initial" and maskid ne . then output;

proc sort data=adverseevent_a;
  by MASKID severitynum;
```

```
data adverseevent_a_short;
  set adverseevent_a;
  by MASKID severitynum;
  if last.MASKID then output;

data table2_1;
  merge table1(in=in_tab1)
        adverseevent_a_short
        ;
  by MASKID;
  if in_tab1;
  if treatmentname=' ' then delete;
  if severitynum=. then severitynum=0;

data table2_2a;
  merge table1(in=in_tab1)
        adverseevent_a
        ;
  by MASKID;
  if in_tab1;
  if treatmentname=' ' then delete;
  if severitynum=. then delete;

proc sort data=table2_2a out=table2_2b nodupkeys;
  by MASKID CATEGORY;

proc freq data=table2_1;
  tables Severitynum*treatmentname/missing nopercent norow;
  title3 'Table 2';

proc freq data=table2_2a;
  tables Category*treatmentname/missing norow nopercent nocol;
  title3 'Table 3 part a (No of events)';
  where category ne " ";

proc freq data=table2_2b;
  tables Category*treatmentname/missing norow nopercent nocol;
  title3 'Table 3 part b (No of subjects)';
  where category ne " ";
```