USER GUIDE

National Adult Medicaid (NAM)
Consumer Assessment of Healthcare Providers and Systems Survey (CAHPS)
Public Use File, 2014-2015

1. Introduction

The first-ever National Adult Medicaid Consumer Assessment of Healthcare Providers and Systems (NAM CAHPS) survey was identified as one of the core set of adult quality measures to monitor health care access and experiences of care among adult enrollees. The goal of the 2014-2015 NAM CAHPS survey was to obtain nationwide and state estimates of adult Medicaid enrollees' experiences of care, including access to and utilization of services, across different financing and delivery models (e.g., managed care and fee-for-service) and population groups (e.g., individuals who are dually eligible for Medicare and Medicaid and individuals with disabilities who are not dually eligible).

The development of a streamlined de-identified Public Use File (PUF) is not intended to replace the more detailed Limited Data Set (LDS) files, rather it will improve data access, lead to more and different types of researchers using the NAM CAHPS, and allow preliminary research to be conducted.

2. Background of NAM CAHPS

The 2014-2015 Nationwide Adult Medicaid (NAM) CAHPS survey, sponsored by the Centers for Medicare & Medicaid Services (CMS), Center for Medicaid and CHIP Services (CMCS), will provide states, stakeholders, and CMCS with national and state-specific data on the health care experiences of the non-institutionalized adult Medicaid population across financing and delivery models. It will also inform the development of standard measures anchoring a robust health care quality strategy for the adult Medicaid population. A rigorous CAHPS survey design, data collection effort, analytic plan, and strategy for documentation and dissemination of findings are critical to facilitating quality improvement efforts for the adult Medicaid population.

3. PUF Technical and programming information

General Information: The NAM CAHPS PUF includes 272,679 survey respondents and 96 variables. All records begin with an -PUF_ID, a unique number for each beneficiary in the public use file. This ID serves to identify records in the PUF only and cannot be used for linking to other sources of data.

Data File Information: Variable name, type, label and values are available in the PUF codebook that is available with the data download. The codebook also includes unweighted frequencies and percentages for each variable for users' reference. For most variables, the label refers to the questionnaire text that was asked during the survey. The label also indicates if the variable is a recode or combination of two other variables (e.g. STATE_STRATUM).

If you have downloaded the NAM CAHPS PUF export (NAMCAHPSPUF16.xpt) file in the folder "**C:\NAMCAHPS\DOWNLOAD**". You can then use the following SAS code to import it into SAS.LIBNAME PUFLIB 'C:\NAMCAHPS\SASDATA'; FILENAME F "C:\NAMCAHPS\DOWNLOAD\NAMCAHPSPUF16.XPT";

PROC IMPORT LIBRARY=PUFLIB INFILE=F; RUN; Additionally, a comma-separated values (CSV) file is available for use with other statistical software packages such as R® and Stata®.

Recoded Variables: The PUF contains coarsened versions of AGE, EDUCATION and RACE variables that have been globally recoded to have fewer categories. These variables make a less precise version of suppressed variables available to the data user. We include on the data file three different coarsened versions of these variables as follows:

a. The original 'Level 1' version of the variable with all response categories from the survey intact (this version is the same as what is available in the LDS File);

b. A 'Level 2' version of the variable where some response categories are collapsed (this version is the same as what is available in the NAM CAHPS Data Resource Website);

c. A new 'Level 3' version of the variable where even more response categories are collapsed.

Comparison to LDS: The NAM CAHPS PUF differs from the NAM CAHPS LDS files because it has been evaluated for disclosure risk and additional steps were taken to protect respondent confidentiality. Twenty-seven variables were dropped from the LDS and 11 recoded variables were added to the PUF.

A summary of the differences between the two data products appears in Table 1 below:

Table 1. Summary of differences between NAM CAHPS PUF and NAM CAHPS LDS

NAM CAHPS PUF	**NAM CAHPS LDS**
Number of variables: 96	Number of variables: 112
PUF_ID – new PUFID Randomly generated, can't be linked back to BASE ID	SID – BASEID Randomly generated
Suppressed cells -- YES	Suppressed cells -- NO
Reserved values: Appropriately Skipped—77 Multiple mark—88 Missing—99 Missing/multi/suppressed—98	Reserved values: Appropriately Skipped—7 or 77 Multiple mark—8 or 88 Missing—9 or 99
Variables in PUF, but not in LDS: AGE1, AGE2, AGE3, RACE1, RACE2, RACE3, EDUCATION1, EDUCATION2, EDUCATION3, ETHNICITY1, SEX1	Variables in LDS, but not in PUF: AGENCY, MC_NAME, LANGUAGE, MODE, Q65 (recoded to AGE in PUF), Q66 (recoded to SEX in PUF), Q67 (recoded to EDUCATION in PUF), Q68a, Q68b, Q68c, Q68d, Q68e, (recoded to ETHNICITY in PUF) Q69a, Q69b, Q69c, Q69d, Q69e, Q69f, Q69g, Q69h, Q69i, Q69j, Q69k, Q69l, Q69m, Q69m, Q69o, (recoded to RACE in PUF)

Note: it was not necessary to adjust the weights from the NAM CAHPS LDS for the NAM CAHPS PUF because the two sampling variables, state and stratum, are already revealed in the PUF. Therefore, the weighting variable, **finalweight**, is the same in both data files.

4. Survey sample information

Background on design and sampling frame: The target population for the survey was adults ages 18 and older as of December 31, 2013, who were continuously enrolled in Medicaid during the first quarter of federal fiscal year (FFY) 2014 (FFY 2014 Q1, October 2013 – December 2013), who were not residing in an institutional setting, and were not part of a Family Planning Waiver. The sample was drawn from adults who were enrolled in Medicaid prior to a state's expansion of coverage to the new adult group that took effect on or after January 1, 2014.

The sampling frame for the survey was designed to capture four key subgroups of adult Medicaid enrollees. The main criteria for stratifying the frame were state (including the District of Columbia) and the following four mutually exclusive enrollee groups based on program eligibility:

- Adults dually eligible for Medicaid and Medicare (Full Duals);
- Adults (non-duals) with disabilities (Persons with Disabilities);
- Adults (non-duals, non-disabled) enrolled in a managed care organization (Managed Care, or MC); and,
- Adults (non-duals, non-disabled) who obtained care from a fee-for-service (FFS) provider or were enrolled in a primary care case management plan (FFS-PCCM).

Group classification was mutually exclusive, determined using the hierarchy shown above. In total, 46 states plus the District of Columbia provided sampling data from which we were able to draw the sample. We drew a total sample of over 1.2 million cases, averaging approximately 29,000 sampled enrollees per state. Data collection was conducted from December 14, 2014 through July 27, 2015, across four waves, with a roughly similar number of states per wave. The questionnaire was administered first through mail, and then with telephone follow-up where necessary. This effort resulted in 272,679 enrollees completing the survey, with an overall response rate of 23.6%. For more details about the construction of each state's sampling frame, the final sample size and sample selection, please refer to the *2014-2015 NAM CAHPS Methodology Report*.

Weights and variance estimation: The purpose of estimation in the 2014-2015 Nationwide Adult Medicaid CAHPS survey is to infer enrollee experiences in the target population based on the data collected from the sample. The weighting procedures make adjustments for variation in sampling rates and differential response rates to produce robust estimates at state level. The weighting scheme involved the following steps:

1) Base sampling weights;

Sample was selected within each of the four strata, within each state. The weighting process started with computing the baseweights of the sampled enrollees, where the base weight was the reciprocal of the selection probability of an enrollee. This resulted in four distinct base weights for each state, one per stratum (noting that some states did not have all four strata).

2) Adjustment for ineligible adults;

An adjustment to the contacted cases was necessary to account for cases that were deemed ineligible. To make the weighting adjustment, the number of adjustment cells within each stratum and state was formed by controlling for known covariates. Adjustment cells for the sample within each state and stratum were defined by age group and gender. Logistic regression analysis was performed to determine variables available on the sampling frame that are associated with being contacted.

3) Adjustment for nonresponse among eligible adults.

Not all contacted survey-eligible enrollees completed the 2014-2015 NAM CAHPS interview. The weight assigned to an enrollee responding to the interview was adjusted to account for the nonresponse of other survey-eligible enrollees.

A Taylor-series approximation is used to calculate the variance for estimates. Since the survey used a stratified single-stage sampling design, where the stratifiers were state and group within state, the user will need to be sure to define the stratification done both by state and strata (i.e., Full Duals, Disabled, Managed Care, FFS-PCCM). There is an additional variable that was created that concatenates both state and stratum together for this purpose, STATE_STRATUM. Also, statistical languages such as SAS, Sudaan, and Stata are all able to compute these variance estimates for analysis. For detailed information on summary of final survey weights and variance estimation, please refer to the *2014-2015 NAM CAHPS Methodology Report.*

5. Sample code to generate cross-tabulations of weighted data with standard errors

The sample code below shows how a data user with access to the 2014-2015 NAM CAHPS PUF could generate a **cross-tabulation** of weighted data and calculate standard errors using Taylor series linearization using different statistical packages.

SAS ANALYSIS STATEMENTS

```
proc surveyfreq data=DATANAME;
        table VARIABLE1*VARIABLE2 / row col ;
        strata STATE_STRATUM;
        weight FINALWEIGHT;
ods output crosstabs=TABLENAME;
run;
```

STATA ANALYSIS STATEMENTS

```
svyset [pweight= FINALWEIGHT], strata(STATE_STRATUM)
svy: tab VARIABLE1 VARIABLE2
```

R ANALYSIS STATEMENTS

mydesign<-svydesign(weights=~ **FINALWEIGHT**, strata=~ **STATE_STRATUM** ,data=mydata)

svytable(~**VARIABLE1+VARIABLE2**, design=mydesign)

6. Sample code to create a combined race_ethnicity variable

One variable likely to be of interest to PUF users is a combined variable for race_ethnicity. Users can easily create a combined race_ethnicity variable using the race and ethnicity variables included in the PUF according to their needs.

For example, the most detailed version of the race_ethnicity variable with all response categories intact can be created using the PUF variables RACE1 and ETHNICITY1 as shown in Table 2 below.

Table 2. Example of creation of combined variable for race_ethnicity

SAS Variable:	**RACE_ETHNICITY1**
Question:	Combined variable of RACE1 and ETHNICITY1
Value:	Description
1	White, Not Hispanic/Latino
2	Black/African American, Not Hispanic/Latino origin
3	American Indian or Alaska Native, Not Hispanic/Latino origin
4	Asian, Not Hispanic/Latino origin
5	Pacific Islander, Not Hispanic/Latino origin
6	Other Race, Not Hispanic/Latino origin
7	Multiple races, Not Hispanic/Latino origin
8	Hispanic, Latino/a, or Spanish origin
98	Missing/multi/ suppressed

The example SAS code to create the race_ethnicity variable is below:

```
If   ethnicity1=1 then race_ethnicity1 = 8;
else if race1 = 1 then race_ethnicity1 = 1;
else if race1 = 2 then race_ethnicity1 = 2;
else if race1 = 3 then race_ethnicity1 = 3;
else if race1 = 4 then race_ethnicity1 = 4;
else if race1 = 5 then race_ethnicity1 = 5;
else if race1 = 6 then race_ethnicity1 = 6;
else if race1 = 7 then race_ethnicity1 = 7;
else if ethnicity1=98 or race1 = 98 then race_ethnicity1 = 98;
```