

Dataset Integrity Check for
Prediction of Arteriovenous Fistula
Clinical Maturation from Postoperative
Ultrasound Measurements: Findings
from the Hemodialysis Fistula
Maturation Study – Robbin et al

Prepared by Jane Rideau Demuth

IMS Inc.

3901 Calverton Blvd, Suite 200 Calverton, MD 20705

May 28, 2020

Contents

1 Standard Disclaimer	2
2 Study Background	2
3 Archived Datasets	2
4 Statistical Methods	2
5 Results	3
6 Conclusions	3
7 References	3
Table A: Variables used to replicate Table 1: Baseline demographic and clinical characteristics of the study cohort (n=602)	4
Table B: Comparison of values computed in integrity check to reference article Table 1 values	5
Attachment A: SAS Code	6

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

The Hemodialysis Fistula Maturation (HFM) study was a multicenter, prospective cohort study of 602 patients in the United States who received new single-stage arteriovenous fistula (AVFs). 22 Participants underwent standardized ultrasound examinations to map the upper extremity vessels preoperatively and to obtain multiple measurements postoperatively. The study investigated the relationships of AVF blood flow, diameter, and depth, measured postoperatively at 1 day (0–3, targeting 1 day) and 2 and 6 weeks, with unassisted and overall (assisted and unassisted) AVF clinical maturation, and whether these ultrasound measurements could predict clinical maturation accurately enough for practical use.

3 Archived Datasets

All the SAS data files, as provided by the Data Coordinating Center (DCC), are located in the HFMC folder in the data package. For this replication, variables were taken from the “f201.sas7bdat,” “f202.sas7bdat,” “f203.sas7bdat,” “f204.sas7bdat,” “f230.sas7bdat,” “f701.sas7bdat” and “f240.sas7bdat” datasets.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Robbin et al [1] in the Journal of the American Society of Nephrology in 2018. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Baseline demographic and clinical characteristics of the study cohort (n=602), 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

With the exception of non-reproducible vascular calcification, the results of the replication are almost an exact match to the published results.

7 References

[1] Robbin M, Greene T, Allon M, et al. Prediction of Arteriovenous Fistula Clinical Maturation from Postoperative Ultrasound Measurements: Findings from the Hemodialysis Fistula Maturation Study. *Journal of the American Society of Nephrology* volume 29, issue 11, pages 2735-2744 (2018).

Table A: Variables used to replicate Table 1: Baseline demographic and clinical characteristics of the study cohort (n=602)

Table Variable	dataset.variable
Age	f201.age
Women	f201.gender
Black	f202.race
Hemodialysis	f204.hemo_d
Diabetes	f202.diabetes
Peripheral artery disease	f202.l_leg_amp, f202.r_leg_amp, f202.l_leg_rsn, f202.r_leg_rsn
Coronary artery disease	f202.mi, f202.angina, f202.bypass, f202.angioplasty
Upper arm fistula	f231.cannulated, f231.artery, f231.vein
Body mass index	f203.ht_cm, f203.wt_kg
Vascular calcification	Not reproducible

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

Variable	MS n (%) or Median (Percentiles)	IMS n (%) or Median (Percentiles)	Difference
Baseline population	602	602	0
Age, yr, median (10 th and 90 th percentiles)	56.4 (35.8, 71.9)	56 (35, 71)	0.4 (0.8, 0.9)
Women, n (%)	180 (30)	179 (30)	1 (0)
Black, n (%)	264 (44)	264 (44)	0 (0)
Hemodialysis, n (%)	383 (64)	382 (63)	1 (1)
Diabetes, n (%)	353 (59)	353 (59)	0 (0)
Peripheral artery disease, n (%)	91 (15)	91 (15)	0 (0)
Coronary artery disease, n (%)	156 (26)	156 (26)	0 (0)
Upper arm fistula, n (%)	459 (76)	459 (76)	0 (0)
Body mass index, kg/m ² , median (15 th -85 th percentiles)	29.3 (21.7, 40.1)	29.3 (22.7, 38.2)	0 (1.0, 1.9)
Vascular calcification, n (%)	265 (44)	N/A	N/A

Attachment A: SAS Code

```
options mprint nocentre linesize=147 validvarname=uppercase;

title "Program:
/prj/niddk/ims_analysis/HFMC/prog_initial_analysis/DSIC.paper.check.HFMC.y2020m03d03.sas";
title2 "This program reviews the HFMC data.";

/*****

programmer: Jane Rideau Demuth

platform: LINUX SASv9.4

date: 14th March 2019

purpose: See title2.

*****/

*****;
*** formats ***;
*****;
proc format;
  value nmsgf
    . = ' '
    low-high = '###'
  ;
  value $cmsgf
    ' ' = ' '
    other = '$$$'
  ;
  value racef
    1 = 'Native American'
    2 = 'Asian'
    3 = 'Hawai''ian/PacIslander'
    4 = 'Black'
    5 = 'White'
    6 = 'Multiracial'
    9 = 'Unknown'
  ;
run;

%include "/prj/niddk/ims_analysis/HFMC/private_orig_data/forms_HFMformat.sas";

*****;
*** input files ***;
*****;
libname pcsasin "/prj/niddk/ims_analysis/HFMC/private_created_data/masked_by_IMS/";
libname pcsasor "/prj/niddk/ims_analysis/HFMC/private_orig_data/HFM_Archive_03032020/";

libname pcsas04 "/prj/niddk/ims_analysis/HFMC/private_orig_data/HFM_Archive_03042020/";

data f201;
  set pcsasin.f201;
  title3 "Input file:
/prj/niddk/ims_analysis/HFMC/private_created_data/masked_by_IMS/f201.sas7bdat";
proc contents data=f201 varnum;
run;

data f202;
  set pcsasin.f202;
  title3 "Input file:
/prj/niddk/ims_analysis/HFMC/private_created_data/masked_by_IMS/f202.sas7bdat";
proc contents data=f202 varnum;
run;

data f203;
  set pcsasin.f203;
```

```

title3 "Input file:
/prj/niddk/ims_analysis/HFMC/private_created_data/masked_by_IMS/f203.sas7bdat";
proc contents data=f203 varnum;
run;

data f204;
  set pcsasin.f204;
title3 "Input file:
/prj/niddk/ims_analysis/HFMC/private_created_data/masked_by_IMS/f204.sas7bdat";
proc contents data=f204 varnum;
run;

data f230;
  set pcsasin.f230;
title3 "Input file:
/prj/niddk/ims_analysis/HFMC/private_created_data/masked_by_IMS/f230.sas7bdat";
proc contents data=f230 varnum;
run;

data f240;
  set pcsasin.f240;
title3 "Input file:
/prj/niddk/ims_analysis/HFMC/private_created_data/masked_by_IMS/f240.sas7bdat";
proc contents data=f240 varnum;
run;

data f701 (drop=notes);
  set pcsasor.f701;
title3 "Input file:
/prj/niddk/ims_analysis/HFMC/private_orig_data/HFM_Archive_03032020/f701.sas7bdat";
proc contents data=f701 varnum;
run;

data f231;
  set pcsas04.f231;
title3 "Input file:
/prj/niddk/ims_analysis/HFMC/private_orig_data/HFM_Archive_03042020/f231.sas7bdat";
proc contents data=f231 varnum;
run;

* final spreadsheet w/ classifications for 'upper arm fistula';
proc import
datafile="/prj/niddk/ims_analysis/HFMC/private_orig_data/HFM_FistConfAdjudication_09Apr2014_final
.csv"
dbms=csv
out=work.fistconf;
run;

proc contents data=fistconf;
run;

proc freq data=fistconf;
  format pid nmsgf.;
run;

data fistconf (keep=pid pidnum fist_loc_adj ADJ_CONFIGURATIONFINAL);
  set fistconf (rename=(pid=pidnum));

*
"... We had series of conference calls to agree on the type of study fistula and classified
them using ?short? and ?long? classification categories below.
The adjudication process was handled through spreadsheets, e-mails and discussions.

"Long" list
  1 = "UA CV (and prox FA CV, median antecubital vein) and brachial artery"
  2 = "FA CV and wrist radial artery"
  3 = "UA BaV (and median antecubital vein) and brachial artery - transposition"

```

```

4 = "FA BaV and radial artery - transposition"
5 = "UA CV and proximal radial artery"
6 = "UA BaV and proximal radial artery - transposition"
7 = "UA BaV and proximal ulnar artery- transposition"
8 = "UA BrV and brachial artery - transposition"
9 = "UA BrV and proximal radial artery - transposition"
10 = "FA CV and proximal ulnar artery - transposition"
11 = "UA CV and antecubital or proximal forearm radial or ulnar artery"
12 = "UA BaV and antecubital or proximal forearm radial or ulnar - transposition"
13 = "UA BrV and antecubital artery - transposition"
16 = "FA CV and wrist ulnar artery - transposition"

...

;
fist_loc_adj = .;
if ADJ_CONFIGURATIONFINAL in (2,4,10,16) then fist_loc_adj = 1 ; * forearm;
if ADJ_CONFIGURATIONFINAL in (1,3,5,6,7,8,9,11,12,13) then fist_loc_adj = 2 ; ** upper arm;

pid = "0" || compress(put(pidnum,5.));

label fist_loc_adj = "Fistula Location";
run;

proc freq data=fistconf;
tables fist_loc_adj* ADJ_CONFIGURATIONFINAL/list missing;
* tables pid*pidnum/list missing;
title3 "check upper arm fistula";
run;

*****;
*** determine base population ***;
*****;

proc sort data=f201;
by pid;
run;

proc sort data=f204 NODUPKEY;
by pid HEMO_D;
run;

proc sort data=f202;
by pid;
run;

data f202;
set f202;

* Peripheral artery disease (PAD) was defined as history of at least one of these
comorbidities: ;
* - Claudication: F202, Q30k: CLAUD
;
* - Lower extremity angioplasty or bypass surgery: F202, Q30m: LOW_BYPASS
;
* - Non-traumatic amputation: based on F202, Q25. Here is the code used:
;
amput_pad = 0; ** non-traumatic amputation ;
if (l_leg_amp + r_leg_amp = 0) then amput_pad = 0 ;
if (l_leg_rsn in (., 1) and r_leg_rsn in (., 1) ) then amput_pad = 0;
if ((l_leg_amp in (1,2,3,4) and l_leg_rsn = 0) OR (r_leg_amp in (1,2,3,4) and r_leg_rsn = 0))
then amput_pad = 1;

if (claud = 1) or (low_bypass=1) or (amput_pad=1) then pad=1;
else pad=0;

* Coronary artery disease (CAD) was defined as history of at least one of these
comorbidities: ;
* - Myocardial infraction: F202, Q30b: MI ;
* - Angina: F202, Q30c: ANGINA ;

```

```

* - Coronary artery bypass surgery: F202, Q30d: BYPASS ;
* - Percutaneous coronary intervention: F202, Q30e: ANGIOPLASTY ;
if (mi = 1) or (angina = 1) or (bypass = 1) or (angioplasty = 1) then cad = 1;
else cad = 0;

label pad = "Peripheral Artery Disease"
cad = "Coronary Artery Disease"
;

run;

proc freq data=f202;
tables amput_pad*l_leg_amp*r_leg_amp*1_leg_rsn*r_leg_rsn/list missing;
tables pad*amput_pad*claud*low_bypass/list missing;
tables cad*mi*angina*bypass*angioplasty/list missing;
title3 "check PAD and CAD";
run;

proc sort data=f203;
by pid visit_dt;
run;

data f203;
set f203;
by pid visit_dt;
if not(first.pid and last.pid) then put pid= visit_dt= vist= ht_cm= wt_kg=;
if first.pid then output;
run;

proc sort data=f230;
by pid;
run;

proc sort data=f240;
by pid;
run;

proc sort data=f701;
by pid;
run;

data f701dups;
set f701;
by pid;
if not(first.pid and last.pid);
run;

proc print data=f701dups;
by pid;
id pid;
var intima media;
title3 "f701 dups";
run;

data f701;
set f701;
* vascular calcification;
vein_ca = sum(intima, media);

if vein_ca = 0 then vein_ca_2 = 0;
if vein_ca gt 0 then vein_ca_2 = 1;

label vein_ca = "Vein Calcification - 3 levels"
vein_ca_2 = "Vein Calcification - Y/N" ;
run;

proc freq data=f701;
tables vein_ca*intima*media/list missing;

```

```

tables vein_ca_2*vein_ca/list missing;
title3 "check vascular calcification";
run;

proc sort data=f701 nodupkey;
  by pid vein_ca_2;
run;

proc sort data=fistconf;
  by pid;
run;

data basepop(keep=pid in_recruit in_avfsurg avf_created vein_ca_2 in_dropout
              age gender race dial_stat diabetes pad cad ht_cm wt_kg bmi fist_loc_adj
              HEMO_D);
  merge f201(in=in201 keep=pid age gender dial_stat)
        f202(in=in202 keep=pid race diabetes pad cad)
        f203(in=in203 keep=pid ht_cm wt_kg)
        f204 (in= in204 keep = pid HEMO_D)
        f230(in=in230 keep=pid avf_created)
        f240(in=in240 keep=pid)
        f701      (keep=pid vein_ca_2)
        fistconf  (keep=pid fist_loc_adj) ;
  by pid;
  if not(first.pid and last.pid) then abort;

  in_recruit = in201;
  in_avfsurg = in230;
  in_dropout = in240;

  bmi = wt_kg / ((ht_cm/100)**2);

  label bmi = 'BMI';
run;

proc freq data=basepop;
  title3 'Does RECRUIT + AVFSURG + AVF_CREATED - DROPOUT = the 602 baseline pop? YES!';
  tables in_recruit*in_avfsurg*avf_created*in_dropout / missing list;
run;

*****;
*** replicate numbers from table 1 ***;
*****;

proc freq data=basepop;
  title3 'Table 1. Baseline demographic and clinical characteristics of the study cohort
(n=602)';
  where in_recruit and in_avfsurg and avf_created and NOT in_dropout;
  tables gender race dial_stat HEMO_D diabetes pad cad fist_loc_adj vein_ca_2/ missing list;
  format race racef.;
run;

proc univariate data=basepop noprint;
  where in_recruit and in_avfsurg and avf_created and NOT in_dropout;
  var age bmi;
  output out=tluni
         pctlpts=10 15 50 85 90
         pctlpre=age_ bmi_
         pctlname=p10 p15_mdn p85 p90;
run;

proc print data=tluni noobs;
  title4 '          (from univariate)';
  var age_p10 age_mdn age_p90 bmi_p15 bmi_mdn bmi_p85;
run;

```