

Integrity Check for Liver Transplant Database (LTD) Baseline Files

As a partial check of the integrity of the LTD datasets archived in the NIDDK data repository, a set of tabulations was performed to verify that published results from the LTD study can be reproduced using the archived datasets. A small number of analyses were performed to duplicate published results for the data reported by the Wei et al [1] in Liver Transplantation and Surgery in January, 1997, and in Charlton et al [2] in Liver Transplantation in September, 2004. The results of this integrity check are described below. The full text of the articles can be found in Attachments 1 and 2, and the SAS code for our tabulations is included in Attachment 3.

Publication using Complete Baseline Data. Wei et al [1] report results for all evaluated subjects (N=1563) and donors (N=1002) over the baseline study period representing recruitment from April 15, 1990 to Jun 30, 1994. These and other totals in the baseline timeframe were derived and compared to totals reported in Wei et al. The derivations reported here match those of Wei et al for evaluated subjects and donors and other categories described below.

Data and Structures. In the data structure of the LTD study, there is no single baseline dataset. The following datasets were used in this baseline stage of the integrity check [Note 1]:

- CE initial evaluations (candidates before transplant)
- FS fulminant liver failures
- LD living donors
- DR donor liver recovery
- CIs hort-term transplant follow-up
- CP immediate post-operative assessment
- MD death reports
- IS intra-operative surgeon assessment form

To produce the total of evaluated patients, unique ID variables based on the latest dates for an ID, from the CE dataset were merged with unique IDs from the FS dataset. The totals for the demographic categories race, blood type and sex, from this stage of the integrity check match the demographic totals given on page 15, Table 1 of Wei et al [1] (see Table 1 below).

Table 1: Comparison of Values for Demographic Categories Computed in Integrity Check to Reference Article Values

Variable	Integrity Check	Wei et al (1997:15)	Attachment 2 Page Number
Male	845	845	12
Female	718	718	12
ABO = A (or "1")	607	607	12
ABO = B	196	196	12
ABO = AB	63	63	12
ABO = O	691	691	12
Race = Caucasian	1216	1216	12
Race = Black	64	64	12
Race = Hispanic	170	170	12
Race = Asia Pacific	70	70	12
Race = (other)	43	43	12
Karnofsky Score = 1,2,3	491	491	13
Karnofsky Score = 4 or 5	594	594	13
Karnofsky Score = 6 or 7	230	230	13
Karnofsky Score = 8,9,10	158	248 ¹	13

Karnofsky Score = missing	90	Not totaled	13
Encephalopathy Stage= B (blank) or none	1108	1109 ²	13
Encephalopathy Stage= Lethargy/Asterixis	288	287 ¹	13
Encephalopathy Stage= Confusion/disorientation	98	98	13
Encephalopathy Stage= Stupor/Coma, but arousable	51	51	13
Encephalopathy Stage= Deep Coma	16	16	13
Encephalopathy Stage= Drug-induced coma	2	2	13
Date of initial evaluation 4/90 – 12/90	351	351	13
Date of initial evaluation 1/91 – 12/91	381	381	13
Date of initial evaluation 1/92 – 12/92	440	440	13
Date of initial evaluation 1/93 – 12/93	384	384	13
Date of initial evaluation 1/94 – 6/94	6	7 ²	13
Date of initial evaluation = missing	1	Not totaled	13
Notes: ¹ Wei et al appear to have included missing values in the 8,9,10 scores ¹ A minor discrepancy was noted			

To produce the total of donors, IDs for living and dead donors (LD+DR datasets) were combined in one dataset. These totals of 1002 donors and 1563 evaluated patients computed in the integrity check are the same as in Wei et al (1997:10,14).

To get the total number of transplants (n=916) reported in Wei et al (1997), non-duplicate IDs from CI were combined with non-duplicate IDs from CP and cases where death (MD) occurred within two days before or after the immediate post-operative assessment were deleted.

To produce the totals of fulminant and single/multi organ transplants as shown on page 14, figure 2 of Wei et al (1997), the 916 transplants identified as described above (with a flag for “fulminant” based on origin in the FS dataset) were merged by ID with the IS dataset of unique ID cases with OTX variable values of 1 or 0 [Note 2]. A case of OTX = 1 (yes) means that other transplants were done involving the recipient. A case of OTX = 0 (no) means that no other transplants were done involving the recipient. The totals computed in the integrity check are the same as in Wei et al (see Table 2 below).

Table 2: Comparison of Values for Transplant Totals Computed in Integrity Check to Reference Article Values

Transplant totals	Integrity Check	Wei et al (1997:14)	Attachment 2 Page Number
All transplants	916	916	9
Fulminant	58	58	10
Non-fulminant	858	858	10

In addition, these basic manipulations of the baseline data provide an indirect confirmation of the total number of adult transplant recipients given in Charlton et al (2004):1121, where 805 is given as the total number of adult transplant recipients in the LTD. This number is just the total number of adult (age greater than or equal to 16 at initial examination or transplant surgery) recipients with either missing or 0 OTX variable values (751) plus 2 times the the number of multiple transplants (ie where OTX = 1): (27 X 2 = 54 + 751 = 805).

Notes

1. Analysis was done using a copy of the database provided by NIDDK Data Repository on June 29, 2005. The SAS datasets used were created on January 26, 1998. There are about 50 other baseline datasets that this analysis does not examine. Moreover, since follow-up data was collected until November 2002, (see Charlton M, et al (2004):1121) this analysis clearly applies only to the baseline data.
2. The 1997 paper and the 2004 paper seem to use the OTX variable in different ways since the 1997 paper seems to exclude patients with missing OTX values and the 2004 paper seems to include them.

Procedural Summary of the SAS Programs

A. baseadult.sas

1. The CE (initial examinations) data is sorted by patient id and date of initial examination.
2. For each patient, the data from their latest exam is taken and all other exams are omitted.
3. Then the fulminant data is read and each record is marked as "fulmin=1"
4. Then the fulminant and CE data is placed in a dataset called CEFS.
5. From CEFS, totals of males and females are made.
6. To get the total of donors, datasets with living and dead donors were combined.
7. The next set of steps has to do with getting transplant totals.
8. The first step in getting transplant totals is to take the short-term transplant follow-up records (CI) and take only one per patient id.
9. Then take the immediate post-operative assessments (CP) and take only one per patient id.
10. Then combine CI and CP and mark them as "transplant = 1"
11. Then take the death records (MD) and sort them by date.
12. Then take the CI and CP records and sort them by date.
13. Then combine the CI and CP and MD records by patient ID.
14. If the difference between the CI or CP (transplant) date and death date is less than three days either way then remove the patient ID from the transplant data. The patients who remain in the transplant data at this point are the successful transplants.
15. To find the successful adult transplants, mark successful transplants of greater than 16 years of age at CE exam as adults.
16. To find multiple transplants, use the OTX variable from the IS (intra-operative surgical assessment). An OTX value of 1 is a case of another transplant, ie OTX = 'YES'.
17. From the IS, take patients with OTX values of 1 and 0.
18. With the resulting IS, merge the successful transplants and call the result Trans2.
19. From TRANS2, totals of transplants of different types are made.

B. basedemog.sas

1. The CE (initial examinations) data is sorted by patient id and date of initial examination.
2. For each patient, the data from their latest exam is taken and all other exams are omitted.
3. Then the fulminant data is read and each record is marked as "fulmin=1"
4. Then the fulminant and CE data is placed in a dataset called CEFS.
5. From CEFS, the other demographic totals are made.

C. initcedemog.sas

1. The CE (initial examinations) data is sorted by patient id and date of initial examination.
2. For each patient, the data from their earliest exam is taken and all other exams are omitted.
3. Then the fulminant data is read and each record is marked as "fulmin=1"
4. Then the fulminant and CE data is placed in a dataset called CEFS.
5. From CEFS, the other demographic totals are made.

References

1. Wei Y, Everhart J, Detre K (1997). Long-term results and modeling to predict outcomes in recipients with HCV infection: results of the NIDDK liver transplantation database. **Liver Transplantation and Surgery**, Vol 3, No 1 (January):10-22.
2. Charlton M, Ruppert K, Belle SH, Bass N, Schafer D, Wiesner RH, Detre K, Wei Y, Everhart J. (2004). Long-term results and modeling to predict outcomes in recipients with HCV infection: results of the NIDDK liver transplantation database. **Liver Transplantation** 10:1120-30.

ATTACHMENT 1

"The full text of the article referenced will be provided to approved requestors along with the data archive."

Wei Y, Everhart J, Detre K (1997). Long-term results and modeling to predict outcomes in recipients with HCV infection: results of the NIDDK liver transplantation database. Liver Transplantation and Surgery, Vol 3, No 1 (January):10-22.

ATTACHMENT 2

"The full text of the article referenced will be provided to approved requestors along with the data archive."

Charlton M, Ruppert K, Belle SH, Bass N, Schafer D, Wiesner RH, Detre K, Wei Y, Everhart J. (2004). Long-term results and modeling to predict outcomes in recipients with HCV infection: results of the NIDDK liver transplantation database. Liver Transplantation 10:1120-30.

ATTACHMENT 3

SAS Code and Results for Baseline Tabulations from LTD Datasets in the NIDDK Repository

```

/*****
*
*          baseadult.sas
*
*          Project: NIDDK - LTD
*
*          Data Roadmap and Integrity Check
*
*          Shows and documents a method for deriving
*          the baseline populations of transplant types
*
*          Programmer: P Hickman, RTI Health Solutions
*          Created: 11JAN2006
*
*          Modification log:
*
*
*****

```

Notes (personal communication from Yuling L. Wei):
 "Candidates who were 'fulminant' cases did not have a CE form, but an FS form instead. Also, some patients had more than one CE form if there was a long lapse (>1 year) and had to be re-evaluated. The difference in the numbers is not easily calculated unless you look at each patient to figure out whether they had > 1 CE form, and whether they had FS form. I don't recall any patients having both CE and FS forms, but I may be wrong."

*/

***** this sets up site-specific libraries *****;

```

%include "T:\biostatistics\liver_trp_dbms\Analytic_LTD\programs
\bONID_init.sas" ;
%BONID_init;

```

**** 'raw' is the arbitrary name for the library
 *****;
 /***** 'CE' is the Initial Clinical Evaluation
 5. Date first seen at transplant center for liver transplant
 evaluation

```

OR      Date of re-evaluation      EOM  EOD  EOY
                                MM / DD / YY

```

**** age selection is study-specific
 ****/

```

data CEX;
  set raw.ce;
  CE_AGE = age;
  CE_date = MDY(EOM,EOD,EOY);
  Format CE_date date9.;
  ***keep id ce_age center sex race CE_date ;

run;

```



```

proc sort data = CEX; by ID CE_date; run;

data CEX01;
  set CEX;
  by ID CE_date;
  if last.ID;
  ce_age = age;
run;

Data FS;
  set raw.FS;
  fulmin = 1;
  ****keep id sex cmvgr hcvr ntxm ntxd ntxy bom bod boy ;
run;

**** the 1563 evalutated patients that result from the next data step
**** are the number expected on the basis of an article in
      "The NIDDK Liver Transplantation Database"
**** Liver transplantation and surgery, vol3, no1 (January),
1997:pp10-22;

data CEFS;
  set CEX01 FS;
run;

title1 "Total Evaluated Patients";
title2 "          ";
proc freq data = CEFS; tables sex / list missing missprint; run;

**** 'raw' is the arbitrary name for the library
*****;
****
****
**** The number of living (LD) plus dead donors (DR) corresponds to
the number of donors: 1002
in "The NIDDK Liver Transplantation Database"
Liver transplantation and surgery, vol3, no1 (January), 1997:pp10-
22;
****;

data LDX;
  set raw.ld;
  keep id sex;

run;

data drx;
  set raw.dr;
  keep id sex;

run;

data Donors;
  set drx ldx;
run;

```

```

title1 "Total Donors";
title2 " ";
proc freq data = donors; tables sex / list missing missprint; run;

***** the population of successful transplants that result from the
        next data steps (N=916) corresponds to
***** the number expected on the basis of an article in
        "The NIDDK Liver Transplantation Database"
***** Liver transplantation and surgery, vol3, no1 (January),
1997:pp10-22;

Data cix;
  set raw.ci;
  hello = 1;

  ****keep id sex hello;

run;

proc sort data = cix nodupkey; by id;

run;

Data cpx;
  set raw.cp;
  hello = 1;
  surdate = mdy(eom,eod,eoy);
  format surdate date9.;
  keep id sex hello eom eod eoy surdate;

run;

proc sort data = cpx nodupkey; by id;

run;

Data cpxcix;
  set cpx cix;
  transplant = 1;

run;

proc sort data = cpxcix nodupkey; by id;

run;

data deathz;

```

```

    set raw.md;
    keep id doddate;
run;

data baserep;
merge cpxcix (in = hh) deathz (in=dd);
by id;
if transplant = 1;
durx = doddate - surdate;
if durx in ( 1 -2 2 0 -1) then delete;

run;

title1 "Successful Transplants";
title2 " ";
proc freq data = baserep; tables transplant / list missing missprint;
run;

proc sort data = CEFS nodupkey; by id;

*** Characteristics of Transplant population ;

data trans;
merge baserep (in = bb) CEFS (in = CCCC);
by id;
if bb;
if ce_age >16 then adult = 1;
run;

data trans1;
merge trans (in = bb) cpx (in = CCCC);
by id;
if bb;
fsbdate = mdy(bom,bod,boy);
fagex = surdate - fsbdate;
fage = round((fagex/365.25),.1);
if fage => 16 then adult = 1;
if fage = . then do;
if ce_age => 16 then adult = 1;
end;
run;

data is;
set raw.is;
if otx in ( 0, 1);
ootxoo = otx;
keep id otx ootxoo;

run;

proc sort data = is nodupkey;
by id; run;

data trans2;
merge trans1 (in = bb) IS (in = CCCC);
by id;
if bb;

run;

```

```
/*
  Frequency of single/multi-organ transplants and fulminant cases. The
  multi-organ and fulimant results equal those on
  page 14 figure 2 of Wei et al (1997).
*/
```

```
title1 "Fulminant and Multiple Transplants";
title2 " ";
```

```
proc freq data = trans2; tables fulmin*otx ootxoo*otx / list missing
missprint; run;
```

```
data Adult;
  set trans2;
  if '30JUN1994'd => surdate => '15APR1990'd;
  if fage => 16 then adult = 1;
  if adult ne 1 then do;
    if ce_age => 16 then adult = 1;
  end;
```

```
run;
```

```
title1 "Adult Transplants";
title2 " ";
```

```
proc freq data = adult; tables otx*adult / list missing missprint;
run;
```

Total Evaluated Patients
1

The FREQ Procedure

SEX				
SEX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Male	845	54.06	845	54.06
Female	718	45.94	1563	100.00

generated on 15FEB2006

Total Donors
2

The FREQ Procedure

SEX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	983	98.10	983	98.10
Male	8	0.80	991	98.90
Female	11	1.10	1002	100.00

generated on 15FEB2006

Successful Transplants
3

The FREQ Procedure

transplant	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	916	100.00	916	100.00

generated on 15FEB2006

Fulminant and Multiple Transplants
4

The FREQ Procedure

fulmin	OTX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	.	7	0.76	7	0.76
.	No	819	89.41	826	90.17
.	Yes	32	3.49	858	93.67
1	No	57	6.22	915	99.89
1	Yes	1	0.11	916	100.00

ootxoo	OTX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	.	7	0.76	7	0.76
0	No	876	95.63	883	96.40
1	Yes	33	3.60	916	100.00

Adult Transplants
5

The FREQ Procedure

OTX	adult	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	1	7	0.80	7	0.80
No	.	93	10.60	100	11.40
No	1	744	84.83	844	96.24
Yes	.	6	0.68	850	96.92
Yes	1	27	3.08	877	100.00

generated on 15FEB2006

```

/*****
*
*       basedemog.sas
*
*       Project: NIDDK - LTD
*
*       Data Roadmap and Integrity Check
*
*       Shows and documents a method for deriving
*       the baseline populations of transplant types
*       and matches Table 1 in Wei 1997 for the total population
*       for 3 demographic variables.
*
*       Programmer: P Hickman, RTI Health Solutions
*       Created: 11JAN2006
*
*       Modification log:
*
*
*****

Notes (personal communication from Yuling L. Wei):
    "Candidates who were 'fulminant' cases did not have a CE form, but
an FS
form instead. Also, some patients had more than one CE form if there
was a long
lapse (>1 year) and had to be re-evaluated. The difference in the
numbers is not easily
calculated unless you look at each patient to figure out whether they
had > 1 CE form,
and whether they had FS form.
I don't recall any patients having both CE and FS forms, but I may be
wrong."

    */

***** this sets up site-specific libraries *****/

%include "T:\biostatistics\liver_trp_dbms\Analytic_LTD\programs
\bONID_init.sas" ;
%BONID_init;

**** 'raw' is the arbitrary name for the library
*****/
/***** 'CE' is the Initial Clinical Evaluation
5. Date first seen at transplant center for liver transplant
evaluation

OR    Date of re-evaluation          EOM  EOD  EOY
                                MM / DD / YY

***** age selection is study-specific
*****/

data CEX;
  set raw.ce;
  CE_AGE = age;
  CE_date = MDY(EOM,EOD,EOY);
  Format CE_date date9.;
  ****keep id ce_age center sex race CE_date ;

```

```

run;

proc sort data = CEX; by ID CE_date; run;

data CEX01;
  set CEX;
  by ID CE_date;
  if last.ID;
  ce_age = age;
run;

Data FS;
  set raw.FS;
  fulmin = 1;
  ****keep id sex cmvgr hcvr ntxm ntxd ntxy bom bod boy ;
run;

***** the 1563 evalutated patients that result from the next data step
***** are the number expected on the basis of an article in
          "The NIDDK Liver Transplantation Database"
***** Liver transplantation and surgery, vol3, no1 (January),
1997:pp10-22;

data CEFS;
  set CEX01 FS;
  run;

title1 "Total Evaluated Patients";
title2 "          ";

proc freq data = CEFS; tables sex ABO race/ list missing missprint;
run;

```

Total Evaluated Patients

1

The FREQ Procedure

SEX

SEX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Male	845	54.06	845	54.06
Female	718	45.94	1563	100.00

BLOOD TYPE

ABO	Frequency	Percent	Cumulative Frequency	Cumulative Percent
B	6	0.38	6	0.38
1	607	38.84	613	39.22
B	196	12.54	809	51.76
AB	63	4.03	872	55.79
O	691	44.21	1563	100.00

RACE

RACE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Caucasian	1216	77.80	1216	77.80
Black	64	4.09	1280	81.89
Am. Indian/Eskimo	20	1.28	1300	83.17
Hispanic	170	10.88	1470	94.05
Oriental Pacific	70	4.48	1540	98.53
Mideast. Arab	8	0.51	1548	99.04
Indian Subcont.	12	0.77	1560	99.81
Other	3	0.19	1563	100.00

generated on 03MAR2006

```

/*****
*
*       initcedemog.sas
*
*       Project: NIDDK - LTD
*
*       Data Roadmap and Integrity Check
*
*       Shows and documents a method for deriving
*       the baseline populations of transplant types
*       and matches Table 1 in Wei 1997 for the total population
*       for 4 demographic variables.
*
*       Programmer: P Hickman, RTI Health Solutions
*       Created: 03MAR2006
*
*       Modification log:
*
*
*****

Notes (personal communication from Yuling L. Wei):
    "Candidates who were 'fulminant' cases did not have a CE form, but
an FS
form instead.  Also, some patients had more than one CE form if there
was a long
lapse (>1 year) and had to be re-evaluated.  The difference in the
numbers is not easily
calculated unless you look at each patient to figure out whether they
had > 1 CE form,
and whether they had FS form.
I don't recall any patients having both CE and FS forms, but I may be
wrong."

    */

***** this sets up site-specific libraries *****/

%include "T:\biostatistics\liver_trp_dbms\Analytic_LTD\programs
\bONID_init.sas" ;
%BONID_init;

**** 'raw' is the arbitrary name for the library
*****/
/***** 'CE' is the Initial Clinical Evaluation
5. Date first seen at transplant center for liver transplant
evaluation

OR    Date of re-evaluation          EOM  EOD  EOY
                                MM / DD / YY

***** age selection is study-specific
*****/

data CEX;
  set raw.ce;
  CE_AGE = age;
  CE_date = MDY(EOM,EOD,EOY);
  Format CE_date date9.;
  ****keep id ce_age center sex race CE_date ;

```

```

run;

proc sort data = CEX; by ID CE_date; run;

data CEX01;
  set CEX;
  by ID CE_date;
  if first.ID;
  ce_age = age;
run;

Data FS;
  set raw.FS;
  fulmin = 1;
  ****keep id sex cmvgr hcvr ntxm ntxd ntxy bom bod boy ;
run;

***** the 1563 evalutated patients that result from the next data step
***** are the number expected on the basis of an article in
        "The NIDDK Liver Transplantation Database"
***** Liver transplantation and surgery, vol3, nol (January),
1997:pp10-22;

data CEFS;
  set CEX01 FS;
  length cdx $ 10;
  d490 = '01APR1990'd;
  d1290 = '31DEC1990'd;
  d0191 = '01JAN1991'd;
  d1291 = '31DEC1991'd;
  d0192 = '01JAN1992'd;
  d1292 = '31DEC1992'd;
  d0193 = '01JAN1993'd;
  d1293 = '31DEC1993'd;
  d0694 = '01JUN1994'd;

  if d1290 => eodate =>d490 then eodatex = '4/90 to 12/90  ';
  if d1291 => eodate =>d0191 then eodatex = '01/91 to 12/91  ';
  if d1292 => eodate =>d0192 then eodatex = '01/92 to 12/92  ';
  if d1293 => eodate =>d0193 then eodatex = '01/93 to 12/93  ';
  if d0694 => eodate >d1293 then eodatex = '01/94 to 06/94  ';

  if kar in ( 1 2 3) then karx = '1to3  ';
  if kar in ( 4 5) then karx = '4or5  ';
  if kar in ( 6 7) then karx = '6or7  ';
  if kar in ( 8 9 10) then karx = '8to10';

  oencs = encs;
  cdx = compress(cdx1||'/'||cdx2||'/'||cdx3);
  if cdx1 in (8 11 12 13 15 17 18 19 20 21 22 23 24 25 26 30 31 32
33 34 35) then bigdiag = 'Chronic  ';
  else if cdx1 in (1 2 3 5 6 ) then bigdiag = 'ACUTE  ';
  else bigdiag = 'Other  ';

  keep sex ABO race karx  cdx1 cdx2 cdx3 cdx eodatex encs oencs
bigdiag;
run;

/*
proc contents data = CEFS; run;

```

```
*/  
title1 "Total Evaluated Patients (Using the First Evaluation)";  
title2 "          ";  
  
proc freq data = CEFS;  
  tables karx encs*oencs eodatex / nopercent list missing missprint;  
run;
```

Total Evaluated Patients (Using the First Evaluation)

1

The FREQ Procedure

karx	Frequency	Cumulative Frequency
1to3	90	90
4or5	491	581
6or7	594	1175
8to10	230	1405
	158	1563

ENCS	oencs	Frequency	Cumulative Frequency
B	B	1108	1108
Lethargy/Asterixis	1	288	1396
Confusion/Disorient	2	98	1494
Stupor/Coma Arousabl	3	51	1545
Deep coma	4	16	1561
Drug-induced coma	5	2	1563

eodatex	Frequency	Cumulative Frequency
01/91 to 12/91	381	382
01/92 to 12/92	440	822
01/93 to 12/93	384	1206
01/94 to 06/94	6	1212
4/90 to 12/90	351	1563

generated on 17MAR2006