



REPORT ON THE FINAL DATA BASE

Contract #270-96-0007

Volume II: Users Guide, Source Code and Data Dictionary

November 2, 1999

Submitted to:

Project Officer: Joan Dilonardo
SAMHSA Services Administration
Center for Substance Abuse Treatment (CSAT)
Rockwall II Building, Suite 740
5600 Fishers Lane
Rockville, MD 20852

Submitted by:

The MEDSTAT Group, Inc.
5425 Hollister Avenue
Suite 140
Santa Barbara, CA 93111

Unless disclosure is required by the Freedom of Information act 5 U.S.C. 552, as amended, this proposal or quotation includes data that shall not be disclosed outside the Government and shall not be duplicated, used or disclosed - in whole or in part -- for any purpose other than to evaluate this proposal or quotation. If, however, a contract is awarded to this offeror or quoter as a result of -- or in connection with -- the submission of this data, the Government shall have the right to duplicate, use, or disclose the data to the extent provided in the resulting contract. This restriction does not limit the Government's right to use information contained in this data if it is obtained from another source without restriction. The data subject to this restriction are contained in sheets all.

TABLE OF CONTENTS

Report on the Final Data Base i

Overview 1

Integrated Data Base Design..... 1

 Data Base Files 1

 Person Summary Data Set 2

 Processing Overview 2

SAS Loads and Recodes, Medicaid Data – Program 100..... 4

 Eligibility Data..... 5

 Claims/Encounters Data..... 8

 Provider Data 15

SAS Loads and Recodes, MH/AOD Data – Program 200..... 15

 Input Data..... 16

 Output Data..... 16

 Temporary Storage..... 23

 Processing Information 23

Client Linking – Program 300 27

 Programs 27

 Input Data..... 28

 Output Data..... 28

 Temporary Storage..... 29

ID Mapping – Program 400 30

 Programs 30

 Input Data..... 31

 Output Data..... 31

 Temporary Storage..... 32

Data Base Build – Program 600..... 33

 Programs 35

 Input Data..... 36

 Output Data..... 36

 Temporary Storage..... 45

APPENDIX A - Consistency Codes 1

APPENDIX B - Sample Code Diagnosis Based Flags 1

APPENDIX C - Inclusion Criteria Code 1

APPENDIX D – Delaware Source CODE AND DATA DICTIONARY..... 1

APPENDIX E – OKLAHOMA Source CODE AND DATA DICTIONARY 1

APPENDIX F – WASHINGTON Source CODE AND DATA DICTIONARY..... 1

Appendix A – Consistency Code

Appendix B – Sample Code Diagnosis Based Flags

Appendix C – Inclusion Criteria Code

Appendix D – Delaware Source Code and Data Dictionary

Appendix E – Oklahoma Source Code and Data Dictionary

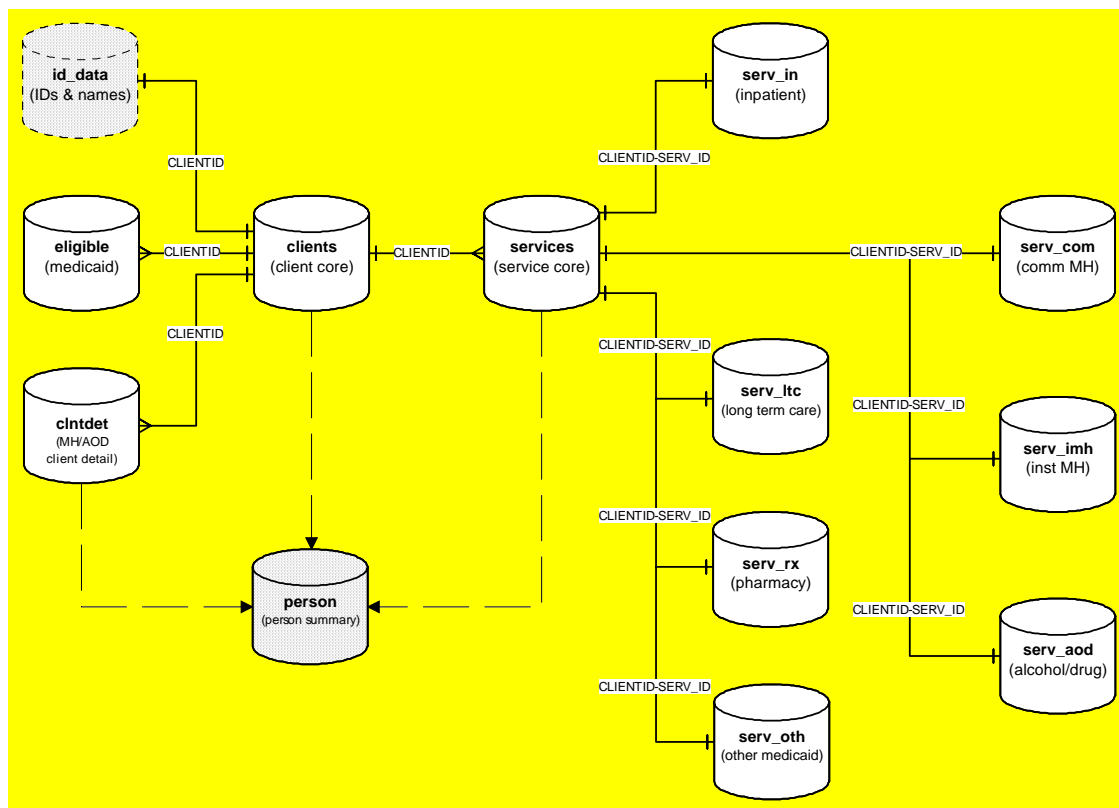
Appendix F – Washington Source Code and Data Dictionary

OVERVIEW

This guide documents the series of steps used to create an integrated data base of MH/AOD and Medical information, specifically a combination of MH/AOD and Medicaid data. As such, it provides a general overview towards creating uniform files from a variety of data sources and formats. The intended audience for this document consists of programmers and planners considering the implementation of a State- or regional-level IDB.

Implementation and programming specifics can be found in Appendixes D, E, and F. The task of creating an integrated data base has been divided into several program groups; each designed to address a specific issue. In turn, each program group has been further divided into a number of steps. The users guide describes the program group organization and develops guidelines for the division of program groups into steps. In addition, the users guide describes the desired output and key issues in creating that output. Issues addressed include file input and output, naming conventions, and required variables. Specifics regarding the steps of a program group can be found in Appendixes D, E, and F.

INTEGRATED DATA BASE DESIGN



DATA BASE FILES

Client Data Sets

- Client Core – CLIENTS
- Medicaid Eligibility – ELIGIBLE

IDB Users Guide

- MH/AOD Client Detail – CLNTDET
- Client Identifying Data – ID_DATA
This data set is used for data base construction only, it is not a deliverable file.

Service Data Sets

- Service Core – SERVICES
- Medicaid Inpatient Service Detail – SERV_IN
- Medicaid Long Term Care Service Detail – SERV_LTC
- Medicaid Pharmacy Service Detail – SERV_RX
- Outpatient/Other Medicaid Service Detail – SERV_OTH
- Community Mental Health Treatment Detail – SERV_COM
- Institutional Mental Health Treatment Detail – SERV_IMH
- Alcohol/Other Drug Treatment Detail – SERV_AOD

PERSON SUMMARY DATA SET

The person summary data set is a supplemental file which is created after production is complete in a separate process.

PROCESSING OVERVIEW

Processing data into the Integrated Data Base (IDB) proceeded through five separate tasks, or program sequences. Together, these five program sequences transformed the many raw data files into the IDB. Along the way, data was recoded, standardized, and integrated with data from other sources. The five programming sequences are:

- Program 100 – Data loads and processing of Medicaid data,
- Program 200 – Data loads and processing of MH/AOD data,
- Program 300 – Client ID linking
- Program 400 – Client ID Mapping
- Program 600 – Data Base Build

For each State, five file directories have been created, one for each program sequence (i.e. program100 and program200). All subtasks within a program sequence are stored, with names relating to the program sequence, in these directories. Subtask program files are named stepNxx where N relates to the program sequence, and xx indicates the program files ordinal place within the program sequence. Consequently, SAS programs in the 100 sequence have names such as step110 and step120.

Most programs make use of uniform processing modules, usually in the form of SAS macro code. Common tasks such as defining variable attributes and standardization of name variables are handled with these SAS macro modules. These modules are common to multiple programs and States, and are stored in

a separate "include library" directory. All include files are listed in the "Common Modules" section. These files generally contain one SAS macro per file with the file name reflecting the name of the macro.

Data Loads and Processing

Two general types of Medicaid data are processed: eligibility and claims/encounters. These programs are named step1xx, where xx are digits to indicate the relative order of the program in the processing stream (i.e. step110 and step120). Variable recodes to standardized values occur at this stage.

Processing programs of MH/AOD data are named step2xx, where xx are digits to indicate the relative order of the program in the processing stream (i.e. step210 and step220). Variable recodes to standardized values occur at this stage.

Client Linking

Client linking is divided into two programs in order to accommodate a manual review. The two client linking programs are named step350 and step352. Step350 performs the initial data joins and testing, creating a report used for the manual review. Step352 incorporates the results of the manual review and finalizes the ID links.

ID Mapping

ID mapping completes the Client linking process and determines the IDB population. Using the links determined in the earlier program stream and the list of Medicaid IDs which received some type of MH/AOD services through Medicaid, the first ID mapping program (step410) determines the population of the IDB to include

- all persons receiving treatment from an MH/AOD agency,
- all persons receiving MH/AOD services billed through Medicaid, and
- all IDs linked to the two previous conditions.

The second program in the sequence, step420 maps the source IDs to a newly generated CLIENTID. The final program (step420) in the sequence generates a simple report of observation counts.

Data Base Build

The final data base build programs use the files created from the program 100 and 200 sequences with the ID mapping from program 400 to build the data sets of the IDB. There are six programs in this sequence:

- step610 – creates the CLIENT data set of client core information, along with the ID_DATA data set (not included in the IDB).
- step620 – creates the ELIGIBLE data set of Medicaid eligibility information.
- step630 – creates the CLNTDET data set of MH/AOD detail information.
- step640 – creates the MH/AOD service detail data sets: SERV_COM, SERV_IMH, and SERV_AOD.

IDB Users Guide

- step650 – creates the Medicaid service detail data sets: SERV_IN, SERV_LTC, SERV_RX, and SERV_OTH.
- step660 – creates the SERVICES data set of service core information.

SAS LOADS AND RECODES, MEDICAID DATA – PROGRAM 100

This section describes the program names, locations and attributes of SAS routines used in processing State data. File and directory locations are reflect those used during the production of the final data base and can be modified as necessary.

Program Names

Name programs “step1nn” where nn are digits (i.e. step110, step115).

Program Locations

Programs should be saved on the server in the appropriate directories:

- Delaware /sas612/samhsa/de/program100/
- Oklahoma /sas612/samhsa/ok/program100/
- Washington /sas612/samhsa/wa/program100/

Macros and File Includes

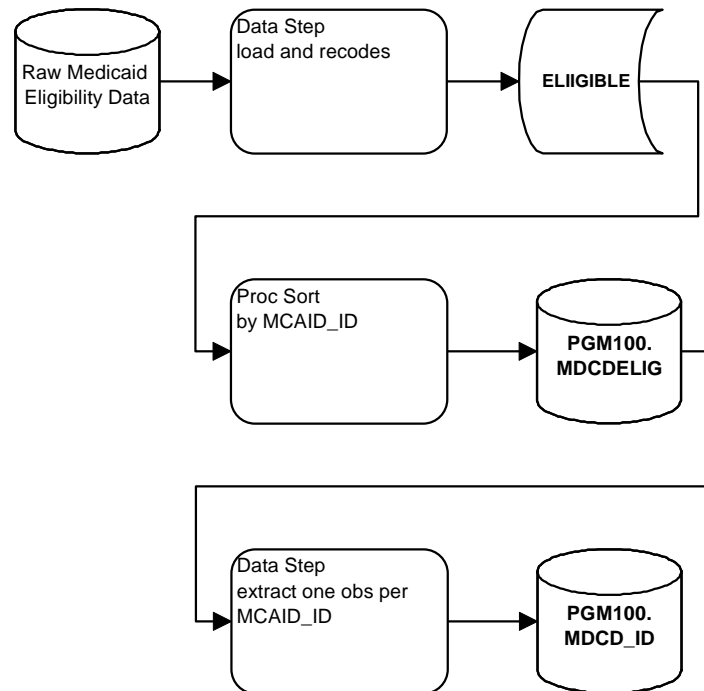
All macros and file includes used by these programs can be found on the following server directory:
/sas612/samhsa/@inclib/

SAS Formats

All custom SAS formats used by these programs can be found on the following server directory:
/sas612/samhsa/@fmtlib/

Documentation

All programs should document the program name, programmer, and purpose of job at the top of the listing.

ELIGIBILITY DATA**Processing****Input Files**

Input files should be found on the server:

- Delaware /samhsa/de/1996/rawdata/medicaid/
- Oklahoma /samhsa/ok/1996/rawdata/medicaid/
- Washington /samhsa/wa/1996/rawdata/medicaid/

Output Data Sets***File Location***

All program 100 output should be placed in the appropriate server directory:

- Delaware /samhsa/de/1996/program100/
- Oklahoma /samhsa/ok/1996/program100/
- Washington /samhsa/wa/1996/program100/

Data Sets

MDCD_ID – Identifying Data

- Sorted by Medicaid ID (MCAID_ID)
- One observation per Medicaid ID
- When multiple eligibility records are available for a recipient, populate variables with the most recent non-missing data.
- Contains variables found in the data sets CLIENT and ID_DATA.

MDCD_ID Variables

MCAID_ID	From source data
BYEAR	Derived from DOB
DOB	From source data
FNAME1	Parsed from IFNAME
FNAME2	Parsed from IFNAME
FNAME3	Parsed from IFNAME
FNAME4	Parsed from IFNAME
GENDER	From source data
LNAME1	Parsed from ILNAME
LNAME2	Parsed from ILNAME
LNAME3	Parsed from ILNAME
LNAME4	Parsed from ILNAME
MAID1	Parsed from IMAID
MAID2	Parsed from IMAID
MAID3	Parsed from IMAID
MAID4	Parsed from IMAID
MCAID_ID	From source data
MINIT	From source data
NYFRST	Derived from FNAME1
NYLAST	Derived from LNAME1
RACE	From source data
SSN	From source data
ZIP	From source data

MDCDELIG – Medicaid Eligibility

- Sorted by Medicaid ID (MCAID_ID)
- Multiple observations per Medicaid ID

MDCDELIG Variables

MCAID_ID	From source data
COUNTY	Recode from ICOUNTY
ELIG_BEG	From source data
ELIG_CAT	From source data
ELIG_END	From source data
IMC_TYPE	From source data
MC_BEG	From source data
MC_END	From source data
MC_IND	Recode from source data
MC_PLAN	From source data
MC_TYPE	Recode from IMC_TYPE
PCP_ID	From source data
ZIP	From source data

Temporary Storage

Temporary storage is available for data sets that are needed for longer than a single job, but are not part of the output for a program can be stored for approximately one to two weeks (depending on space availability). The directories for temporary storage are:

- Delaware /saswrk/de/program100/
- Oklahoma /saswrk/ok/program100/
- Washington /saswrk/wa/program100/

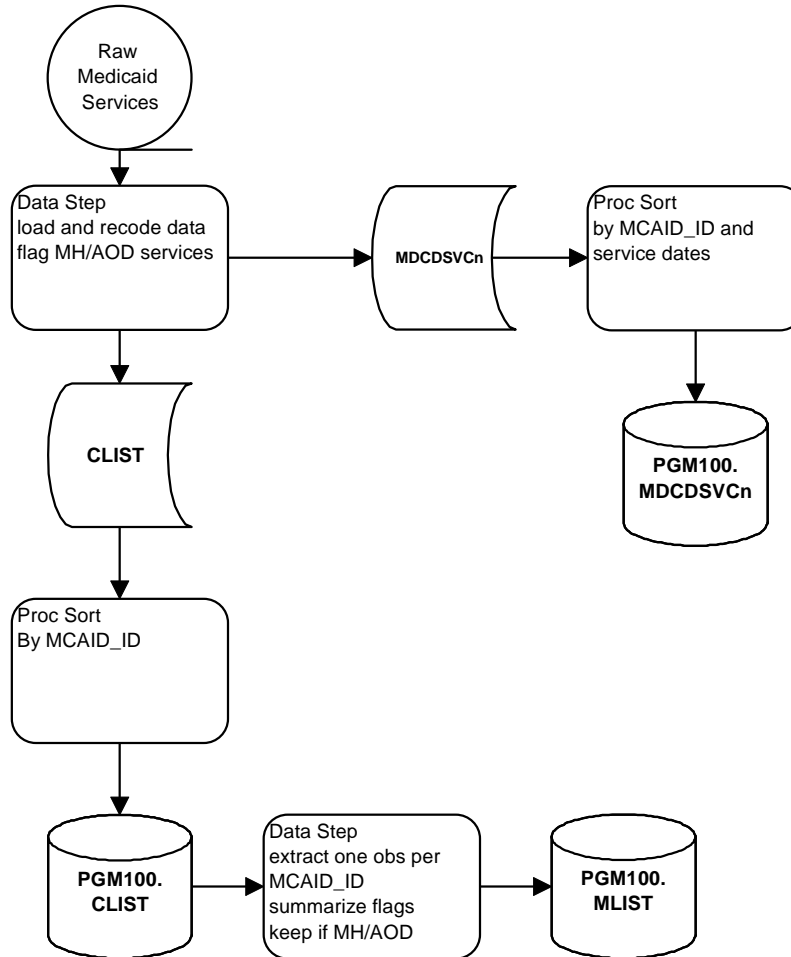
Processing Notes

- Programmers need to include all variables for each data set. Code has been written with appropriate Keep and Attribute statements, this should be included as the first statement of the data step creating the output data sets. The include file is #mcdclnt.sas

```
Usage is:
Filename INCLIB '/sas612/samhsa/@inclub/';
%Inc INCLIB(#mcdclnt.sas);
...
Data MDCD_ID (Keep=&MDCD_ID)
      MDCDELIG (Keep=&MDCDELIG);
      &MDCDCLNT;
...
Run;
```

- Proper coding of consistency codes are the responsibility of programmers. Consult the data dictionary and section on consistency codes for more information.
- Remove any dashes from SSN
- Use the SSNSCRUB macro to check for “special” missing values in the SSN variable (i.e. 9 filled or 101010101)
- If the client’s name is built into the person ID, compare that information to the name variables, using the ID to fill in any missing values.
- Use the NAMES macro to standardize both first and last names

Processing



Input Files

Input files should be found on the server:

- Delaware /samhsa/de/1996/rawdata/medicaid/
- Oklahoma /samhsa/ok/1996/rawdata/medicaid/
- Washington /samhsa/wa/1996/rawdata/medicaid/

Output Data Sets

File Location

All program 100 output should be placed in the appropriate server directory:

- Delaware /samhsa/de/1996/program100/

- Oklahoma /samhsa/ok/1996/program100/
- Washington /samhsa/wa/1996/program100/

Data Sets

MDCDSVC1 – Outpatient/Other Medicaid Claims

- Sorted by Medicaid ID (MCAID_ID) and service dates (SERV_BEG and SERV_END)
- Multiple observations per Medicaid ID
- Set Variable SERVFILE to 'OTH'
- Set Variable SERVSRC to 'MCAID'

MDCDSVC1 Variables	
MCAID_ID	From source data
ISERVCAT	From source data
SERV_BEG	From source data
SERV_END	From source data
SERVCAT	Recode from ISERVCAT
SERVFILE	Assigned value 'OTH'
SERVSRC	Assigned (see below)
ADJ_IND	Recode from IADJ_IND
AMT_CAID	From source data
AMT_CARE	From source data
AMT_CHGD	From source data
AMT_COIN	From source data
AMT_DED	From source data
AMT_PL	From source data
AMT_TPL	From source data
CLNT_ZIP	From source data
CSO	From source data
DX1	From source data
DX2	From source data
DX3	From source data
DX4	From source data
DX5	From source data
DX6	From source data
DX7	From source data
DX8	From source data
DX9	From source data
FFS_IND	From source data
IADJ_IND	From source data
IMCARE_X	From source data
IND_CAID	Coded
IND_CARE	Coded
IND_CHGD	Coded
IND_COIN	Coded
IND_DED	Coded
IND_PL	Coded
IND_TPL	Coded
IPROVSPC	From source data
IPROVTYP	From source data
MC_PLAN	From source data
MC_TYPE	Recode from source data
MCARE_X	Recode from IMCARE_X
PAY_DATE	From source data
PROC_CD	From source data
PROC_IND	From source data
PROC_MOD	From source data
PROV_ID	From source data
PROVCNTY	From source data
PROVSPC	Recode from IPROVSPC

IDB Users Guide

MDCDSVC1 Variables	
PROVTYP	Recode from IPROVTYP
REV_CD	From source data
SERV_QTY	From source data
SERV_UNT	From source data
SERVPLC	Recode from ISERVPLC
SERVPROV	From source data

MDCDSVC2 – Inpatient Medicaid Claims

- Sorted by Medicaid ID (MCAID_ID) and service dates (SERV_BEG and SERV_END)
- Multiple observations per Medicaid ID
- Set Variable SERVFILE to 'IN'
- Set Variable SERVSRC to 'MCAID'

MDCDSVC2 Variables	
MCAID_ID	From source data
ISERVCAT	From source data
SERV_BEG	From source data
SERV_END	From source data
SERVCAT	Recode from ISERVCAT
SERVFILE	Assigned value 'IN'
SERVSRC	Assigned (see below)
ADATE	From source data
ADJ_IND	Recode from IADJ_IND
AMT_CAID	From source data
AMT_CARE	From source data
AMT_CHGD	From source data
AMT_COIN	From source data
AMT_DED	From source data
AMT_PL	From source data
AMT_TPL	From source data
CLNT_ZIP	From source data
CSO	From source data
DAYS	From source data
DDATE	From source data
DRG	From source data
DSTATUS	Recode from source data
DX1	From source data
DX2	From source data
DX3	From source data
DX4	From source data
DX5	From source data
DX6	From source data
DX7	From source data
DX8	From source data
DX9	From source data
FFS_IND	Recode from source data
IADJ_IND	From source data
IMCARE_X	From source data
IND_CAID	Coded
IND_CARE	Coded
IND_CHGD	Coded
IND_COIN	Coded
IND_DED	Coded
IND_PL	Coded
IND_TPL	Coded
IPROVTYP	From source data
MC_PLAN	From source data
MC_TYPE	Recode from source data
MCARE_X	Recode from IMCARE_X
PAY_DATE	From source data

MDCDSVC2 Variables	
PR1	From source data
PR1_DT	From source data
PR1_IND	Coded from source data
PR1_MOD	From source data
PR2	From source data
PR2_DT	From source data
PR2_IND	Coded from source data
PR2_MOD	From source data
PR3	From source data
PR3_DT	From source data
PR3_IND	Coded from source data
PR3_MOD	From source data
PR4	From source data
PR4_DT	From source data
PR4_IND	Coded from source data
PR4_MOD	From source data
PR5	From source data
PR5_DT	From source data
PR5_IND	Coded from source data
PR5_MOD	From source data
PR6	From source data
PR6_DT	From source data
PR6_IND	Coded from source data
PR6_MOD	From source data
PR7	From source data
PR7_DT	From source data
PR7_IND	Coded from source data
PR7_MOD	From source data
PR8	From source data
PR8_DT	From source data
PR8_IND	Coded from source data
PR8_MOD	From source data
PR9	From source data
PR9_DT	From source data
PR9_IND	Coded from source data
PR9_MOD	From source data
PROV_ID	From source data
PROVCNTY	From source data
PROVTYPE	Recode from IPROVTYP

MDCDSVC3 – Long Term Care Medicaid Claims

- Sorted by Medicaid ID (MCAID_ID) and service dates (SERV_BEG and SERV_END)
- Multiple observations per Medicaid ID
- Set Variable SERVFILE to 'LTC'
- Set Variable SERVSRC to 'MCAID'

MDCDSVC3 Variables	
MCAID_ID	From source data
ISERVCAT	From source data
SERV_BEG	From source data
SERV_END	From source data
SERVCAT	Recode from ISERVCAT
SERVFILE	Assigned value 'LTC'
SERVSRC	Assigned (see below)
ADATE	From source data
ADJ_IND	Recode from IADJ_IND
AMT_CAID	From source data
AMT_CARE	From source data
AMT_CHGD	From source data
AMT_COIN	From source data

IDB Users Guide

MDCDSVC3 Variables	
AMT_DED	From source data
AMT_PL	From source data
AMT_TPL	From source data
CLNT_ZIP	From source data
COVDAYS	From source data
CSO	From source data
DSTATUS	Recode from source data
DX1	From source data
DX2	From source data
DX3	From source data
DX4	From source data
DX5	From source data
DX6	From source data
DX7	From source data
DX8	From source data
DX9	From source data
FFS_IND	From source data
IADJ_IND	From source data
IMCARE_X	From source data
IND_CAID	Coded
IND_CARE	Coded
IND_CHGD	Coded
IND_COIN	Coded
IND_DED	Coded
IND_PL	Coded
IND_TPL	Coded
IPROVTYP	From source data
LDAYS	From source data
LEV_CARE	From source data
MC_PLAN	From source data
MC_TYPE	Recode from source data
MCARE_X	Recode from IMCARE_X
PAY_DATE	From source data
PROV_ID	From source data
PROVCNTY	From source data
PROVTYPE	Recode from IPROVTYP
REVCODE	From source data

MDCDSVC4 – Pharmacy Medicaid Claims

- Sorted by Medicaid ID (MCAID_ID) and service dates (SERV_BEG and SERV_END)
- Multiple observations per Medicaid ID
- Set Variable SERVFILE to 'RX'
- Set Variable SERVSRC to 'MCAID'

MDCDSVC4 Variables	
MCAID_ID	From source data
ISERVCAT	From source data
SERV_BEG	From source data
SERV_END	From source data
SERVCAT	Recode from ISERVCAT
SERVFILE	Assigned value 'RX'
SERVSRC	Assigned (see below)
ADJ_IND	Recode from IADJ_IND
AMT_CAID	From source data
AMT_CHGD	From source data
AMT_PL	From source data
AMT_TPL	From source data
CLNT_ZIP	From source data
CSO	From source data
DRUG_CD	From source data

MDCDSVC4 Variables	
FFS_IND	From source data
IADJ_IND	From source data
IND_CAID	Coded
IND_CHGD	Coded
IND_PL	Coded
IND_TPL	Coded
IPROVTYP	From source data
MC_PLAN	From source data
MC_TYPE	Recode from source data
PAY_DATE	From source data
PROV_ID	From source data
PROVCNTY	From source data
PROVTYP	Recode from IPROVTYP
REFILL	From source data
RX_DATE	From source data
RX_DAYS	From source data
RX_FILL	From source data
RX_PHYS	From source data
RX_QTY	From source data
TH_CLASS	From source data

CLIST – Medicaid Services, Flagged

- Sorted by Medicaid ID (MCAID_ID)
- Multiple observations per Medicaid ID are possible, one observation for each service
- Contains client flags (see section on diagnosis-based flags for coding details)

CLIST Variables	
MCAID_ID	From source data
FAFFPSYC	Coded - See Appendix
FALCABSE	Coded - See Appendix
FALCPSYC	Coded - See Appendix
FALZHMR	Coded - See Appendix
FCHILDIS	Coded - See Appendix
FDRGABSE	Coded - See Appendix
FDRGPSYC	Coded - See Appendix
FHIWAIDS	Coded - See Appendix
FMAJDEP	Coded - See Appendix
FMENTDEV	Coded - See Appendix
FOTHAOD	Coded - See Appendix
FOTHMENT	Coded - See Appendix
FOTHMOOD	Coded - See Appendix
FOTHORG	Coded - See Appendix
FOTHPSYC	Coded - See Appendix
FPERSDIS	Coded - See Appendix
FPTNALC	Coded - See Appendix
FPREG	Coded - See Appendix
FPROC	Coded - See Appendix
FPRVTYPE	Coded - See Appendix
FREVCTR	Coded - See Appendix
FSCHIZO	Coded - See Appendix
FSTRESS	Coded - See Appendix
FTOBACCO	Coded - See Appendix
FTYPSERV	Coded - See Appendix
FVCODE	Coded - See Appendix

MLIST – Medicaid Recipients With MH/AOD Services

- Sorted by Medicaid ID (MCAID_ID)
- Single observation per Medicaid ID

IDB Users Guide

- Contains client flags at most relevant level

MLIST Variables	
MCAID_ID	From source data
FAFFPSYC	Coded - See Appendix
FALCABSE	Coded - See Appendix
FALCPSYC	Coded - See Appendix
FALZHMR	Coded - See Appendix
FCHILDIS	Coded - See Appendix
FDRGABSE	Coded - See Appendix
FDRGPSYC	Coded - See Appendix
FHIVAIDS	Coded - See Appendix
FMAJDEP	Coded - See Appendix
FMENTDEV	Coded - See Appendix
FOTHAOD	Coded - See Appendix
FOTHMENT	Coded - See Appendix
FOTHMOOD	Coded - See Appendix
FOTHORG	Coded - See Appendix
FOTHPSYC	Coded - See Appendix
FPERSDIS	Coded - See Appendix
FPOTNALC	Coded - See Appendix
FPREG	Coded - See Appendix
FPROC	Coded - See Appendix
FPRVTYPE	Coded - See Appendix
FREVCTR	Coded - See Appendix
FSCHIZO	Coded - See Appendix
FSTRESS	Coded - See Appendix
FTOBACCO	Coded - See Appendix
FTYPSERV	Coded - See Appendix
FVCODE	Coded - See Appendix

Processing Notes

- Programmers need to include all variables for each data set. Code has been written with appropriate Keep and Attribute statements, this should be included as the first statement of the data step creating the output data sets. The include file is #mcdclms.sas

```
Usage is:
%Let State_ = wa;  ** use one of de/ok/wa;
Filename INCLIB '/sas612/samhsa/@inclib/';
%Inc INCLIB(#mcdclms.sas);
...
Data MDCDSVC1 (Keep=&MDCDSVC1)
      MDCDSVC2 (Keep=&MDCDSVC2)
      MDCDSVC3 (Keep=&MDCDSVC3)
      MDCDSVC4 (Keep=&MDCDSVC4)
      CLIST    (Keep=&CLIST);
&MDCDCLMS;
...
Run;
```

- Proper coding of consistency codes are the responsibility of programmers. Consult the data dictionary and Appendix for more information.
- Recodes for SERVCAT, PROVTYPE, and PROVSPEC should follow the uniform mappings developed specifically for these variables.
- Use the DXSCRUB macro to standardize diagnosis codes
- The \$DSM format should be used to set flags

- ◆ Four additional flags based on \$DSM have been added since the prototype: FHIVAIDS, FPOTNALC, FVCODE, and FMENTDEV
- ◆ The name of one flag has been changed (from FPSYCHDS to FOTHORG)
- ◆ See the appendix
- The pregnancy complication flag (FPREG) is set from the diagnoses separately from the other flags, using the \$PREG format.

Temporary Storage

Temporary storage is available for data sets that are needed for longer than a single job, but are not part of the output for a program can be stored for approximately one to two weeks (depending on space availability). The directories for temporary storage are:

- Delaware /saswrk/de/program200/
- Oklahoma /saswrk/ok/program200/
- Washington/saswrk/wa/program200/

PROVIDER DATA

Provider files are not processed for 1996 data.

SAS LOADS AND RECODES, MH/AOD DATA – PROGRAM 200

Program Names

Name programs “step2nn” where nn are digits (i.e. step210, step215).

Location

Programs should be saved on the server in the appropriate directories:

- Delaware /sas612/samhsa/de/program200/
- Oklahoma /sas612/samhsa/ok/program200/
- Washington/sas612/samhsa/wa/program200/

Macros and File Includes

All macros and file includes used by these programs can be found on the following server directory:
/sas612/samhsa/@inclib/

SAS Formats

All custom SAS formats used by these programs can be found on the following server directory:
/sas612/samhsa/@fmtlib/

IDB Users Guide
Documentation

All programs should document the program name, programmer, and purpose of job at the top of the listing.

INPUT DATA

- Input files should be found on the server:
- Delaware /samhsa/de/1996/rawdata/mh_adult/
and /samhsa/de/1996/rawdata/mh_youth/
- Oklahoma /samhsa/ok/1996/rawdata/mhsa/
- Washington /samhsa/wa/1996/rawdata/dasa/
and /samhsa/wa/1996/rawdata/smha/

OUTPUT DATA

File Location

All program 200 output should be placed in the appropriate server directory:

- Delaware /samhsa/de/1996/program200/
- Oklahoma /samhsa/ok/1996/program200/
- Washington /samhsa/wa/1996/program200/

Prefixes for Data Set Name

All program 200 output should be named using the following prefixes:

- Delaware
 - ◆ Adult MH/AOD ALDT
 - ◆ Youth MH/AOD CHLD
- Oklahoma MHSA
- Washington
 - ◆ AOD DASA
 - ◆ MH SMHA

Substitute the appropriate prefix for <pre> in the data set names that follow

Client Data Sets

<pre>_ID – Client Identifying Data

- Sorted by person ID (STATE_ID)
- One observation per person ID

- When multiple client records are available for an individual, populate variables with the most recent non-missing data.
- Contains client flags at most relevant level (see Appendix for coding details)
- Contains variables found in the data sets CLIENT and ID_DATA.

<pre>_ID Variables	
STATE_ID	From source data
BYEAR	Derived from DOB
DOB	From source data
FNAME1	Parsed from IFNAME
FNAME2	Parsed from IFNAME
FNAME3	Parsed from IFNAME
FNAME4	Parsed from IFNAME
GENDER	From source data
LNAME1	Parsed from ILNAME
LNAME2	Parsed from ILNAME
LNAME3	Parsed from ILNAME
LNAME4	Parsed from ILNAME
MAID1	Parsed from IMAID
MAID2	Parsed from IMAID
MAID3	Parsed from IMAID
MAID4	Parsed from IMAID
MCAID_ID	From source data
MINIT	From source data
NYFRST	Derived from FNAME1
NYLAST	Derived from LNAME1
SED	From source data
SMI	From source data
RACE	From source data
SSN	From source data
ZIP	From source data
FAFFPSYC	Coded - See Appendix
FALCABSE	Coded - See Appendix
FALCPSYC	Coded - See Appendix
FALZHMR	Coded - See Appendix
FCHILDIS	Coded - See Appendix
FDRGABSE	Coded - See Appendix
FDRGPSYC	Coded - See Appendix
FHIVAIDS	Coded - See Appendix
FMAJDEP	Coded - See Appendix
FMENTDEV	Coded - See Appendix
FOTHAOD	Coded - See Appendix
FOTHMENT	Coded - See Appendix
FOTHMOOD	Coded - See Appendix
FOTHORG	Coded - See Appendix
FOTHPSYC	Coded - See Appendix
FPERSDIS	Coded - See Appendix
FPOTNALC	Coded - See Appendix
FPREG	Coded - See Appendix
FSCHIZO	Coded - See Appendix
FSTRESS	Coded - See Appendix
FTOBACCO	Coded - See Appendix
FVCODE	Coded - See Appendix

<pre>DET – Client Detail

- Sorted by person ID (STATE_ID) and update date (UP_DATE)
- Zero to multiple observations per person ID

Where available, this file should contain information for each input record.

IDB Users Guide

<pre>DET Variables	
STATE_ID	From source data
AOD_HIST	Recode from source data
COUNTY	Recode from source data
DAGE1	Recode from source data
DAGE2	Recode from source data
DAGE3	Recode from source data
DISCNT	Recode from source data
DRUG1	Recode from source data
DRUG2	Recode from source data
DRUG3	Recode from source data
DX_A1_1	From source data
DX_A1_2	From source data
DX_A1_3	From source data
DX_A2_1	From source data
DX_A2_2	From source data
DX_A2_3	From source data
DX_A3_1	From source data
DX_A3_2	From source data
DX_A3_3	From source data
DX_A4_A	From source data
DX_A4_B	From source data
DX_A4_C	From source data
DX_A4_D	From source data
DX_A4_E	From source data
DX_A4_F	From source data
DX_A4_G	From source data
DX_A4_H	From source data
DX_A4_I	From source data
DX_A5	From source data
E_AFDC	Recode from source data
E_MCAID	Recode from source data
EDUCAT	Recode from source data
EMPLOY	Recode from source data
FREQ1	Recode from source data
FREQ2	Recode from source data
FREQ3	Recode from source data
HLTHINS	Recode from source data
LEGAL	Recode from source data
LIVING	Recode from source data
LOF	From source data
MARITAL	Recode from source data
METHSERV	Recode from source data
MH_HLST	Recode from source data
PAY_SRC	Recode from source data
PPROB1	Recode from source data
PPROB2	Recode from source data
PPROB3	Recode from source data
PR_TREAT	Recode from source data
PREGADMT	Recode from source data
PRINAXIS	Recode from source data
PROGTYP	Recode from source data
REF_DIS	Recode from source data
REF_SRC	Recode from source data
ROUTE1	Recode from source data
ROUTE2	Recode from source data
ROUTE3	Recode from source data
TRANTYP	Recode from source data
UP_DATE	From source data
VETERAN	Recode from source data
ZIP	From source data

Processing Notes

- The Client Detail data set replaces the Program data set from the prototype data base. This data set contains information only from client files.

- Programmers need to include all variables for each data set. Code has been written with appropriate Keep and Attribute statements, this should be included as the first statement of the data step creating the output data sets. The include file is #mhsaclnt.sas

```
Usage is
Filename INCLIB '/sas612/samhsa/@inclub/';
%Inc INCLIB(#mhsaclnt.sas);
...
Data MHSA_ID (Keep=&MHSA_ID)
      MHSADET (Keep=&MHSADET);
      &MHSACLNT;
...
Run;
```

- Proper coding of consistency codes are the responsibility of programmers. Consult the data dictionary and Appendix for more information.
- Recodes for PROGTYP should follow the uniform mappings developed specifically for this variable.
- Remove any dashes from SSN
- Use the SSNSCRUB macro to check for “special” missing values in the SSN variable (i.e. 9 filled or 101010101)
- If the client’s name is built into the person ID, compare that information to the name variables, using the ID to fill in missing values.
- Use the NAMES macro to standardize both first and last names
- Use the DXSCRUB macro to standardize diagnosis codes (where available)
- The \$DSM format should be used to set flags
 - ◆ Three additional flags based on \$DSM have been added since the prototype: FHIVAIDS, FPOTNALC, FVCODE, and FMENTDEV
 - ◆ The name of one flag has been changed (from FPSYCHDS to FOTHORG)
 - ◆ See the appendix

Service Data Sets

<pre>COM – Community Mental Health Services

- Sorted by person ID (STATE_ID) and service dates (SERV_BEG and SERV_END)
- Multiple observations per person ID
- Set Variable SERVSRC as appropriate
 - ◆ Delaware

Adult AOD	HSSA
Adult MH	HSSM
Youth MH/AOD	CMHS
 - ◆ Oklahoma

	MH/AOD	MHSA
◆ Washington		
	AOD	DASA
	MH	SMHA

<pre>COM Variables

| | |
|----------|---|
| STATE_ID | From source data |
| CAP_AMT | From source data |
| COST | From source data –OR – derived from rates and quantity of service |
| COST_IND | Coded based on COST |
| DXCODE | From source data |
| ISERVICE | From source data |
| SERV_LOC | Recode from source data |
| SERV_QTY | From source data |
| SERV_UNT | Recode from source data |
| SERVPROV | From source data |
| ISERVCAT | From source data |
| SERV_BEG | From source data |
| SERV_END | From source data |
| SERVCAT | Recode from ISERVCAT |
| SERVFILE | Set to 'COM' |
| SERVSRC | Set as described above |

<pre>IMH – Institutional Mental Health Services

- Sorted by person ID (STATE_ID) and service dates (SERV_BEG and SERV_END)
- Multiple observations per person ID
- Set Variable SERVSRC as appropriate

| | | |
|--------------|--------------|------|
| ◆ Delaware | | |
| | Adult AOD | HSSA |
| | Adult MH | HSSM |
| | Youth MH/AOD | CMHS |
| ◆ Oklahoma | | |
| | MH/AOD | MHSA |
| ◆ Washington | | |
| | AOD | DASA |
| | MH | SMHA |

<pre>IMH Variables

| | |
|----------|--|
| STATE_ID | From source data |
| CAP_AMT | From source data |
| COST | From source data –OR – derived from rates and days of care |
| COST_IND | Coded based on COST |
| CDAYS | From source data |
| DXCODE | From source data |
| ISERVICE | From source data |
| SERV_LOC | Recode from source data |
| SERVPROV | From source data |
| STAYTYP | Recode from source data |


```

<pre>IMH Variables
ISERVCAT      From source data
SERV_BEG      From source data
SERV_END      From source data
SERVCAT       Recode from ISERVCAT
SERVFILE      Set to 'IMH'
SERVSRC       Set as described above

```

<pre>AOD – Alcohol/Other Drug Treatments

- Sorted by person ID (STATE_ID) and service dates (SERV_BEG and SERV_END)
- Multiple observations per person ID
- Set Variable SERVSRC as appropriate

- ◆ Delaware

| | |
|--------------|------|
| Adult AOD | HSSA |
| Adult MH | HSSM |
| Youth MH/AOD | CMHS |

- ◆ Oklahoma

| | |
|--------|------|
| MH/AOD | MHSA |
|--------|------|

- ◆ Washington

| | |
|-----|------|
| AOD | DASA |
| MH | SMHA |

```

<pre>AOD Variables
STATE_ID      From source data
CAP_AMT       From source data
COST          From source data –OR– derived from rates and quantity of service
COST_IND      Coded based on COST
DXCODE        From source data
ISERVICE      From source data
SERV_LOC      Recode from source data
SERV_QTY      From source data
SERV_UNT      Recode from source data
SERVPROV      From source data
ISERVCAT      From source data
SERV_BEG      From source data
SERV_END      From source data
SERVCAT       Recode from ISERVCAT
SERVFILE      Set to 'AOD'
SERVSRC       Set as described above

```

<pre>FLGS –Treatment/Service Diagnosis Flags

- Sorted by person ID (STATE_ID)
- Single observation per person ID
- Create only if there treatment level diagnosis codes are present

IDB Users Guide

| <pre>FLGS Variables | |
|---------------------|----------------------|
| STATE_ID | From source data |
| FAPFPSYC | Coded - See Appendix |
| FALCABSE | Coded - See Appendix |
| FALCPSYC | Coded - See Appendix |
| FALZHMR | Coded - See Appendix |
| FCHILDIS | Coded - See Appendix |
| FDRGABSE | Coded - See Appendix |
| FDRGPSYC | Coded - See Appendix |
| FHIVADS | Coded - See Appendix |
| FMAJDEP | Coded - See Appendix |
| FMENTDEV | Coded - See Appendix |
| FOTHAOD | Coded - See Appendix |
| FOTHMENT | Coded - See Appendix |
| FOTHMOOD | Coded - See Appendix |
| FOTHORG | Coded - See Appendix |
| FOTHPSYC | Coded - See Appendix |
| FPERSDIS | Coded - See Appendix |
| FPOTNALC | Coded - See Appendix |
| FPREG | Coded - See Appendix |
| FSCHIZO | Coded - See Appendix |
| FSTRESS | Coded - See Appendix |
| FTOBACCO | Coded - See Appendix |
| FVCODE | Coded - See Appendix |

Processing Notes

- CAP_AMT is intended to record capitation amounts received by the MH/AOD agency – it does not record any specific service or treatment. Capitation should be coded with SERVCAT set to '24' (Capitated premium), RATE set to zero (0), and RATE_IND set to '1' (Monthly bundled capitation rate). Specific services for capitated clients, should be coded with CAP_AMT set to . A (Service for capitated client), while services for non-capitated clients should be coded with CAP_AMT set to . B (Service for non-capitated client).
- The COST variable is intended to reflect the cost of the service or treatment to the MH/AOD agency. When available, this amount should be recorded from source data. If the source data does not record costs, they should be imputed from quantity of service (or days of care) and rate estimates supplied by the agency.
- Treatment level diagnoses should be loaded into the variable DXCODE rather than moved to the client level data sets. If more than one diagnosis is available for a treatment, load the primary or principal diagnosis.
- Proper coding of consistency codes are the responsibility of programmers. Consult the data dictionary and Appendix for more information.
- Recodes for SERVCAT should follow the uniform mappings developed specifically for this variable.
- Programmers need to include all variables for each data set. Code has been written with appropriate Keep and Attribute statements, this should be included as the first statement of the data step creating the output data sets. The include file is #mhsaclms.sas

```
Usage is:
Filename INCLIB '/sas612/samhsa/@inclib/';
%Inc INCLIB(#mhsaclms.sas);
...
Data MHSACOM (Keep=&MHSACOM)
      MHSAIMH (Keep=&MHSAIMH)
```

```

MHSAAOD (Keep=&MHSAAOD)
MHS AFLGS (Keep=&MHS AFLGS) ;
&MHS ACLMS ;
...
Run ;

```

- Use the DXSCRUB macro to standardize diagnosis codes
- The \$DSM format should be used to set flags
 - ◆ Three additional flags based on \$DSM have been added since the prototype: FHIVAIDS, FPOTNALC, FV CODE, and FMENTDEV
 - ◆ The name of one flag has been changed (from FPSYCHDS to FOTHORG)
 - ◆ See the appendix
- The pregnancy complication flag (FPREG) is set from the diagnoses separately from the other flags, using the \$PREG format.

TEMPORARY STORAGE

Temporary storage is available for data sets that are needed for longer than a single job, but are not part of the output for a program can be stored for approximately one to two weeks (depending on space availability). The directories for temporary storage are:

- Delaware /saswrk/de/program200/
- Oklahoma /saswrk/ok/program200/
- Washington /saswrk/wa/program200/

PROCESSING INFORMATION

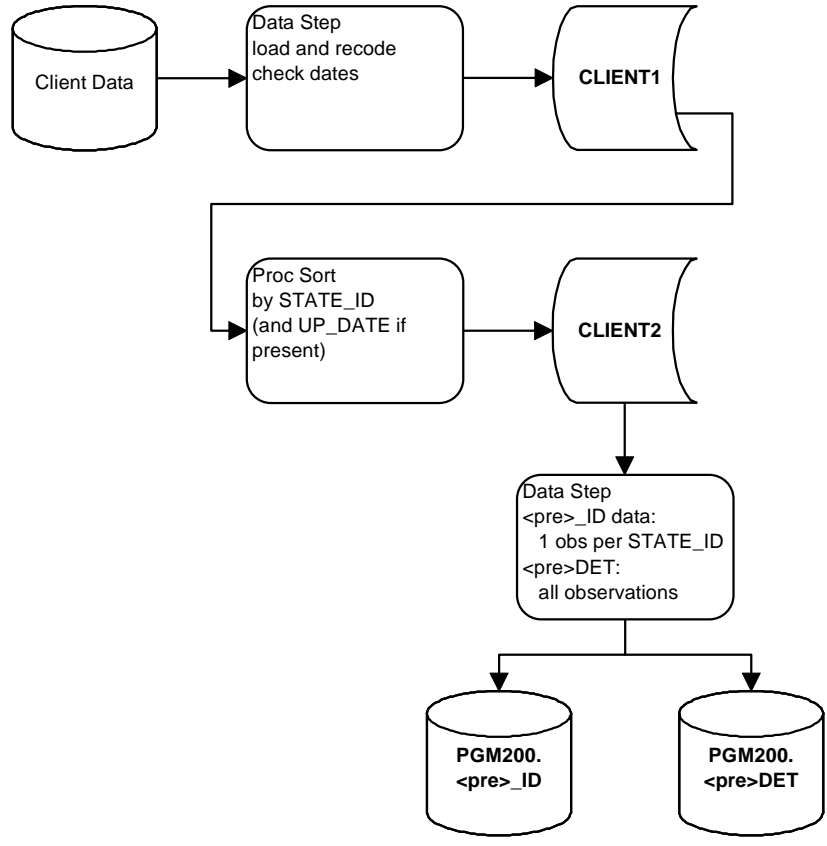
General Principles

- Client and service information are separate
- Client data, ultimately found in the Client Core and Client Detail data sets only describe a person.
- Service data, ultimately found in the Service Core and assorted service detail data sets describe services only.
- No variables, or other information, will be transferred from service data to client data. Nor from client data to service data.

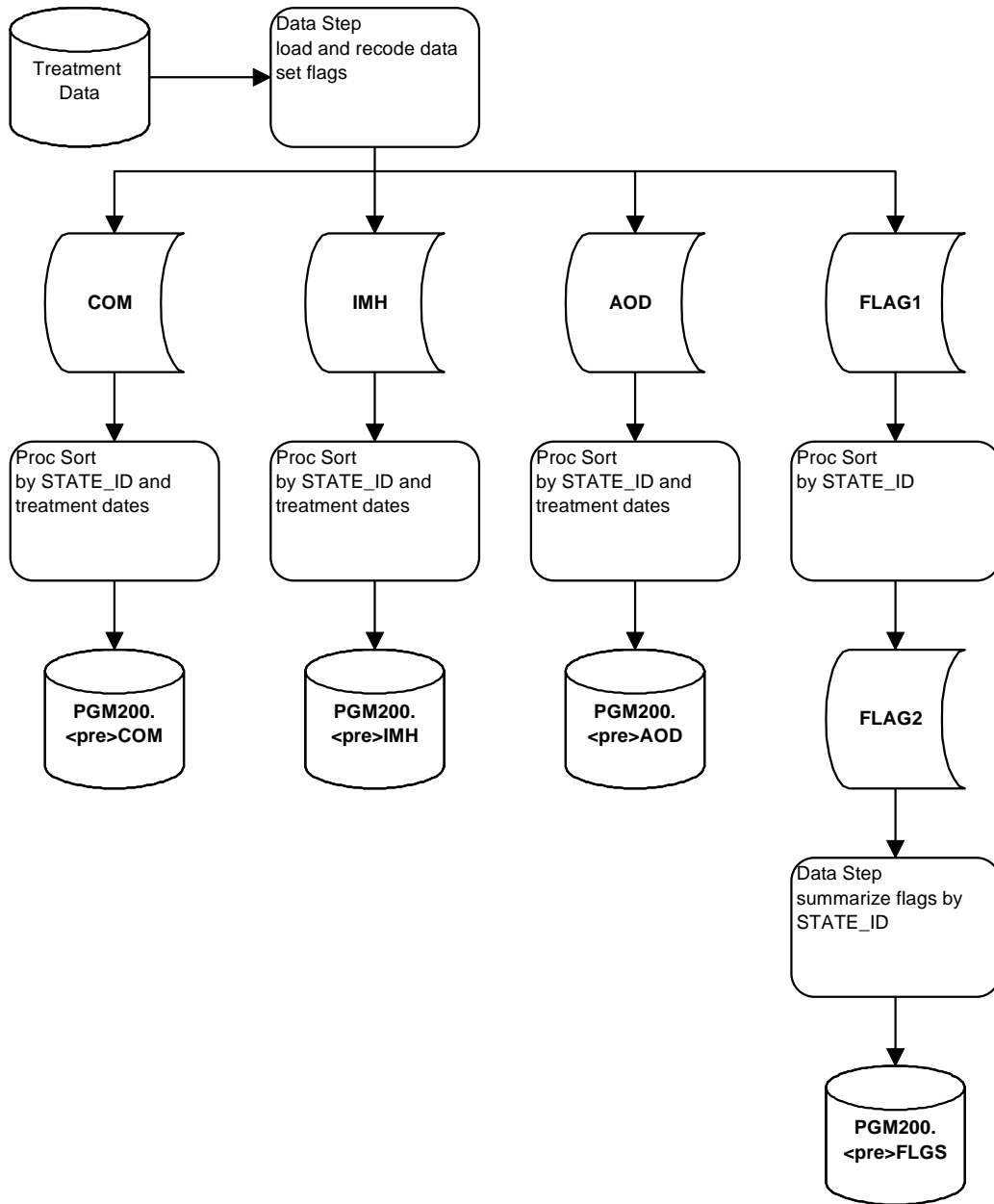
When Source data separates client information from service information

- Client data used to populate the ID and Detail data sets
- Service data used to populate the service data sets

IDB Users Guide
Client Data

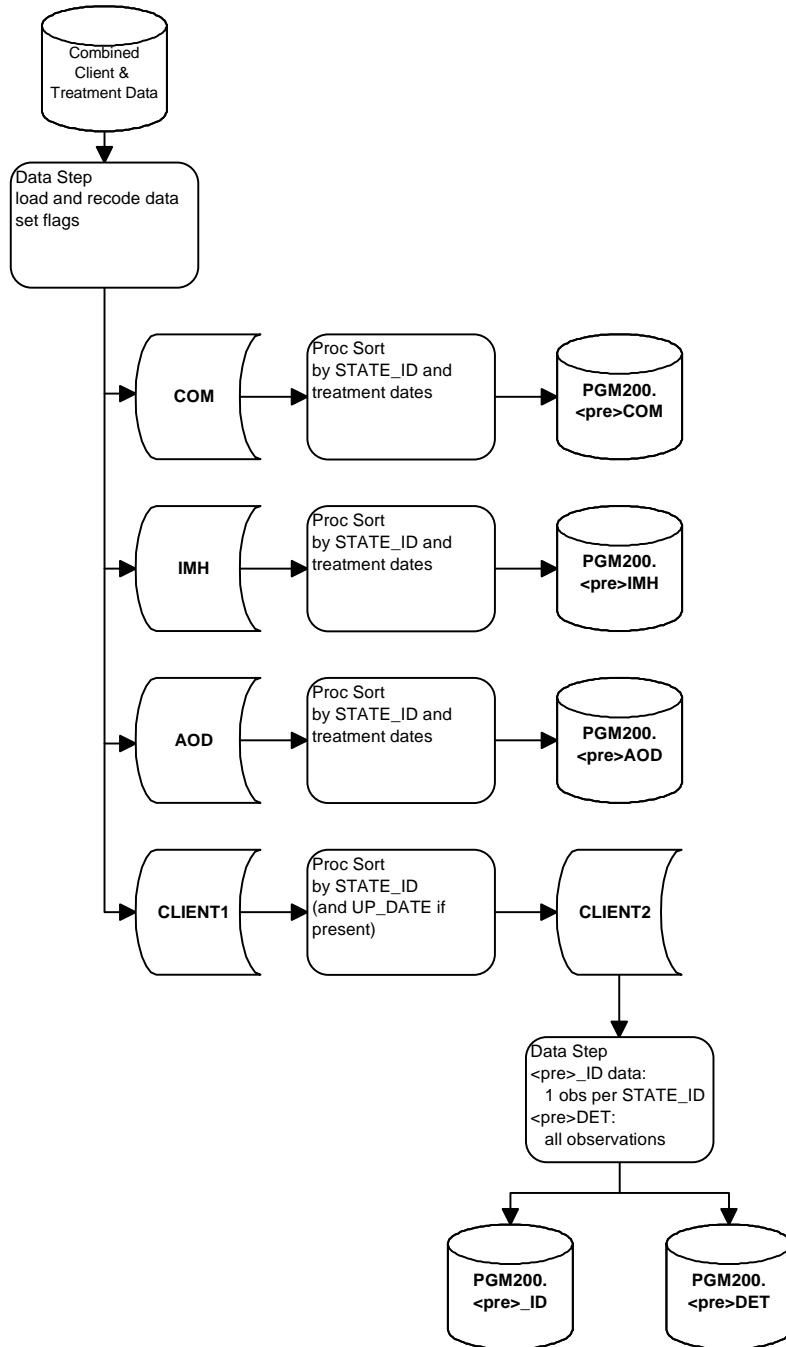


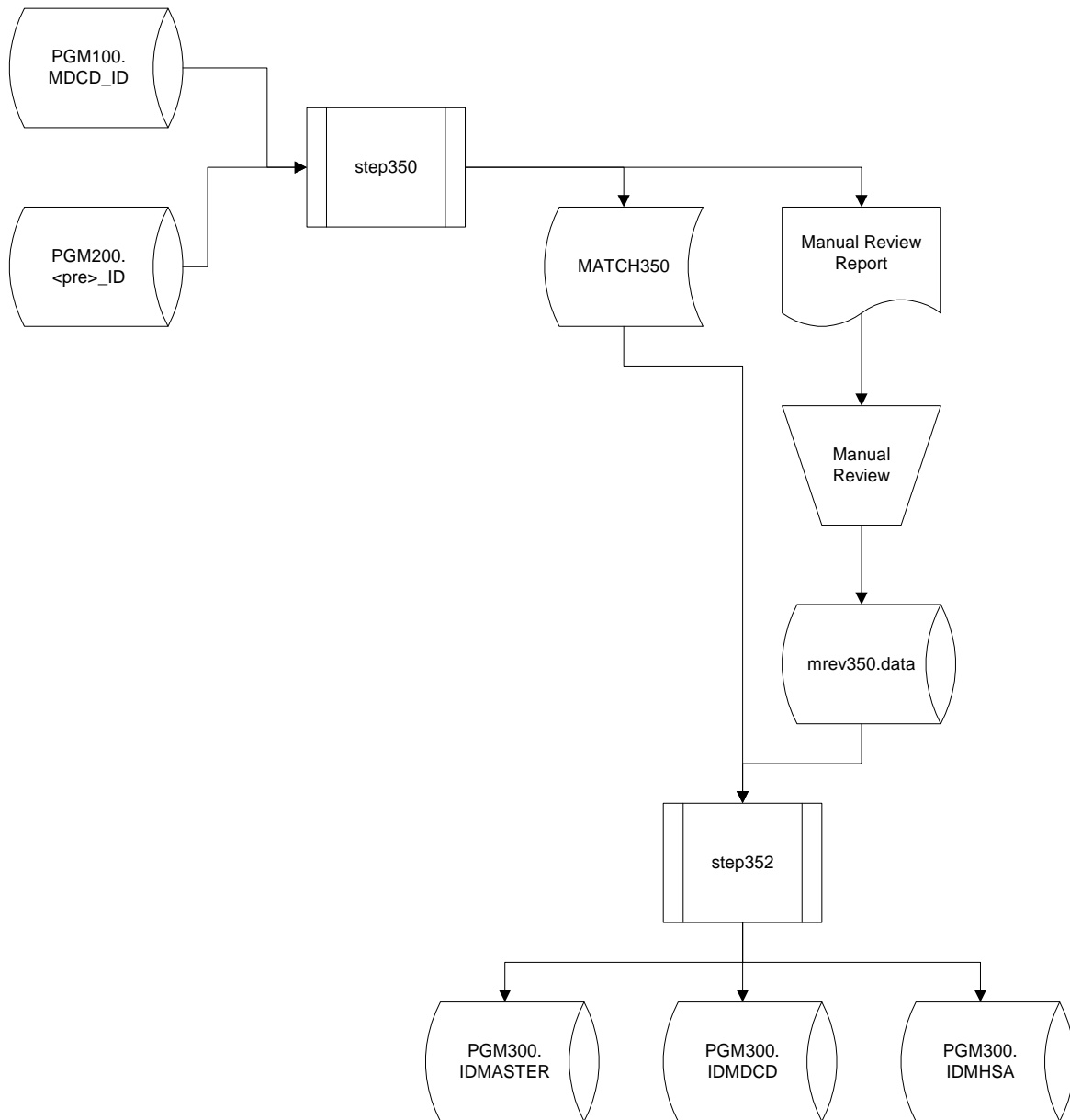
Treatment Data



Source Combines Client and Service Data

- Group data by person ID and provider
- Create detail observations using the first service for each person-provider



CLIENT LINKING – PROGRAM 300**PROGRAMS****Program Names**

Name programs “step3nn” where nn are digits (i.e. step350, step352).

Location

Programs should be saved on the server in the appropriate directories:

- Delaware /sas612/samhsa/de/program300/

IDB Users Guide

- Oklahoma /sas612/samhsa/ok/program300/
- Washington /sas612/samhsa/wa/program300/

Macros and File Includes

All macros and file includes used by these programs can be found on the following server directory:
/sas612/samhsa/@includlib/

SAS Formats

All custom SAS formats used by these programs can be found on the following server directory:
/sas612/samhsa/@fmtlib/

Documentation

All programs should document the program name, programmer, and purpose of job at the top of the listing.

INPUT DATA

Client linking uses the ID data sets from programs 100 (MDCD_ID) and 200 (<pre>_ID) found on the server:

- Delaware /samhsa/de/1996/program100/
/samhsa/de/1996/program200/
- Oklahoma /samhsa/ok/1996/program100/
/samhsa/ok/1996/program200/
- Washington /samhsa/wa/1996/program100/
/samhsa/wa/1996/program200/

OUTPUT DATA

Location

All program 300 output should be placed in the appropriate server directory:

- Delaware /samhsa/de/1996/program300/
- Oklahoma /samhsa/ok/1996/program300/
- Washington /samhsa/wa/1996/program300/

Data Sets

IDMASTER – Combined Identifying information

- Sorted by CLIENTID

- One observation per client

| IDMASTER Variables | |
|--------------------|----------------------------|
| CLIENTID | Assigned during processing |
| BYEAR | From program 100/200 data |
| DOB | From program 100/200 data |
| FNAME1 | From program 100/200 data |
| FNAME2 | From program 100/200 data |
| FNAME3 | From program 100/200 data |
| FNAME4 | From program 100/200 data |
| GENDER | From program 100/200 data |
| LNAME1 | From program 100/200 data |
| LNAME2 | From program 100/200 data |
| LNAME3 | From program 100/200 data |
| LNAME4 | From program 100/200 data |
| MAID1 | From program 100/200 data |
| MAID2 | From program 100/200 data |
| MAID3 | From program 100/200 data |
| MAID4 | From program 100/200 data |
| MINIT | From program 100/200 data |
| NYFRST | From program 100/200 data |
| NYLAST | From program 100/200 data |
| RACE | From program 100/200 data |
| SSN | From program 100/200 data |
| ZIP | From program 100/200 data |

IDMDCD – Complete Mapping from CLIENTID to MCAID_ID

- Sorted by CLIENTID
- Multiple observations per client

| IDMDCD Variables | |
|------------------|----------------------------|
| CLIENTID | Assigned during processing |
| MCAID_ID | From program 100 data |

IDMHSA – Complete Mapping from CLIENTID to STATE_ID

- Sorted by CLIENTID
- Multiple observations per client

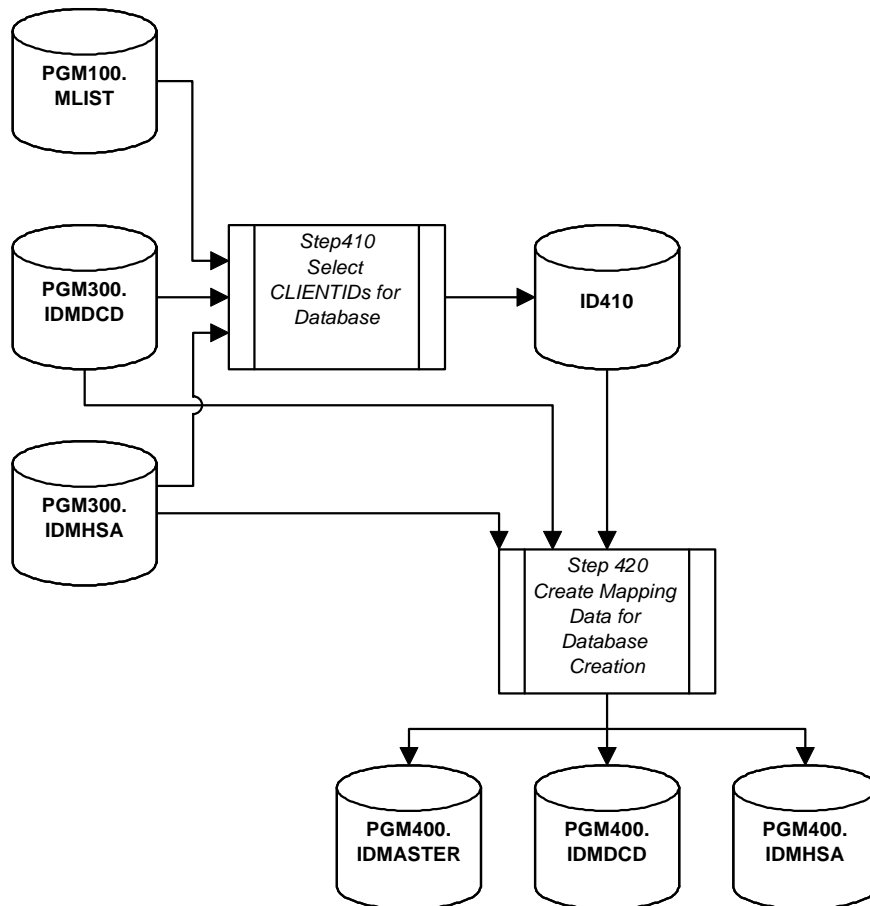
| IDMHSA Variables | |
|------------------|----------------------------|
| CLIENTID | Assigned during processing |
| STATE_ID | From program 200 data |

TEMPORARY STORAGE

Temporary storage is available for data sets that are needed for longer than a single job, but are not part of the output for a program can be stored for approximately one to two weeks (depending on space availability). The directories for temporary storage are:

- Delaware /saswrk/de/program300/
- Oklahoma /saswrk/ok/program300/
- Washington /saswrk/wa/program300/

ID MAPPING – PROGRAM 400



PROGRAMS

Program Names

Name programs “step4nn” where nn are digits (i.e. step410, step415).

Location

Programs should be saved on the server in the appropriate directories:

- Delaware /sas612/samhsa/de/program400/
- Oklahoma /sas612/samhsa/ok/program400/
- Washington /sas612/samhsa/wa/program400/

Macros and File Includes

All macros and file includes used by these programs can be found on the following server directory:

/sas612/samhsa/@includ/

SAS Formats

All custom SAS formats used by these programs can be found on the following server directory:
/sas612/samhsa/@fmtlib/

Documentation

All programs should document the program name, programmer, and purpose of job at the top of the listing.

INPUT DATA

ID mapping uses the output from program 300 and the MLIST data from program 100 found on the server:

- Delaware /samhsa/de/1996/program100/
/samhsa/de/1996/program300/
- Oklahoma /samhsa/ok/1996/program100/
/samhsa/ok/1996/program300/
- Washington /samhsa/wa/1996/program100/
/samhsa/wa/1996/program300/

OUTPUT DATA

Location

All program 400 output should be placed in the appropriate server directory:

- Delaware /samhsa/de/1996/program400/
- Oklahoma /samhsa/ok/1996/program400/
- Washington /samhsa/wa/1996/program400/

IDMASTER – Identifying information for Data Base Population

- Sorted by CLIENTID
- One observation per client

IDMASTER Variables

| | |
|----------|-----------------------|
| CLIENTID | From program 300 data |
| BYEAR | From program 300 data |
| DOB | From program 300 data |
| FNAME1 | From program 300 data |
| FNAME2 | From program 300 data |
| FNAME3 | From program 300 data |
| FNAME4 | From program 300 data |
| GENDER | From program 300 data |
| LNAME1 | From program 300 data |
| LNAME2 | From program 300 data |
| LNAME3 | From program 300 data |
| LNAME4 | From program 300 data |
| MAID1 | From program 300 data |
| MAID2 | From program 300 data |
| MAID3 | From program 300 data |
| MAID4 | From program 300 data |
| MINIT | From program 300 data |
| NYFRST | From program 300 data |
| NYLAST | From program 300 data |
| RACE | From program 300 data |
| SSN | From program 300 data |
| ZIP | From program 300 data |

IDMDCD – Mapping from CLIENTID to MCAID_ID for Data Base Population

- Sorted by CLIENTID
- Multiple observations per client

IDMDCD Variables

| | |
|----------|-----------------------|
| CLIENTID | From program 300 data |
| MCAID_ID | From program 300 data |

IDMHSA – Mapping from CLIENTID to STATE_ID for Data Base Population

- Sorted by CLIENTID
- Multiple observations per client

IDMHSA Variables

| | |
|----------|-----------------------|
| CLIENTID | From program 300 data |
| STATE_ID | From program 300 data |

TEMPORARY STORAGE

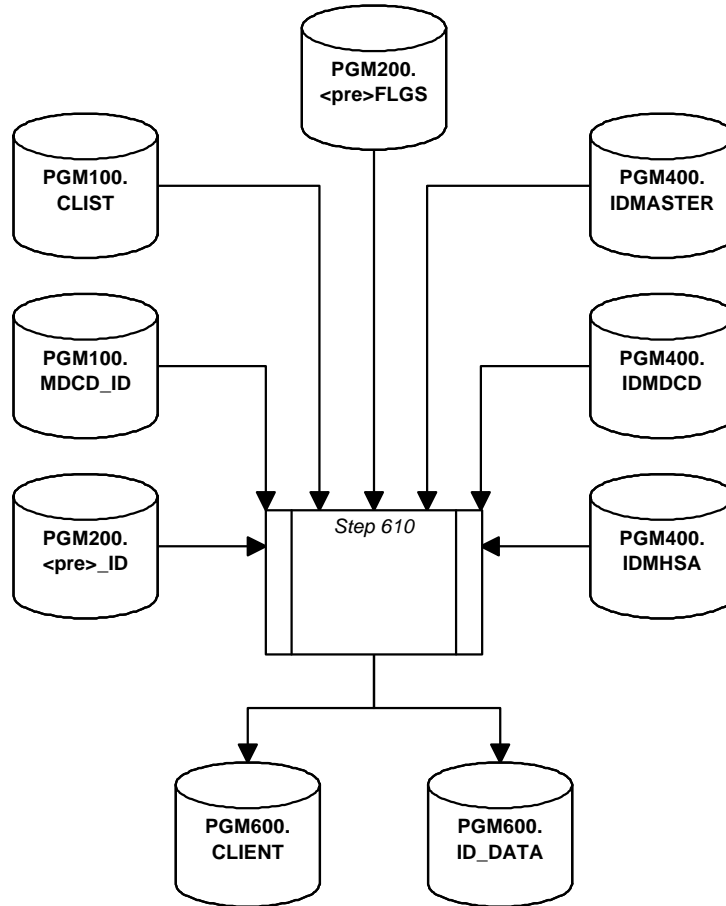
Temporary storage is available for data sets that are needed for longer than a single job, but are not part of the output for a program can be stored for approximately one to two weeks (depending on space availability). The directories for temporary storage are:

- Delaware /saswrk/de/program400/
- Oklahoma /saswrk/ok/program400/

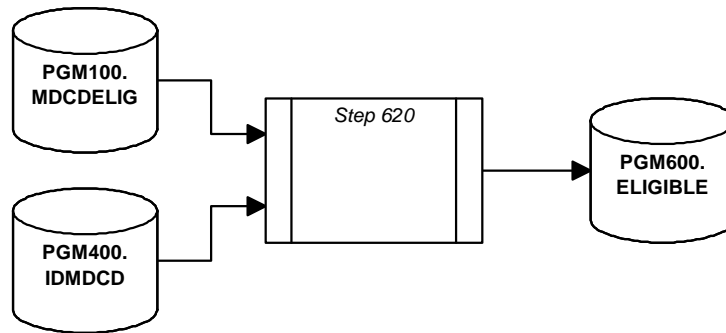
- Washington/saswrk/wa/program400/

DATA BASE BUILD – PROGRAM 600

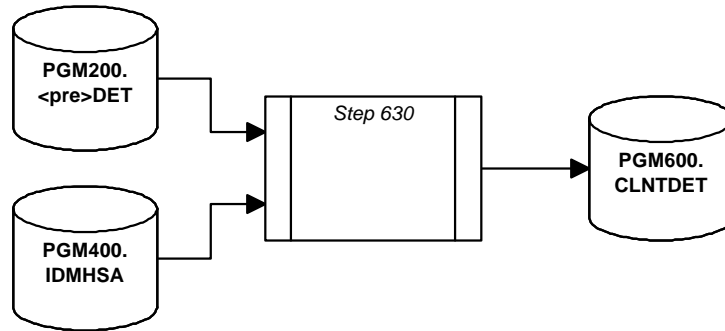
Client Core and ID Data Sets



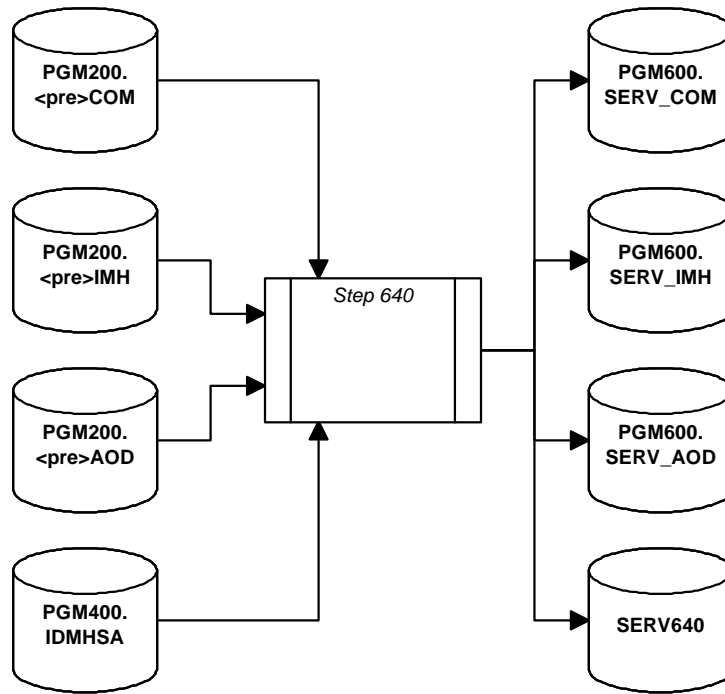
Medicaid Eligibility Data Set



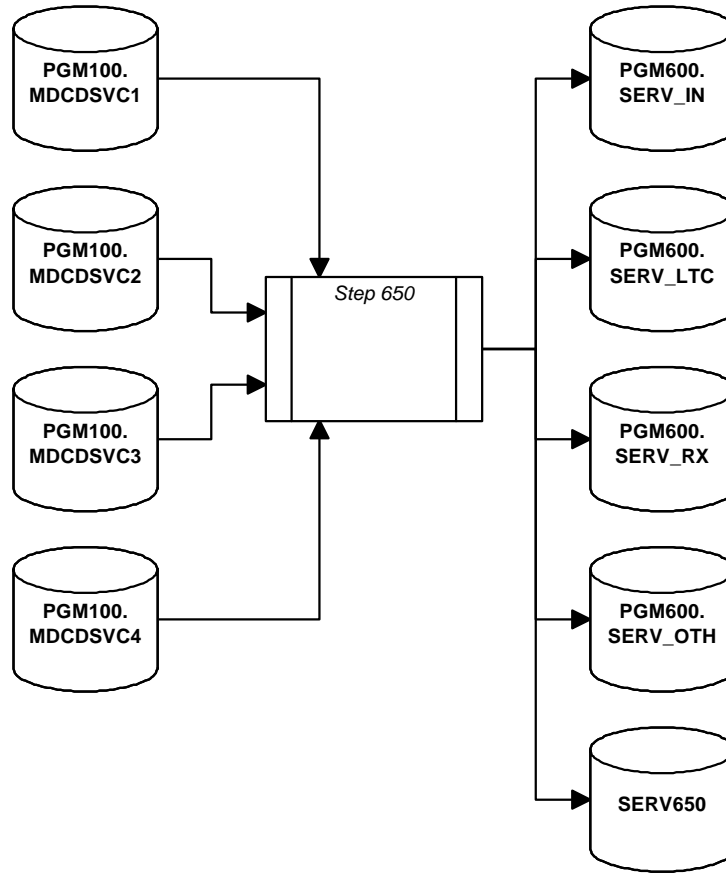
MH/AOD Client Detail Data Set



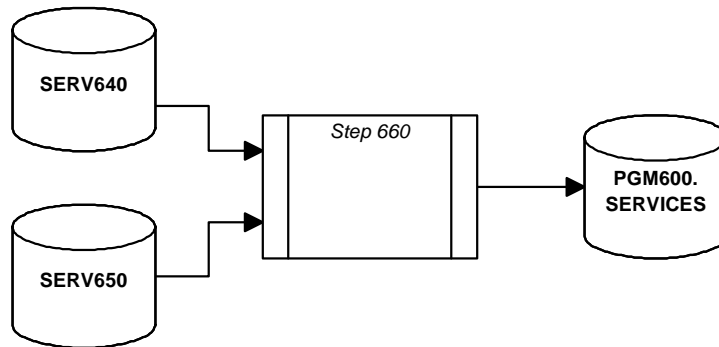
MH/AOD Service/Treatment Data Sets



Medicaid Service Detail Data Sets



Service Core Data Set



PROGRAMS

Program Names

Name programs “step6nn” where nn are digits (i.e. step610, step615).

Location

Programs should be saved on the server in the appropriate directories:

- Delaware /sas612/samhsa/de/program600/
- Oklahoma /sas612/samhsa/ok/program600/
- Washington /sas612/samhsa/wa/program600/

Macros and File Includes

All macros and file includes used by these programs can be found on the following server directory:
/sas612/samhsa/@includlib/

SAS Formats

All custom SAS formats used by these programs can be found on the following server directory:
/sas612/samhsa/@fmtlib/

Documentation

All programs should document the program name, programmer, and purpose of job at the top of the listing.

INPUT DATA

The data base build uses the output from program 400 and the data sets created in programs 100 and 200 found on the server:

- Delaware /samhsa/de/1996/program100/
/samhsa/de/1996/program200/
/samhsa/de/1996/program300/
- Oklahoma /samhsa/ok/1996/program100/
/samhsa/ok/1996/program200/
/samhsa/ok/1996/program300/
- Washington /samhsa/wa/1996/program100/
/samhsa/wa/1996/program200/
/samhsa/ws/1996/program300/

OUTPUT DATA

Location

All program 600 output should be placed in the appropriate server directory:

- Delaware /samhsa/de/1996/program600/
- Oklahoma /samhsa/ok/1996/program600/

- Washington /samhsa/wa/1996/program600/

Data Sets

CLIENT – Client Core

- Sorted by CLIENTID
- One observation per client

Client Core Variables – CLIENTS

| | |
|----------|------------------------------|
| CLIENTID | From program 400 data |
| C_CMHS | Deriverd from source of data |
| C_DASA | Deriverd from source of data |
| C_HSSA | Deriverd from source of data |
| C_MCAID | Deriverd from source of data |
| C_MHAA | Deriverd from source of data |
| C_MHSA | Deriverd from source of data |
| C_SMHA | Deriverd from source of data |
| DOB_S | Created from DOB |
| FAFFPSYC | From program 100/200 data |
| FALCABSE | From program 100/200 data |
| FALCPSYC | From program 100/200 data |
| FALZHMR | From program 100/200 data |
| FCHILDIS | From program 100/200 data |
| FDRGABSE | From program 100/200 data |
| FDRGPSYC | From program 100/200 data |
| FHIWAIDS | From program 100/200 data |
| FMAJDEP | From program 100/200 data |
| FMENTDEV | From program 100/200 data |
| FOTHAOD | From program 100/200 data |
| FOTHMENT | From program 100/200 data |
| FOTHMOOD | From program 100/200 data |
| FOTHORG | From program 100/200 data |
| FOTHPSYC | From program 100/200 data |
| FPERSDIS | From program 100/200 data |
| FPOTNALC | From program 100/200 data |
| FPREG | From program 100/200 data |
| FPROC | From program 100 data |
| FPRVTYPE | From program 100 data |
| FREVCTR | From program 100 data |
| FSCHIZO | From program 100/200 data |
| FSTRESS | From program 100/200 data |
| FTOBACCO | From program 100/200 data |
| FTYPSERV | From program 100 data |
| FVCODE | From program 100/200 data |
| GENDER | From program 400 data |
| RACE | From program 400 data |
| SED | From program 200 data |
| SMI | From program 200 data |

ID_Data – Identifying Information

- Sorted by CLIENTID
- One observation per client

Client Identifying Data Variables – ID_DATA

| | |
|----------|-----------------------|
| CLIENTID | From program 400 data |
| BYEAR | From program 400 data |
| DOB | From program 400 data |
| FNAME1 | From program 400 data |
| FNAME2 | From program 400 data |
| FNAME3 | From program 400 data |
| FNAME4 | From program 400 data |
| LNAME1 | From program 400 data |
| LNAME2 | From program 400 data |
| LNAME3 | From program 400 data |
| LNAME4 | From program 400 data |
| MAID1 | From program 400 data |
| MAID2 | From program 400 data |
| MAID3 | From program 400 data |
| MAID4 | From program 400 data |
| MINIT | From program 400 data |
| NYFRST | From program 400 data |
| NYLAST | From program 400 data |
| SSN | From program 400 data |
| ZIP | From program 400 data |

ELIGIBLE – Medicaid Eligibility

- Sorted by CLIENTID
- Multiple observation per client

Medicaid Eligibility Variables – ELIGIBLE

| | |
|----------|-----------------------|
| CLIENTID | From program 400 data |
| COUNTY | From program 100 data |
| ELIG_BEG | From program 100 data |
| ELIG_CAT | From program 100 data |
| ELIG_END | From program 100 data |
| MC_BEG | From program 100 data |
| MC_END | From program 100 data |
| MC_IND | From program 100 data |
| MC_PLAN | From program 100 data |
| MC_TYPE | From program 100 data |
| PCP_ID | From program 100 data |
| ZIP | From program 100 data |

CLNTDET – Client Detail (MH/AOD)

- Sorted by CLIENTID
- Multiple observation per client

MH/AOD Client Detail Variables – CLNTDET

| | |
|----------|-----------------------|
| CLIENTID | From program 400 data |
| AOD_HIST | From program 200 data |
| COUNTY | From program 200 data |
| DAGE1 | From program 200 data |
| DAGE2 | From program 200 data |
| DAGE3 | From program 200 data |
| DISCNT | From program 200 data |
| DRUG1 | From program 200 data |
| DRUG2 | From program 200 data |
| DRUG3 | From program 200 data |
| DX_A1_1 | From program 200 data |
| DX_A1_2 | From program 200 data |
| DX_A1_3 | From program 200 data |
| DX_A2_1 | From program 200 data |

MH/AOD Client Detail Variables – CLNTDET

| | |
|----------|-----------------------|
| DX_A2_2 | From program 200 data |
| DX_A2_3 | From program 200 data |
| DX_A3_1 | From program 200 data |
| DX_A3_2 | From program 200 data |
| DX_A3_3 | From program 200 data |
| DX_A4_A | From program 200 data |
| DX_A4_B | From program 200 data |
| DX_A4_C | From program 200 data |
| DX_A4_D | From program 200 data |
| DX_A4_E | From program 200 data |
| DX_A4_F | From program 200 data |
| DX_A4_G | From program 200 data |
| DX_A4_H | From program 200 data |
| DX_A4_I | From program 200 data |
| DX_A5 | From program 200 data |
| E_AFDC | From program 200 data |
| E_MCAID | From program 200 data |
| EDUCAT | From program 200 data |
| EMPLOY | From program 200 data |
| FREQ1 | From program 200 data |
| FREQ2 | From program 200 data |
| FREQ3 | From program 200 data |
| HLTHINS | From program 200 data |
| LEGAL | From program 200 data |
| LIVING | From program 200 data |
| LOF | From program 200 data |
| MARITAL | From program 200 data |
| METHSERV | From program 200 data |
| MH_HIST | From program 200 data |
| PAY_SRC | From program 200 data |
| PPROB1 | From program 200 data |
| PPROB2 | From program 200 data |
| PPROB3 | From program 200 data |
| PR_TREAT | From program 200 data |
| PREGADMT | From program 200 data |
| PRINAXIS | From program 200 data |
| PROGTYP | From program 200 data |
| REF_DIS | From program 200 data |
| REF_SRC | From program 200 data |
| ROUTE1 | From program 200 data |
| ROUTE2 | From program 200 data |
| ROUTE3 | From program 200 data |
| TRANTYP | From program 200 data |
| UP_DATE | From program 200 data |
| VETERAN | From program 200 data |
| ZIP | From program 200 data |

SERVICE – Service Core

- Sorted by CLIENTID and SERV_ID
- Multiple observation per client

Service Core Variables – SERVICES

| | |
|----------|-----------------------------------|
| CLIENTID | From program 400 data |
| AOD_CLM | Derived from program 100/200 data |
| DETOX | Derived from program 100/200 data |
| ISERVCAT | From program 100/200 data |
| MH_CLM | Derived from program 100/200 data |
| MH_OTH | Derived from program 100/200 data |
| SERV_BEG | From program 100/200 data |
| SERV_END | From program 100/200 data |
| SERV_ID | Assigned during processing |
| SERVCAT | From program 100/200 data |
| SERVFILE | From program 100/200 data |
| SERVSRC | From program 100/200 data |

SERV_IN – Medicaid Inpatient Claims

- Sorted by CLIENTID and SERV_ID
- Multiple observation per client

Medicaid Inpatient Service Detail Variables – SERV_IN

| | |
|----------|---|
| CLIENTID | From program 400 data |
| ADATE | From program 100 data |
| ADJ_IND | From program 100 data |
| AMT_CAID | From program 100 data |
| AMT_CARE | From program 100 data |
| AMT_CHGD | From program 100 data |
| AMT_COIN | From program 100 data |
| AMT_DED | From program 100 data |
| AMT_PL | From program 100 data |
| AMT_TPL | From program 100 data |
| CLNT_ZIP | From program 100 data |
| CSO | From program 100 data (Washington only) |
| DAYS | From program 100 data |
| DDATE | From program 100 data |
| DRG | From program 100 data |
| DSTATUS | From program 100 data |
| DX1 | From program 100 data |
| DX2 | From program 100 data |
| DX3 | From program 100 data |
| DX4 | From program 100 data |
| DX5 | From program 100 data |
| DX6 | From program 100 data |
| DX7 | From program 100 data |
| DX8 | From program 100 data |
| DX9 | From program 100 data |
| FFS_IND | From program 100 data |
| IADJ_IND | From program 100 data |
| IMCARE_X | From program 100 data |
| IND_CAID | From program 100 data |
| IND_CARE | From program 100 data |
| IND_CHGD | From program 100 data |
| IND_COIN | From program 100 data |
| IND_DED | From program 100 data |
| IND_PL | From program 100 data |
| IND_TPL | From program 100 data |
| IPROVTYP | From program 100 data |
| MC_PLAN | From program 100 data |
| MC_TYPE | From program 100 data |
| MCARE_X | From program 100 data |
| PAY_DATE | From program 100 data |
| PR1 | From program 100 data |
| PR1_DT | From program 100 data |
| PR1_IND | From program 100 data |
| PR1_MOD | From program 100 data |
| PR2 | From program 100 data |

Medicaid Inpatient Service Detail Variables – SERV_IN

| | |
|----------|----------------------------|
| PR2_DT | From program 100 data |
| PR2_IND | From program 100 data |
| PR2_MOD | From program 100 data |
| PR3 | From program 100 data |
| PR3_DT | From program 100 data |
| PR3_IND | From program 100 data |
| PR3_MOD | From program 100 data |
| PR4 | From program 100 data |
| PR4_DT | From program 100 data |
| PR4_IND | From program 100 data |
| PR4_MOD | From program 100 data |
| PR5 | From program 100 data |
| PR5_DT | From program 100 data |
| PR5_IND | From program 100 data |
| PR5_MOD | From program 100 data |
| PR6 | From program 100 data |
| PR6_DT | From program 100 data |
| PR6_IND | From program 100 data |
| PR6_MOD | From program 100 data |
| PR7 | From program 100 data |
| PR7_DT | From program 100 data |
| PR7_IND | From program 100 data |
| PR7_MOD | From program 100 data |
| PR8 | From program 100 data |
| PR8_DT | From program 100 data |
| PR8_IND | From program 100 data |
| PR8_MOD | From program 100 data |
| PR9 | From program 100 data |
| PR9_DT | From program 100 data |
| PR9_IND | From program 100 data |
| PR9_MOD | From program 100 data |
| PROV_ID | From program 100 data |
| PROVCNTY | From program 100 data |
| PROVTYPE | From program 100 data |
| SERV_ID | Assigned during processing |

SERV_LTC – Medicaid Long Term Care Claims

- Sorted by CLIENTID and SERV_ID
- Multiple observation per client

Medicaid Long Term Care Service Detail Variables – SERV_LTC

| | |
|----------|---|
| CLIENTID | From program 400 data |
| AMT_CAID | From program 100 data |
| AMT_CARE | From program 100 data |
| AMT_CHGD | From program 100 data |
| AMT_COIN | From program 100 data |
| AMT_DED | From program 100 data |
| AMT_PL | From program 100 data |
| AMT_TPL | From program 100 data |
| CLNT_ZIP | From program 100 data |
| COVDAYS | From program 100 data |
| CSO | From program 100 data (Washington only) |
| DSTATUS | From program 100 data |
| DX1 | From program 100 data |
| DX2 | From program 100 data |
| DX3 | From program 100 data |
| DX4 | From program 100 data |
| DX5 | From program 100 data |
| DX6 | From program 100 data |
| DX7 | From program 100 data |
| DX8 | From program 100 data |
| DX9 | From program 100 data |
| FFS_IND | From program 100 data |

Medicaid Long Term Care Service Detail Variables – SERV_LTC

| | |
|----------|----------------------------|
| IADJ_IND | From program 100 data |
| IMCARE_X | From program 100 data |
| IND_CAID | From program 100 data |
| IND_CARE | From program 100 data |
| IND_CHGD | From program 100 data |
| IND_COIN | From program 100 data |
| IND_DED | From program 100 data |
| IND_PL | From program 100 data |
| IND_TPL | From program 100 data |
| IPROVTYP | From program 100 data |
| LDAYS | From program 100 data |
| LEV_CARE | From program 100 data |
| MC_PLAN | From program 100 data |
| MC_TYPE | From program 100 data |
| MCARE_X | From program 100 data |
| PAY_DATE | From program 100 data |
| PROV_ID | From program 100 data |
| PROVCNTY | From program 100 data |
| PROVTYPE | From program 100 data |
| REVCODE | From program 100 data |
| SERV_ID | Assigned during processing |

SERV_RX – Medicaid Pharmacy Claims

- Sorted by CLIENTID and SERV_ID
- Multiple observation per client

Medicaid Pharmacy Service Detail Variables – SERV_RX

| | |
|----------|---|
| CLIENTID | From program 400 data |
| ADJ_IND | From program 100 data |
| AMT_CAID | From program 100 data |
| AMT_CHGD | From program 100 data |
| AMT_PL | From program 100 data |
| AMT_TPL | From program 100 data |
| CLNT_ZIP | From program 100 data |
| CSO | From program 100 data (Washington only) |
| DRUG_CD | From program 100 data |
| FFS_IND | From program 100 data |
| IADJ_IND | From program 100 data |
| IND_CAID | From program 100 data |
| IND_CHGD | From program 100 data |
| IND_PL | From program 100 data |
| IND_TPL | From program 100 data |
| IPROVTYP | From program 100 data |
| MC_PLAN | From program 100 data |
| MC_TYPE | From program 100 data |
| PAY_DATE | From program 100 data |
| PROV_ID | From program 100 data |
| PROVCNTY | From program 100 data |
| PROVTYPE | From program 100 data |
| REFILL | From program 100 data |
| RX_DATE | From program 100 data |
| RX_DAYS | From program 100 data |
| RX_FILL | From program 100 data |
| RX_PHYS | From program 100 data |
| RX_QTY | From program 100 data |
| SERV_ID | Assigned during processing |
| TH_CLASS | From program 100 data |

SERV_OTH – Outpatient/Other Medicaid Claims

- Sorted by CLIENTID and SERV_ID

- Multiple observation per client

Outpatient/Other Medicaid Service Detail Variables – SERV_OTH

| | |
|----------|---|
| CLIENTID | From program 400 data |
| ADJ_IND | From program 100 data |
| AMT_CAID | From program 100 data |
| AMT_CARE | From program 100 data |
| AMT_CHGD | From program 100 data |
| AMT_COIN | From program 100 data |
| AMT_DED | From program 100 data |
| AMT_PL | From program 100 data |
| AMT_TPL | From program 100 data |
| CLNT_ZIP | From program 100 data |
| CSO | From program 100 data (Washington only) |
| DX1 | From program 100 data |
| DX2 | From program 100 data |
| DX3 | From program 100 data |
| DX4 | From program 100 data |
| DX5 | From program 100 data |
| DX6 | From program 100 data |
| DX7 | From program 100 data |
| DX8 | From program 100 data |
| DX9 | From program 100 data |
| FFS_IND | From program 100 data |
| IADJ_IND | From program 100 data |
| IMCARE_X | From program 100 data |
| IND_CAID | From program 100 data |
| IND_CARE | From program 100 data |
| IND_CHGD | From program 100 data |
| IND_COIN | From program 100 data |
| IND_DED | From program 100 data |
| IND_PL | From program 100 data |
| IND_TPL | From program 100 data |
| IPROVSPC | From program 100 data |
| IPROVTYP | From program 100 data |
| MC_PLAN | From program 100 data |
| MC_TYPE | From program 100 data |
| MCARE_X | From program 100 data |
| PAY_DATE | From program 100 data |
| PROC_CD | From program 100 data |
| PROC_IND | From program 100 data |
| PROC_MOD | From program 100 data |
| PROV_ID | From program 100 data |
| PROVCNTY | From program 100 data |
| PROVSPC | From program 100 data |
| PROVTYP | From program 100 data |
| REV_CD | From program 100 data |
| SERV_ID | Assigned during processing |
| SERV_QTY | From program 100 data |
| SERV_UNT | From program 100 data |
| SERVPLC | From program 100 data |
| SERVPROV | From program 100 data |

SERV_COM – Community Mental Health Services

- Sorted by CLIENTID and SERV_ID
- Multiple observation per client

Community Mental Health Treatment Detail Variables – SERV_COM

| | |
|----------|----------------------------|
| CLIENTID | From program 400 data |
| CAP_AMT | From program 200 data |
| COST | From program 200 data |
| COST_IND | From program 200 data |
| DXCODE | From program 200 data |
| ISERVICE | From program 200 data |
| SERV_ID | Assigned during processing |
| SERV_LOC | From program 200 data |
| SERV_QTY | From program 200 data |
| SERV_UNT | From program 200 data |
| SERVPROV | From program 200 data |

SERV_IMH – Institutional Mental Health Services

- Sorted by CLIENTID and SERV_ID
- Multiple observation per client

Institutional Mental Health Treatment Detail Variables – SERV_IMH

| | |
|----------|----------------------------|
| CLIENTID | From program 400 data |
| CAP_AMT | From program 200 data |
| CDAYS | From program 200 data |
| COST | From program 200 data |
| COST_IND | From program 200 data |
| DXCODE | From program 200 data |
| ISERVICE | From program 200 data |
| REV_CD | From program 200 data |
| SERV_ID | Assigned during processing |
| SERV_LOC | From program 200 data |
| SERVPROV | From program 200 data |
| STAYTYP | From program 200 data |

SERV_AOD – Alcohol/Other Drug Treatments

- Sorted by CLIENTID and SERV_ID
- Multiple observation per client

Alcohol/Other Drug Treatment Detail Variables – SERV_AOD

| | |
|----------|----------------------------|
| CLIENTID | From program 400 data |
| CAP_AMT | From program 200 data |
| COST | From program 200 data |
| COST_IND | From program 200 data |
| DXCODE | From program 200 data |
| ISERVICE | From program 200 data |
| SERV_ID | Assigned during processing |
| SERV_LOC | From program 200 data |
| SERV_QTY | From program 200 data |
| SERV_UNT | From program 200 data |
| SERVPROV | From program 200 data |

TEMPORARY STORAGE

Temporary storage is available for data sets that are needed for longer than a single job, but are not part of the output for a program can be stored for approximately one to two weeks (depending on space availability). The directories for temporary storage are:

- Delaware /saswrk/de/program600/
- Oklahoma /saswrk/ok/program600/
- Washington /saswrk/wa/program600/

APPENDIX A - CONSISTENCY CODES

The data base utilizes consistency codes for most character and numeric data. Simple missing values (blank for character and "." for numeric variables) should never occur when consistency codes are indicated. Consistency codes should be utilized as shown below.

- Unknown – the information is available on the source data, and the value is documented as missing or unknown.
- Unavailable – the information is not available on the source data.
- Invalid – the information is available on the source data, but the value is not documented.

For example if a source file of client data did not contain gender information, GENDER should be coded as "unavailable". (Hopefully, gender information is always supplied with client data).

Suppose instead that gender information is supplied with the client data, and the documentation specifies these possible values:

- "F" for females,
- "M" for males, and
- "U" for unknown

When the source values are "U" or blank, GENDER should be coded as "unknown", and when the source values are "T" or "Z", GENDER should be coded as "invalid".

Consistency codes for numeric variables are:

- .A Unknown
- .B Unavailable
- .C Invalid

For character variables, the consistency codes are dependent on the length of the variable. Generally character consistency codes are:

- "7", "97", or "997" Unknown
- "8", "98", or "998" Unavailable
- "9", "99", or "999" Invalid
-

APPENDIX B - SAMPLE CODE DIAGNOSIS BASED FLAGS

```

** array for flags <- NEW FLAGS/NAMES **;
Array FLAGS {22} FSCHIZO  FMAJDEP  FAFFPSYC  FOTHPSYC  FALZHMR
                FOTHORG  FSTRESS  FPERSDIS  FCHILDIS  FOTHMOOD
                FOTHMENT  FALCPSYC  FALCABSE  FDRGPSYC  FDRGABSE
                FOTHAOD  FTOBACCO  FHIVAIDS  FPOTNALC  FMENTDEV
                FVCODE   FPREG;

** array for diagnosis codes;
Array DXCODE {9} DX1-DX9;

** set flags to '0' at start;
Do I = 1 To 22;
  FLAGS{I} = '0';
End;

** cycle through diagnosis codes in reverse order to capture
primary;
Do I = 9 To 1 By -1;
  ** check for pregnancy complication <- NEW FLAG **;
  If (FPREG ne '1') Then FPREG = put(DXCODE{I}, $PREG.);

  ** check for selected conditions;
  Select (put(DXCODE{I}, $DSM.));
    When ('01') FSCHIZO = put(min(2, I), 1.);
    When ('02') FMAJDEP = put(min(2, I), 1.);
    When ('03') FAFFPSYC = put(min(2, I), 1.);
    When ('04') FOTHPSYC = put(min(2, I), 1.);
    When ('05') FALZHMR = put(min(2, I), 1.);
    When ('06') FOTHORG = put(min(2, I), 1.);
    When ('07') FSTRESS = put(min(2, I), 1.);
    When ('08') FPERSDIS = put(min(2, I), 1.);
    When ('09') FCHILDIS = put(min(2, I), 1.);
    When ('10') FOTHMOOD = put(min(2, I), 1.);
    When ('11') FOTHMENT = put(min(2, I), 1.);
    When ('21') FALCPSYC = put(min(2, I), 1.);
    When ('22') FALCABSE = put(min(2, I), 1.);
    When ('23') FDRGPSYC = put(min(2, I), 1.);
    When ('24') FDRGABSE = put(min(2, I), 1.);

```

IDB Users Guide

```
When ('25') FOTHAOD = put(min(2, I), 1.);
When ('27') FTOBACCO = put(min(2, I), 1.);
When ('31') FHIVAIDS = put(min(2, I), 1.);
When ('32') FPOTNALC = put(min(2, I), 1.);
When ('33') FMENTDEV = put(min(2, I), 1.);
When ('34') FVCODE = put(min(2, I), 1.);
Otherwise;
End; ** end-select;
End; ** end-do-I;
```

APPENDIX C - INCLUSION CRITERIA CODE

```
** ----- **
** inclusion criteria for final data base <- NEW CRITERIA **;
If (FSCHIZO In ('1', '2')) or
    FMAJDEP In ('1', '2')) or
    FAFFPSYC In ('1', '2')) or
    FOTHPSYC In ('1', '2')) or
    FSTRESS In ('1', '2')) or
    FPERSDIS In ('1', '2')) or
    FCHILDIS In ('1', '2')) or
    FOTHMOOD In ('1', '2')) or
    FOTHMENT In ('1', '2')) or
    FALCPSYC In ('1', '2')) or
    FALCABSE In ('1', '2')) or
    FDRGPSYC In ('1', '2')) or
    FDRGABSE In ('1', '2')) or
    FVCODE In ('1', '2')) or
    FOTHAOD In ('1', '2')) or
    FTOBACCO In ('1', '2')) or
    FTYPSEV In ('1', '2')) or
    FREVCTR In ('1', '2')) or
    FPRVTYPE In ('1', '2')) or
    FPROC In ('1', '2'))
Then WANTED = 1;
```


APPENDIX D – DELAWARE SOURCE CODE AND DATA DICTIONARY

APPENDIX E – OKLAHOMA SOURCE CODE AND DATA DICTIONARY

APPENDIX F – WASHINGTON SOURCE CODE AND DATA DICTIONARY