Instructions for Analyzing Data from CAHPS® Surveys:

Using the CAHPS Analysis Program Version 4.1

Purpose of this Document 1
The CAHPS Analysis Program 1
Computing Requirements 1
Pre-Analysis Decisions 2
What Does the CAHPS Analysis Program Analyze? 6
Preparing Data for Analysis 8
SAS Data Set Requirements 12
Using the CAHPS Analysis Program 20
Interpreting the Results 36
Small Data Set Example 45
Explanation of Statistical Calculations 49
Hypothesis Tests and Assignment of Final Ratings 61
Examining Sample Size Issues for CAHPS Surveys 63

Appendix

Summary of Features Included in Each Version of the CAHPS Analysis Program ... 66

List of Exhibits and Tables

Table 1. Description of test data set variables 22
Table 2. Arguments for CAHPS 4.1 Macro 28
Table 3. Effect size detected with 80 percent power (alpha = 0.05) by number of plans and sample size per plan 64
Table 4. Effect size detected with 80 percent power (alpha = 0.05) by number of plans and sample size for one plan (n = 300 for all other plans) 64

For additional guidance, please e-mail cahps1@ahrq.gov or call the Help Line at (800) 492-9261.

Purpose of this Document

This document explains how the CAHPS Analysis Program works and how sponsors and vendors can use the program to interpret the results of their CAHPS survey. While the program was initially designed for the CAHPS Health Plan Survey, you can use it to analyze data from any of the CAHPS surveys. For most CAHPS surveys, the instructions include a document with analysis guidance specific to that survey.

The CAHPS Analysis Program

The goal of the CAHPS Analysis Program – often referred to as the CAHPS macro – is to provide the user with a flexible way to analyze CAHPS survey data in order to make valid comparisons of performance. Written in SAS, the CAHPS Analysis Program is designed to assist CAHPS survey users in implementing two kinds of statistical adjustments.

Comparing items and composites. CAHPS surveys collect consumers' and patients' reports and ratings of a number of dimensions of health care. Comparing performance based on all the individual CAHPS survey items is a very complex task. Moreover, individual survey items are often less reliable than multiple item combinations. To simplify the interpretation of the data and enhance the reliability of the results, questions that measure similar topics are grouped together. These groups of questions, called **composites**, facilitate comparisons of performance across your unit of analysis (e.g., health plan, medical practice, clinician). (Note: The instructions available for each survey include a document that lists the items in that survey's reporting measures, i.e., composite measures and ratings.)

Adjusting for case mix. The CAHPS Team recommends that you adjust the survey data for respondent age, education, and general health status. This makes it more likely that reported differences are due to real differences in performance, rather than differences in the characteristics of enrollees or patients.

Computing Requirements

The CAHPS Analysis Program was developed using SAS® software. SAS is a data management, analysis, and presentation product produced by the SAS Institute, which is headquartered in Cary, North Carolina. The operation of SAS requires a Base system, but a number of individual modules can be added to perform more complex analyses and data manipulation.

The CAHPS Analysis Program requires Base SAS and the SAS/STAT module. Base SAS, which is required to use any SAS product, provides the data manipulation, print commands, simple plotting capabilities, and procedures for descriptive statistics. Base SAS includes the CORE module. The SAS/STAT module adds several statistical procedures for use by SAS. The CAHPS Analysis Program uses the SAS regression procedure, PROC REG, to do part of its case-mix calculations. If case-mix calculations will not be used, then the macro requires Base SAS only.

The CAHPS Analysis Program was written in version 6.12 of PC SAS and version 6.12 of SAS/STAT. The program should work on all SAS platforms that have version 6.0 or later. It has been extensively tested on UNIX and Windows SAS products and has performed well.

Pre-Analysis Decisions

The CAHPS Analysis Program offers the user a number of options for analyzing the survey data. Before preparing to run the program, analysts should make sure that the project team has agreed upon answers to the following questions. Their implications for the CAHPS Analysis Program are reviewed below. Having these questions answered early will save time when doing the analyses.

What is the reporting unit (entity)?

Any analysis of CAHPS data is intended to assess, compare and report on some type of reporting unit. Examples of possible such units include health plans, hospitals, provider groups, clinics, sites of care, and individual physicians. To avoid confusion, we use the neutral term "entity" in these instructions to refer to the unit whose data will be aggregated into a summary measure. Users of the Analysis Program will have to specify which variable identifies the entity to which each response will be attached. Note that there might be alternative ways of analyzing the same data with different entities, and if the data collection design is suitable, more than one of them might be valid. For example, a dataset might be analyzed to compare provider groups and then, with a different "entity" variable, the same data might be used to assess individual doctors.

The CAHPS Analysis Program was initially written for the CAHPS Health Plan Survey. Therefore, variable names, examples, and other references throughout the program often refer to health plans. This has no bearing on the suitability of the program for analyzing data on other types of entities. For specific instructions on adapting the program for other surveys, please refer to ***Preparing and Analyzing Data (****https://www.cahps.ahrq.gov/Surveys-Guidance/CG/~/media/Files/SurveyDocuments/CG/12%20Month/Prep_Analyze/1035_Preparing_analyzing_data_from_cg.pdf****)*** in the instruction documents available for the survey you are using. This document is not available for all CAHPS surveys.

Will you analyze specific population groups separately?

If the project team has collected data for different groups of people, the team needs to decide whether to analyze the data separately or together. Subgroups that you may be considering for separate analyses can be defined by payer (e.g., Medicare, Medicaid, privately insured), geographic region (e.g., state, county, region), or other factors. If the groups are to be analyzed together, no changes to the CAHPS Analysis Program are necessary. If a team decides to analyze the groups separately and the data file contains more than one group, it is important to set up selection criteria in the CAHPS Analysis Program or split the data set.

Will adult and child surveys be analyzed together or separately?

The Analysis Program allows users to specify how child and adult surveys will be analyzed. The project team needs to decide whether to analyze surveys about adults and children separately or together. If adult and child survey data are to be analyzed together, the team must also decide whether to consider interaction effects. Interaction effects are important to consider in an analysis when the impact of age or health status on one of the reporting items depends on whether an adult or child survey is being analyzed. See the section called ***Explanation of Statistical Calculations*** for a more detailed discussion of interaction effects. We recommend that you consider interaction effects when combining adult and child data.

If the team collected only adult or child surveys, users still need to specify an option listed below in the CAHPS Analysis Program. There are four options depending on which surveys are in the data set and how the user wants to analyze them:

Option value	Surveys in data set	Analysis method
0	Adult and child surveys, only adult surveys, or only child surveys	Combine adult and child survey data; do not consider interaction effects
1	Adult and child surveys	Combine adult and child survey data; consider interaction effects
2	Adult and child surveys or only child surveys	Child surveys only
3	Adult and child surveys or only adult surveys	Adult surveys only

Will high and low users of health care services be analyzed together or separately?

The CAHPS Analysis Program allows users to analyze the data for survey respondents with a high or low number of outpatient visits separately or combined. The criterion for establishing low and high users of outpatient services is based on Question 7 of the ***CAHPS Health Plan Survey – Adult Commercial Questionnaire*** or Question 4 of the ***CAHPS Clinician & Group Surveys – Adult 12-Month and Visit Questionnaires***. High users are defined as respondents who indicated that they have had three or more visits for their own care to a doctor's office or clinic (not including emergency room visits). Low users indicated that they have had fewer than three visits to a doctor's office or clinic.

The CAHPS Consortium recommends reporting data for global ratings (e.g., respondent's rating of their personal doctor or nurse, specialists, health care in the last 12 months, and/or health plan) according to outpatient utilization. It is up to the project team to decide whether to analyze the survey data for low and high users

separately or together. You can then choose from the corresponding Option Values listed below for the CAHPS Analysis Program.

Option value	Analysis decision
1	Combine low and high users.
2	Low users only (< 3 visits)
3	High users only (≥ 3 visits)

What level of significance (p-value) will you use in the analysis?

The CAHPS Consortium recommends a *p*-value of 0.05 to test for statistically significant differences between the entities being compared. The *p*-value the team chooses must be specified in the CAHPS Analysis Program.

What, if any, level of substantive (practical) significance will you use to compare performance?

Substantive significance refers to an absolute difference between the entities being compared (e.g., health plans, provider groups, individual physicians) that must be achieved before a change is considered meaningful. If two health plans, for example, had significantly different average scores based on the *p*-value criteria, the difference between the plans' average scores may still not be large enough to be meaningful.

The CAHPS Analysis Program has two options that allow the user to specify a difference that is substantive. You can use these options simultaneously or specify only one.

First method. The team decides on a percentage of the distance to the nearest bound that is meaningful. The example presented below explains this concept.

Assume the analysis of a global rating question (one that uses a 0-10 rating scale) has the following mean scores for a global rating question across all entities:

Global Rating Question overall mean (0-10 scale) = 6

To determine a level of difference between entities that is substantively large:

1. Determine the distance from the mean to the nearest bound:

   a) Compute the distance from the lower bound

      Mean (6) - Lower bound (0) = 6

   b) Compute the distance from the mean to the upper bound

      Upper bound (10) - mean (6) = 4

c) Determine the smaller difference

Minimum (6,4) = 4

2. The distance from the nearest bound is 4. Now the project team must decide what percentage of this distance is a meaningful difference between entities. This fraction is entered in the CAHPS Analysis Program.

Second method. A much simpler method available in the Analysis Program is to specify an absolute difference that must exist between the entity mean and the mean for all entities in the analysis for a difference to be considered significant. For this method, the user needs only to specify the absolute difference considered to be meaningful.

Do you need to adjust the results for case mix?

Case mix refers to the respondents' health status and sociodemographic characteristics, such as age or educational level, that may affect the ratings that the respondent provides. Without an adjustment, differences between entities could be due to case-mix differences rather than true differences in quality. Each project team must determine if it is appropriate to adjust its data to account for case-mix differences.

What to adjust for

If the project team decides to adjust the data for case-mix, it must choose the appropriate adjusters. The CAHPS Consortium recommends using general health status, age, and education. Individuals in better health and older individuals tend to rate their care, plans, and providers higher. There is also evidence from a number of studies that education affects ratings, with more educated individuals giving lower ratings.

Missing data for case-mix adjusters

If case-mix adjusted results will be used, the project team must decide whether or not to impute missing data for the adjusters at each adjuster's entity-level mean.

Risk of out-of-range values for case-mixed means

In the special cases where there are very few records for an analysis variable or all respondents answered in only one or two response categories, there is the possibility that the case-mix adjusted values will be out of range. For example, if all respondents to a Health Plan Survey answered "Yes," where 0= "No" and 1= "Yes" to a yes/no question, and the adjustment for that entity is up, the adjusted mean for that entity would be greater than 1 and the adjusted frequencies would be less than zero for the "No" category and greater than 1 for the "Yes" category.

The macro does not force a change in these values, since it would change the mean of the means on the adjusted scores but not on the unadjusted scores. It is recommended that, in reports of CAHPS results, you set these out-of-range values to the minimum or maximum value for that category. Then a manual adjustment could be made to the adjacent category if necessary. For example, in the case of three response categories, where the minimum frequency should be zero and the maximum value is 100, the case-mixed frequency results are as follows:

category 1 =	-2.0,
category 2 =	25.0 and
category 3 =	77.0

The results could be adjusted so that

category 1 =	0.0,
category 2 =	23.0 and
category 3 =	77.0

Do results need to be analyzed using weighting and stratification?

As discussed above, the survey sampling plan can be designed to select disproportionately potential respondents from certain geographic or demographic groups in the population. Alternatively, situations can arise after sampling is complete that create the need to combine data for certain sampling units. For example, this can occur when two entities merge their operations and the survey sponsor chooses to report their results as a combined score. Whether entities are merging their operations or a disproportionate stratified sampling design was used, the CAHPS Analysis Program can perform the appropriate analyses, provided the user specifies the correct strata to be combined and the number of members in each stratum out of the entire population.

What Does the CAHPS Analysis Program Analyze?

The CAHPS Analysis Program is designed to analyze multi-item composites and single items from the CAHPS survey data. The output from the program compares the performance of an entity to the overall performance of all entities.

The macro accepts five variable types. For four of these, the macro cleans their response values within the expected minimum and maximum range. For the fifth type, the minimum and maximum response values must be entered as an argument.

The following are the variable types:

Variable type	Min – Max response values
1 – Dichotomous	0 – 1
2 – Global rating	0 – 10
3 – How often or other 4-point response scale	1 – 4
4 – 3-point response scale	1 – 3
5 – Other	min_resp – max_resp

Global ratings are based on survey items that ask the respondents to rate different aspects of health care on a scale from 0 to 10. For example, the global ratings items in the ***CAHPS Health Plan Survey 4.0 – Adult Commercial Questionnaire*** are:

- Health care in the last 12 months (Question 8)
- Personal doctor (Question 15)
- Specialist (Question 19)
- Health plan (Question 27)

Overall Ratings Response Format: 0-10	
Q8	Using any number from 0 to 10, where 0 is the worst health care possible and 10 is the best health care possible, what number would you use to rate all your health care in the last 12 months?
Q15	Using any number from 0 to 10, where 0 is the worst personal doctor possible and 10 is the best personal doctor possible, what number would you use to rate your personal doctor?
Q19	We want to know your rating of the specialist you saw most often in the last 12 months. Using any number from 0 to 10, where 0 is the worst specialist possible and 10 is the best specialist possible, what number would you use to rate the specialist?
Q27	Using any number from 0 to 10, where 0 is the worst health plan possible and 10 is the best health plan possible, what number would you use to rate your health plan?

Composites are groupings of two or more questions that measure the same dimensions of health care or health insurance plan services. Composites usually are developed for survey items that have the same response options.

For example, the composites in the Health Plan Survey represent the experiences of respondents (adult enrollees or the parents/guardians of enrolled children) in the following areas:

- Getting needed care (2 questions for adults; 2 questions for children).
- Getting care quickly (2 questions for adults; 2 questions for children).
- How well doctors communicate (4 questions for adults; 5 questions for children).
- Health plan information and customer service (2 questions for adults; 2 questions for children).

The following table illustrates how items in the Health Plan Survey 4.0 are grouped into the first of these composites.

Getting Needed Care		
Q17	In the last 12 months, how often was it easy to get appointments with specialists?	**Response Format** ▪ Never ▪ Sometimes ▪ Usually ▪ Always
Q21	In the last 12 months, how often was it easy to get the care, tests, or treatment you thought you needed through your health plan?	

A document listing each survey's reporting measures – composites and ratings – is available in the instructions provided for that survey.

Preparing Data for Analysis

Prior to applying the CAHPS Analysis Program, you must perform several tasks to transform raw data from the completed questionnaires into data that the SAS analysis programs can use. (Guidance on determining when a questionnaire is complete is available for most CAHPS surveys in the appendix of the instructions on fielding that survey.)

- Task 1: Identify and exclude ineligible cases.
- Task 2: Code and enter the data.
- Task 3: Clean the data.
- Task 4: Conduct an audit.

Many interim files will be created along the way. Before beginning this process, you must take steps to preserve the original data file created when the raw survey responses were entered. Any changes and corrections made during the cleaning and data preparation phase should be made on duplicate files. There are three reasons for this action:

1. The original data file is an important component of the complete record of the project.
2. Having an original file will allow you to correct data errors that were made during the cleaning process.
3. The existence of an original file is critical if the vendor or sponsor wants to go back later and conduct other analyses or tests, such as extent of error tests or tests of skip patterns.

Data File Specifications

The data file contains the raw data from responses to the survey. Short item handles for the items in each questionnaire can be found in tables provided in the document called ***Overview of the Questionnaires at*** *http://cahpscms.westat.com/Surveys-Guidance/CG/~/media/Files/SurveyDocuments/CG/12%20Month/Get_Surveys/1350_cg_overview_of_questionnaires.pdf*. The responses to each question must use the code numbers, or precodes, contained in the questionnaires.

Users should construct a separate data file for each version of the survey. Do not include data from different survey instruments in the same data file. For example, do not include responses to the Clinician & Group Adult 12-Month Survey and Child 12-Month Survey in the same data file.

If you are interested in submitting your data to the CAHPS Database, please refer to the data file specifications for submission at *https://www.cahps.ahrq.gov/CAHPS-Database/Submitting-Data/CG-Data.aspx.*

The number and scope of the data preparation tasks and the way they are carried out depend on the data collection protocol and the way in which the data were recorded. For example:

- If the vendor/sponsor collected data with a self-administered mailed questionnaire, did respondents record answers on optical scan forms[1] or record them directly on the CAHPS-formatted questionnaires?
- If the vendor/sponsor collected data through telephone interviews, did the interviewer use computer-assisted telephone interviewing (CATI) or paper-and-pencil forms?

[1] Optical scan forms are answer sheets in which respondents fill in the circle that corresponds to their answer choice. These forms are fed through an optical scanning machine, and the data are automatically captured by a computer. Standardized tests for students, such as the SAT, generally use optical scan forms.

Task 1: Identify and Exclude Ineligible Cases

Several situations render a case ineligible for analysis. One common scenario that vendors must be prepared to handle occurs when the respondent reports he or she has not visited the sampled entity (e.g., a physician or medical group). This might be indicated by a "no" response to Question 1 (e.g., "Our records show that you got care from the provider named below in the last 12 months. Is that right?").

Other questionnaires may be considered incomplete and excluded from analysis but are not excluded from the denominator used to calculate the response rate. For example:

- If someone else assisted the respondent or answered the questions (as a proxy) or
- If at least half of the key items on the questionnaire were not filled in. (For most CAHPS surveys, a list of key items is available in the appendix of the instructions on fielding the survey.)

Task 2: Code and Enter the Data

There are a variety of possible methods that can be used to enter data from CAHPS surveys. The exact level of coding required will depend on the method used to capture the data (e.g., questionnaires that require data entry versus questionnaires that are scanned by a computer). Your coding specialist should review each questionnaire to see whether the responses are legible and whether any responses need to be coded. Each item should have a corresponding code, even items that were not answered. The table below shows examples of recommended coding. If you are interested in submitting your data to the CAHPS Database, follow the data file and coding specifications available at *https://www.cahps.ahrq.gov/CAHPS-Database/Submitting-Data/CG-Data.aspx*. After coding is completed, enter the data into a computer file.

Response	Code
Actual response	Use the corresponding survey precode indicated beside the response option. Precodes typically begin with 1 and number consecutively for each response option.
Item was appropriately skipped	Code as 7 (or 77 if precodes exceed 7)
Item showed more than one response option completed when only one is appropriate (i.e. multiple marks)	Code as 8 (or 88 if precodes exceed 7)
Item was left blank and should not have been skipped	Code as 9 (or 99 if precodes exceed 7)

If you use **optical scan forms** for your mailed questionnaires, the scanning equipment automatically enters the data into a computer-readable file. If you do not use optical scan forms, the mail questionnaires are designed for direct data entry without the need for coding most respondent answers. However, if it is unclear which answer the respondent selected (e.g., the respondent's pencil mark does not neatly fit within a single answer category, or two responses are marked), then your coding specialist will have to make a decision about which response the respondent intended. If it is not readily apparent what the respondent intended, the coding specialist should indicate that the answer be entered as missing or multiple marks, as appropriate.

If you use a **CATI system** for a telephone survey, data are entered directly into a data file that has already been programmed to refuse unlikely and invalid responses. If you use paper questionnaires to record answers given in a telephone interview instead, the process for coding and data entry is the same as for the standard paper version of the mailed questionnaires.

To ensure quality, answers from paper-and-pencil questionnaires should be key-entered by two separate data entry specialists. The results from the two should be compared to identify and correct data entry errors. At the end of the coding and data entry process, you will have an electronic data set of responses to all the questionnaire items.

Additional coding and recoding may be necessary prior to using the CAHPS Analysis Program. Refer to the **SAS Data Set Requirements** below for recoding variables for use with the CAHPS Analysis Program.

Task 3: Clean the Data

In many cases, the data set you have created will have imperfections. You will have to take several steps to fix these imperfections before any results are reported.

Check for out-of-range values. Out-of-range responses occur when respondents provide inappropriate responses for a particular question. For example, if the valid response choices for a question are 0 or 1, a value of 2 would be out of range. Similarly, if a respondent circled two categories when he or she was supposed to provide only one answer, the response is out of the acceptable range of the question.

To detect out-of-range values, you need to review question frequencies. This can be done either by visually scanning a report showing the item distributions or frequencies or by running the questionnaire data through a computer program., Both are often used to improve the quality of the data. If a value is found that is impossible (or unlikely) given the response options, then the questionnaire should be reviewed and revisions made to the data. These revisions often involve setting the out-of-range values to "missing," which drops them from the data analysis for that particular question. Carefully document the results from this review process, including any changes to the data set. Maintain an audit trail (electronically and on paper) so it is possible to go back to the original data file.

Check for skip pattern problems. Response inconsistencies generally arise when a respondent misunderstands a question or does not successfully follow instructions to

skip questions. An example of a response inconsistency would be if a respondent answered that he or she had no doctor visits in the past 6 months, but then answered followup questions about visits in the past 6 months. If there are inconsistencies between the response to the screener question and the following question's response, assume the screener response is correct.

Check again for ineligible cases. Identify any questionnaires that are not eligible for analysis and remove them from the data set used for the CAHPS SAS program. Questionnaires in which fewer than half of key items are answered should not be counted as completed surveys. (Guidance on determining when a questionnaire is complete, including a list of key items, is available for most CAHPS surveys in the appendix of the instructions on fielding that survey.)

Check for duplicates. The number of records in the data file should match the number of completes and partials in the sample file. Duplicates can occur if the vendor conducts a followup phone interview, if the mail questionnaire arrives at the same time or soon after, and the case slips through the receipt control system, or if there are errors in data entry. Your policy should be to keep the first questionnaire that comes in.

Task 4: Conduct an Audit

Whether surveys are collected in standard paper format, as optically scanned forms, or as paper telephone questionnaires, a small random sample of the entered data should be audited by comparing hard-copy forms with the results of data entry. This enables you to catch any systematic errors. For example, if the optical scanning program was incorrect, Question 5 may be entered in the file where Question 6 was supposed to be. These types of systematic errors will show up consistently across all questionnaires.

SAS Data Set Requirements

Before running the CAHPS Analysis Program, make sure that the structure and properties of the data file meet the specifications listed below. If the data set does not meet these requirements, the SAS program will not work properly. Many of the variable coding and cleaning requirements are demonstrated in the next section on using the CONTROL.SAS Analysis Program.

Data File Specifications

- Each row or case in the SAS data set represents the data for a unique questionnaire.
- If data from different CAHPS questionnaires are in the same data set and are to be analyzed together, each questionnaire is listed on a separate row. If data from adult and child questionnaires are in the data set, the adult and child questionnaires are also listed on separate rows.

- If data from different CAHPS questionnaires are in the same data set, responses for equivalent questions are listed under the same variable names.

Sample Size Requirements

Number of entities (i.e., such as health plan or providers). The data set must have surveys from at least two entities. If there is only one entity in the data being analyzed, statistical comparisons cannot be performed and some parts of the program will not work properly. If the CAHPS macro is run with data for one entity, a couple of warning messages and notes will appear in the log file that would not be produced if two or more entities were represented. All the reports will still be produced, though some of the results will be of limited value.

Responses. At least two responses per entity are required by the Analysis Program. We recommend analysis of at least 100 responses for each entity. The program flags entities with fewer than 100 responses for an individual measure, but the analysis is performed on all entities with at least two records. Including entities with very little data tends to reduce the precision of comparisons between individual entities or providers and the overall means. The user can consider removing entities with fewer than 100 responses from the data file before analysis.

Note: When analyzing units of analysis such as medical groups or individual physicians, follow the minimum response guidance in the instructions for fielding the survey. Since the program was initially designed for the CAHPS Health Plan Survey, you will receive the program flags in your results when there are fewer than 100 responses even if the target number of completed responses is less for your survey.

Variable Coding and Cleaning Requirements

Numeric variables. All analytic variables used by the CAHPS Analysis Program must be numeric. Analytic variables include any questionnaire item used to compute CAHPS reporting items, case-mix adjustment variables, the dichotomous variable used to identify child and adult surveys, and the variable used to identify high and low users of outpatient services. These variables are discussed in more detail below. To ensure that an error does not occur in the SAS program, all variables created from survey questions should be coded numerically. If the user recodes character variables to numeric, there should be a minimum length of 4. A length of 8 is recommended for the recode.

WARNING: The variables PLAN, CHILD, VISITS, and SPLIT are variable names needed by the CAHPS macro. If the data set has other variables with these names and they do not conform to the specifications below, the macro may produce errors in the log file and the results may be erroneous.

Variable PLAN. The variable PLAN must be included in the data set. Note that this variable represents a numeric code or text descriptor for each entity in the data set. This is the only variable that does not have to be coded numerically. The SAS program accepts alphanumeric, character, and numeric formats for this variable. The maximum variable length for PLAN is 40 characters.

Even if you are not analyzing health plan data, you must use the variable name PLAN to refer to your unit of analysis. This is because the SAS macros use that variable name. However, the variable can be any unit of analysis. For example, PLAN can represent the names of the medical groups (Group A, Group B, etc.) or individual physicians (Dr. A, Dr. B, etc.).

Variable CHILD. The numeric variable CHILD needs to be in the data set if subsetting the data between adult and child records. This variable is used to distinguish between adult and child surveys in the SAS program. CHILD should be coded 0 for adult surveys and 1 for child surveys. If this variable is missing from the data set, the CAHPS macro sets CHILD = 0 when ADULTKID has the values 0, 1, or 3, and sets CHILD = 1 when ADULTKID = 2.

Variable VISITS. The variable VISITS needs to be in the data set if using the VISITS parameter. This variable is used to identify high and low users of health care services (e.g., refer to Question 7 from the ***CAHPS Health Plan Survey 4.0 -- Adult Commercial Questionnaire).***

The table that follows shows the response values based on item 7 in the ***CAHPS Health Plan Survey 4.0 -- Adult Commercial Questionnaire***. We recommend that you use these values while coding your questionnaires. If the VISITS variable is missing from the data set, the CAHPS macro will work as long as the VISITS parameter is not equal to 1.

7. In the last 12 months, not counting times you went to an emergency room, how many times did you go to a doctor's office or clinic to get care for yourself?

Response value	Label/description
0	None
1	1 time
2	2 times
3	3 times
4	4 times
5	5 to 9 times
6	10 or more times
All other values	Not analyzed by the SAS program

Variable SPLIT. The numeric variable SPLIT needs to be in the data set if you are doing separate case-mix adjustments on two different populations as indicated by the macro parameter SPLITFLG = 1. For most cases, the default value 0 for SPLIT does not need to be modified. An example of splitting the case-mix adjustments separately on two populations is when comparing Medicaid Fee-for-Service populations with Medicaid Managed Care populations.

Yes/No Variables. Variables with "yes/no" response categories to be used in the analysis are coded as shown in the table below. Any variable with dichotomous response options should be coded in this manner. For easier interpretation of the

results, the "positive" response should have the highest value. Raw data for this type of variables will need to be recoded as the precodes typically set the values of the responses to 1 and 2 rather than 0 and 1.

Response value	Label/description
0	No
1	Yes
All other values	Not analyzed by the SAS program

Three Response Options. Any variable with three response options should be coded as shown in the table below. For easier interpretation of the results, the "positive" response should have the highest value. Reverse coding may be necessary to ensure that the most positive response has the highest value—for example, where "Yes, definitely" is the most positive response.

Response value	Label/description
1	Yes, definitely
2	Yes, somewhat
3	No
All other values	Not analyzed by the SAS program

Four-Point Frequency scale. Variables with "never" to "always" response options are coded as shown in the table below. Any variable with four response options should be coded in this manner. For easier interpretation, the "positive" response should have the highest value. Reverse coding may be necessary to ensure that the most positive response has the highest value—for example, where "never" is the most positive response.

Response value	Label/description
1	Never
2	Sometimes
3	Usually
4	Always
All other values	Not analyzed by the SAS program

Response value	Label/description
1	Definitely no
2	Somewhat no
3	Somewhat yes
4	Definitely yes
All other values	Not analyzed by the SAS program

Global Ratings. Global rating items with 0-10 response options are coded as shown in the table below:

Response value	Label/description
0	Worst
1	…
2	…
3	…
4	…
5	…
6	…
7	…
8	…
9	…
10	Best
All other values	Not analyzed by the SAS program

Coding for Adjuster Variables. If the project team decides to case-mix adjust the CAHPS survey results, numeric variables must also be properly coded for each adjuster variable. The CAHPS Consortium recommends adjusting the data for age, education, and general health status; however, the program allows for a flexible number of adjuster variables. The user can choose the proper specification for each variable used to adjust the data. Specifications used for age, education, and general health status in the CAHPS Analysis Program are described below. Users may also specify the variables as dichotomous with reference categories (dummy variables). It is important to remember that the SAS program previously cleaned out-of-range values for these variables. However, the added flexibility of a user-specified number of variables and specification of the variables makes it necessary to code out-of-range values to missing before running data through the macro.

The coding specification for the numeric variable EDUCATION is included in the data set. Education refers to the respondent's highest level of school completed. This variable and its response codes should be coded based on the responses to the education item (such as Question 35 in the ***CAHPS Health Plan Survey -- Adult Commercial Questionnaire***).

35. What is the highest grade or level of school that you have completed?

Response value	Label/description
1	8th grade or less
2	Some high school, but did not graduate
3	High school graduate or GED
4	Some college or 2-year degree
5	4-year college graduate
6	More than 4-year college degree
All other values	Code to missing

The coding specification for the numeric variable GENERAL HEALTH RATING (GHR) is included in the data set. The GHR is a rating of the survey respondent's overall health status. This variable and its response codes should be based on the responses to the health status item (for example, Question 28 in the ***CAHPS Health Plan Survey -- Adult Commercial Questionnaire***).

28. In general, how would you rate your overall health now?

Response value	Label/description
1	Excellent
2	Very good
3	Good
4	Fair
5	Poor
All other values	Code to missing

The numeric variable AGE is included in the data set, representing age groupings based on data from the CAHPS survey. This variable and its response value codes should be based on questions about age, such as Question 33 of the ***CAHPS Health Plan Adult Commercial Questionnaire*** and Question 32 of the ***CAHPS Health Plan Child Commercial Questionnaire***. The response values for these questions need to match the values for AGE as follows:

33. What is your age? / 32. What is your child's age?

Response value	Label/description (years)
For child surveys:	
0	< 1
1	1-3
2	4-7
3	8-12
4	13-17
For adult surveys:	
1	18 to 24
2	25 to 34
3	35 to 44
4	45 to 54
5	55 to 64
6	65 to 74
7	75 or older
All other values	Code to missing

Stratified Data. If you want to combine data for reporting from different sampling strata, you must create a text file that identifies the strata and indicates which ones are being combined and the identifier of the entity obtained by combining them. Some examples illustrate situations in which this feature might be used:

- Two health plans are merged that were formerly separate and were treated as such in the survey.
- A hospital decides to sample 200 medical and 200 surgical patients, although this is not proportional to the numbers of discharges in the two services.
- A survey designed to assess providers samples 80 patients from each regardless of the number of sessions each has, but the sponsor decides to also use the data to assess provider groups.

If no file is specified, the macro creates one using the PLAN variable in the data set to set the "Original Plan" and "New Plan" equal to the PLAN variable, the "Population Size" equal to 1, and "Subsetting Code" equal to 1. If stratification is part

of your survey design, an ASCII data set needs to be created with columns separated by one or more spaces for these four variables:

- **Original Plan –** a unique identifier of the units or strata before they are combined. This variable can be coded as alphanumeric, but it cannot exceed 16 characters. This variable is the first column of the data table.
- **New Plan –** identifier for the entities that will be created by combination of strata. This variable can be coded as alphanumeric, but it cannot exceed 16 characters. This variable is the second column of the data table. If no stratification is being done, this column may look identical to the column for original plan.
- **Population Size –** a numeric variable that indicates the size of the population for the unit or stratum. This variable is used to create the weights for combining the strata. The populations for the combined strata should equal the total population of the "new plan." This variable is the third column of the data table. If no stratification is being done, this column may be set to 1s.
- **Subsetting Code –** identifier for the subset (i.e., region, state, county…) that the entity belongs in. This variable can be coded as alphanumeric. If no subsetting is to be done, this column may be set to 1s.

The ASCII file for the plan details should not contain any missing data and each column of data should be separated by spaces. If tabs are used, the macro may not read in the data correctly. Also, be sure to not have any extra records at the bottom of the ASCII file. If you want to make a quick sample plan detail file from the CAHPS data set, use the program make_plandtal_dat.sas as a starting point and change variable names and paths as needed. If the number of plans is small, it is probably easier to create the file by hand.

An example of the plan detail data set is provided for the test program (test.sas). The data file is called "plandtal.dat" and looks like the text below:

```
HMO_A_URBAN HMO_A 5000 Northeast
HMO_B_URBAN HMO_B 8000 Northeast
HMO_C_URBAN HMO_C 15000 Atlantic
HMO_B_RURAL HMO_B 2000 Northeast
HMO_C_RURAL HMO_C 3000 Atlantic
```

The TEST data set provides an example for three health plans (2nd column): HMO_A, HMO_B, and HMO_C. The urban/rural strata for HMO_B and HMO_C are weighted together and the combined plans compared to HMO_A, which had members only in urban areas.

- The first column provides a unique identifier for each plan/region combination (original plan).
- The second column, the new plan variable, indicates which units will be combined.
- The third column, or the unit population size, is used to compute the weights for the plans. Units with greater population sizes receive more weight than smaller units in the combined plan.
- The fourth column is the region (subset) of the country in which each plan does business.

Adult and Child Interactions (Macro Parameter ADULTKID)

When the macro parameter ADULTKID equals 1, the macro creates adult and child interactions for the adjuster variables. The macro creates additional adjuster variables, with the a set naming convention, AC1, AC2, ..., ACn, where n is the total number of adjusters originally submitted in the macro call parameter ADJUSTER. When there is an adult and child interaction, the macro creates the ACx variables by looping through the list of adjusters.

For example:

If &ADJUSTER = GHR AGE EDUCATION, then the following additional interaction adjuster variables are created:

AC1 = GHR * CHILD

AC2 = AGE * CHILD

AC3 = EDUCATION * CHILD

Using the CAHPS Analysis Program

The steps described below assume a basic knowledge of how to use the SAS system.

Step 1: Loading the programs and test data

Version 4.1 of the CAHPS Analysis Program contains three core components:

- a SAS control program,
- a SAS macro, and
- a “Plan Detail” data file.

All program and data sets needed for the CAHPS Analysis Program, Version 4.1, are available for downloading in the instructions provided for every CAHPS survey. Below is a description of the purpose of each file. Each file should be copied to a project folder related to the CAHPS data set that is to be analyzed.

CONTROL.SAS

CONTROL.SAS is a SAS program that contains examples of the macro call parameter arguments that need to be specified to produce the recommended reporting measures for CAHPS surveys. For most surveys, specific examples of the macro call are provided in separate instructions on analyzing the results of that survey. The program also demonstrates the variable cleaning and coding steps needed to perform the analyses for entity-level comparisons. Modifications most likely will need to be made to this program to reflect how variables are named, how variables are coded or formatted, whether or not entity stratification is used, and whether or not the data set includes child surveys, among the many possible combinations.

CAHPS41.SAS

CAHPS40.SAS is the core SAS macro program that performs the analyses the user specifies in the control program. The macro file should not be modified.

PLANDTAL.DAT

This is a sample data set that is used by the CAHPS macro when running the TEST data set. It contains the unique plan names, combined strata names, strata weight, and subsetting code. If this file is not included in your control programs, the macro will create this data set based on the PLAN variable in the input data set.

SMALLTEST.SAS

This program creates a small data set with ten records that can be used to better understand what the CAHPS macro is doing. More details on how this data progresses through the macro can be found in the ***Small Data Set Example*** section in this document.

FORMAT.SAS

FORMAT.SAS is the SAS program that creates formats helpful to view the data with English words instead of the data values assigned in the TEST data set. The formats have the essence of the tables described in the section ***SAS Data Set Requirements***. The program creates the formats in the library named LIBRARY as defined by the libname LIBRARY statement in CONTROL.SAS and TEST.SAS programs. The file, currently set up to work with the test data programs, can be modified for use with other data. Modifications, such as changing the values of the formats, adding new formats or deleting formats, are the most common.

TEST.SAS

TEST.SAS is a test control program for the CAHPS Analysis Program macros. It was designed for use with the test data sets described below. This program was constructed to provide users with a short program and data set that demonstrated the analysis options and output for Version 4.1 of the CAHPS Analysis Program.

The hypothetical example was designed to incorporate a disproportionate sampling design of individuals in rural and urban areas for three health plans. One health plan (HMO_A) has members only in urban areas, while the other two plans (HMO_B, HMO_C) have members in both urban and rural areas. To make comparisons across the three plans, the data for the plans with members sampled from rural and urban regions need to be combined using weights.

Various options are used in the test program to demonstrate the analysis features available to the user. Analyses are performed for all three types of reporting items (single questions, global ratings, and composites). A variety of analysis features are also used, including a varying number of case-mix adjustment variables, turning off the option that creates output data sets, and the weighting option.

TEST.SAS7BDAT, TEST.SSD01, TEST_windows.SAS7BDAT, TEST.SD2

TEST*.* files are SAS data sets that contain the same variables. TEST.SSD01 and TEST.SAS7BDAT are for use with UNIX SAS programs, and TEST.SD2 and TEST_windows.SAS7BDAT are designed for use with the Windows version of SAS. The table below describes the variables in the data sets and provides value labels for each variable. (Note that the test data files and sets were created for the CAHPS Health Plan Survey 3.0 and have not yet been updated.)

Table 1. Description of test data set variables

Variable	Description	Response options
ID	Respondent identification number	Unique numeric value
PlanID	Plan identification number	1 = HMO_A_URBAN ; 2 = HMO_B_URBAN ; 7 = HMO_C_URBAN ; 4 = HMO_B_RURAL ; 5 = HMO_C_RURAL ; 6 = HMO_BE_1 ; . = Missing
Q31	Global rating of care	0 (worst) - 10 (best) .=Missing 98=Inapplicable 99=No Answer Given

Variable	Description	Response options
Q38	Global rating of plan	0 (worst) - 10 (best) .=Missing 98=Inapplicable 99=No Answer Given
Q06	Problem to get doctor or nurse that you were happy with	1=A Big Problem 2=A Small Problem 3=Not a Problem .=Missing 98=Inapplicable 99=No Answer Given
Q10	Problem to get a referral to see a specialist that you needed	1=A Big Problem 2=A Small Problem 3=Not a Problem .=Missing 98=Inapplicable 99=No Answer Given
Q22	Problem in getting the care you or doctor believed was necessary	1=A Big Problem 2=A Small Problem 3=Not a Problem .=Missing 98=Inapplicable 99=No Answer Given
Q23	Problem with delays in health care while waiting for approval from plan	1=A Big Problem 2=A Small Problem 3=Not a Problem .=Missing 98=Inapplicable 99=No Answer Given
Q05	Was it easy to find a personal doctor or nurse	1=Yes 2=No .=Missing 98=Inapplicable 99=No Answer Given
Q15	How often got the help or advice you needed	1=Never 2=Sometimes 3=Usually 4=Always .=Missing 98=Inapplicable 99=No Answer Given

Variable	Description	Response options
Q17	How often got appointment as soon as you wanted	1=Never 2=Sometimes 3=Usually 4=Always .=Missing 98=Inapplicable 99=No Answer Given
Q19	How often got needed care as soon as you wanted	1=Never 2=Sometimes 3=Usually 4=Always .=Missing 98=Inapplicable 99=No Answer Given
Q24	How often wait in doctor's office	1=Never 2=Sometimes 3=Usually 4=Always .=Missing 98=Inapplicable 99=No Answer Given
Q39	General health rating	1=Excellent 2=Very Good 3=Good 4=Fair 5=Poor .=Missing 98=Inapplicable 99=No Answer Given
Q40	Age of respondent	1=18 to 24 2=25 to 34 3=35 to 44 4=45 to 54 5=55 to 64 6=65 to 74 7=75 or older .=Missing 98=Inapplicable 99=No Answer Given
Q41	Gender	1=Male 2=Female .=Missing 98=Inapplicable 99=No Answer Given

Variable	Description	Response options
Q21	Office and clinic visits in the past 6 months	1=None 2=1 time 3=2 times 4=3 to 4 times 5=5 to 9 times 6=10 or more times .=Missing 98=Inapplicable 99=No Answer Given

Step 2: Modifying CONTROL.SAS

CONTROL.SAS is a SAS program that invokes and executes the macro file, CAHPS41.SAS, to perform basic analyses for the CAHPS surveys. This program can be modified to perform the analyses that the team has decided to conduct. Statements from CONTROL.SAS are demonstrated below. The user can alter this program to perform analyses on other data sets. The modifications demonstrated below apply to the test SAS program, TEST.SAS, as well. TEST.SAS demonstrates many of the key concepts for a limited number of variables described earlier. (Note that the test data files and sets were created for the CAHPS Health Plan Survey 3.0 and have not yet been updated.)

Identifying the Data Set, Macros, Program, and Output File Locations

The program statements below specify the library reference and file names for the macros and data sets. These statements should be modified based on the location of the files. Note: The filename statements creating “logfile” and “outfile” are not necessary unless the user wants to save the log information to a file named CONTROL.LOG and the printed results to a file named CONTROL.TXT. The libname statement creating “out” statement is required to identify the location where the data sets of the summary results the program creates will be placed.

```
%let ProgramName = control ;
%let root  = /data/cahpsmmc/analysis_program/version_4.1 ;

libname in  “&root./sasdata/” ;
libname out  “&root./sasdata/control/” ;
libname library “&root./sascatalog/” ;

filename logfile “&root./output/logs/&ProgramName..log” ;
filename outfile “&root./output/&ProgramName..txt” ;
filename cahps “&root./programs/cahps41.sas” ;
filename plan_dat “&root./data_other/plandtal.dat” ;
```

Output Table Titles

The following code clears the SAS titles and footnotes from the output data tables. This step ensures that any titles and footnotes previously created in a SAS session are cleared before you execute the control program or macro.

```
title ;
footnote ;
```

Print the Output to a File

The following statements can be removed if the user does not want to save the results to an explicitly stated external file.

```
proc printto print = outfile new
    log = logfile new
 ;
run ;
```

To return the log and output to the default sources, include the following lines of code at the end of the control file.

```
proc printto ;
run ;
```

Data Set Specifications

The following statements prepare the test data set according to the specifications outlined under ***Computing Requirements***. You may need to make modifications to the following statements depending on the variable names and variable response options in the data set. It is very important that all variables in that section are in a temporary or permanent SAS data set that will be used for the analysis. General health status (Q39) and age (Q40) variables are prepared as case-mix adjusters for illustrative purposes.

```
1. Set permanent or temporary SAS data set.

   data adult ( drop = i ) ;
    set in.test ;
```

```
2. Recodes numeric plan variables to character to simplify
interpretation of the result tables.

       length plan $ 16 ;
       if  planid = 1 then plan = 'HMO_A_URBAN' ;
       else if planid = 2 then plan = 'HMO_B_URBAN' ;
       else if planid = 7 then plan = 'HMO_C_URBAN' ;
       else if planid = 4 then plan = 'HMO_B_RURAL' ;
       else if planid = 5 then plan = 'HMO_C_RURAL' ;
       else if planid = 6 then plan = 'HMO_BE_1' ;


3. Creates visits variable.

       visits = q21 ;

4. Creates child variable by coding it to 0 for all surveys.

       child = 0 ;


5. Recodes dichotomous variables from 1-2 to 1-0.

       array yn q05 q13;
       do i = 1 to dim ( yn ) ;
        if yn [i] = 0  then yn [i] = . ;
        else if yn [i] = 2 then yn [i] = 0 ;
       end ;

6. REVERSE codes item in which never is a positive response and
always is a negative response.

   array rev q24 ;
   do i = 1 to dim ( rev ) ;
    if rev [i] in (1, 2, 3, 4) then rev [i] = 5 - rev [i] ;
    else       rev [i] = . ;
   end ;


7. Version 1.5 and higher of the CAHPS program does not automatically
clean case-mix adjustment variables as previous versions did because it
allows for a varying number and specification of the adjusters. If
adjusters are used in the analysis they must be cleaned first.

   age = q40 ;
   ghr = q39 ;

   if ghr not in (1, 2, 3, 4, 5)  then ghr = . ;
   if age not in (1, 2, 3, 4, 5, 6, 7) then age = . ;
```

CAHPS41.SAS—Specifying Arguments and Options

The following statement includes the macro code CAHPS41.SAS.

```
%include cahps ;
```

The macro call statements for CAHPS41.SAS in CONTROL.SAS require that at least six arguments, (VAR, VARTYPE, NAME, ADULTKID, DATASET, and OUTNAME), be specified for it to work properly. These arguments, along with the 20 optional arguments, are listed in the table below with the valid value ranges. The six arguments must be specified for each analytic run of the global ratings and

composites. If using adjusters, then the ADJUSTER argument is required. The macro call can be repeated any number of times in the SAS program for the different composites and ratings the user wants to compare. Users can also analyze the same composite or global rating more than once by using different macro arguments.

For each composite, the user needs to specify the items listed below in CONTROL.SAS. Arguments with an asterisk (*) are optional and are needed only in specific cases.

Table 2. Arguments for CAHPS 4.1 Macro

Option or argument	Description	Values
`Var`	Name of variable(s) in composite, or global rating	Name of variable(s) from SAS data set to include in composite or global rating. For composites, separate the variable names by a single space.
`Vartype`	Type of variable	1 = Dichotomous Scale (yes/no 0-1) 2 = Global Rating Scale (0-10) 3 = How Often Scale or other four-point response scale (“never” to “always” scale 1-4) 4 = Any type of three-point response scale (1-3) 5 = Other Scale (Must assign a value to min_resp and max_resp arguments)

Option or argument	Description	Values

`* Recode`	Recodes the global rating* and the 'How Often' scales down to three categories before performing the case-mix adjustment and the statistical tests—the default value is 0	0 = For the statistical tests, use default response options for the variables in the Var argument.
		For the "Percent of each response" table and report, split the "Rating" scale into three categories with the following break points, 0-6|7-8|9-10 or 1-2|3|4 for the "How Often" scale.
		Recode option is not needed in the CAHPS macro call if it = 0.
		1 = For the statistical tests, recode the Global Rating Scale (0-10, vartype = 2) and the "How Often" scale (1-4, vartype = 3) as:
		`Rating` `How Often`
		`0 - 6 = 1` `1 - 2 = 1`
		`7 - 8 = 2` `3 = 2`
		`9 - 10 = 3` `4 = 3`
		If vartype is not equal to 2 or 3, then no recoding occurs for the statistical tests.
		For the "Percent of each response" table and report, split the "Rating" scale into three categories with the following break points, 0-6|7-8|9-10 or 1-2|3|4 for the "How Often" scale.
		2 = For the statistical tests, use default response options for the variables in the Var argument.
		For the "Percent of each response" table and report, split the "Rating" scale into three categories with the following break points, 0-7|8-9|10 or 1-2|3|4 for the "How Often" scale.

Option or argument	Description	Values
* Recode (continued)		3 = For the statistical tests, recode the Global Rating Scale (0-10, vartype = 2) and the "How Often" scale (1-4, vartype = 3) as: Rating: 0 – 7 = 1; 8 – 9 = 2; 10 = 3 4 = 3 How Often: 1 – 2 = 1; 3 = 2 If vartype is not equal to 2 or 3, then no recoding occurs for the statistical tests. For the "Percent of each response" table and report split the "Rating" scale into three categories with the following break points, 0-7|8-9|10 or 1-2|3|4 for the "How Often" scale.
* Min_resp	Used with vartype = 5 only—the minimum response value	Can be any numeric value. It will be used as the low value for the valid response options.
* Max_resp	Used with vartype = 5 only—the maximum response value	Can be any numeric value. It will be used as the high value for the valid response options.
Name	Description of composite or global rating	Note: This argument is limited to 40 characters and can be numeric, text, or a combination of both.
* Adjuster	Name(s) of adjuster variables	Name(s) of case-mix adjuster variables—separated by a space if using more than 1.
* Adj_bars	Flag indicating if the frequencies for the response values are to be case-mix adjusted for the triple stacked bar—the default value is 0	0 = Do not case-mix adjust the triple stacked bars. 1 = Case-mix the triple stacked bars and store the adjusted frequencies along with the unadjusted frequencies.

Option or argument	Description	Values
`* Bar_stat`	Flag indicating if permanent data sets for the case-mixed frequencies should be saved—the default value is 0	0 = Do not case save the statistical results in data sets for the case-mix adjusted triple stacked frequency bars. 1 = Save the case-mix adjusted statistical results in permanent data sets for the triple stacked frequency bars.
`* Impute`	Imputation of missing data for adjuster variables—the default value is 0	0 = Do not impute mean values by plan for all adjuster variables. 1 = Impute mean values by plan for all adjuster variables.
`* Even_wgt`	Determines how to weight composite items—the default value is 1	0 = Use item weighting for composites. 1 = Use equal weighting for composites (1 / # of Items). 2 = Apply the respondent level weight, in WGTRESP, to the item weighting for composites.
`* K`	Assign a target minimum response size for equal weighting for composites (even_wgt = 1) - the default value is 1.	Number ≥ 0.
`* Kp_resid`	Flag used to make the residual values from the SAS work data set RES_4_ID in the STD_DATA module. The residuals are the response values after case-mix adjustments have been made – the default value is 0	0 = Do NOT save the residual response values. 1 = Save the residual response values in a permanent data set.

Option or argument	Description	Values
`Adultkid`	Specifies how to analyze child and adult surveys	0 = Combine adult and child surveys in analysis; do not consider interaction effects in case-mix adjustment. This option can also be used if the data set contains only a single type of survey. 1 = Combine adult and child survey data in analysis; consider interaction effects between child and each case-mix adjuster variable. 2 = Analyze child data only. 3 = Analyze adult data only.
`* Visits`	Specifies whether to analyze high and low users together or separately (based on VISITS variable)—the default value is 1	1 = All visits. 2 = Low users only (< 3 visits per 6 months). 3 = High users only (≥ 3 visits per 6 months).
`* Pvalue`	Level of significance for comparisons—the default value is 0.05	0.05 recommended, but valid values are between 0 and 1.
`* Change`	Level of practical significance based on a percentage difference from the minimum absolute theoretical difference from the overall mean (can be used only with '*p*-value' criteria)—the default value is 0	Value between 0 and 1 (i.e., 25% is entered as 0.25).
`* Meandiff`	Level of practical significance based on absolute difference between plan mean and mean of all plans (can be used only with '*p*-value' criteria)—the default value is 0	Number ≥ 0.

Option or argument	Description	Values
`* Wgtdata`	Specifies whether samples are stratified within health plan—the default value is 1	1 = Do not invoke weighting macro 2 = Combine strata, weighting
`* Wgtresp`	Name of the variable storing the weight values for individual respondents. —the default value is blank	Blank or the name of a variable in the data set
`* Wgtmean`	Name of the variable storing the weight values for the plan means—the default value is blank	Blank or the name of a variable in the data set
`* Wgtplan`	Specifies whether to use plan weights for the plan level statistical test or not. The default value is zero.	0 = Do not use the plan weights when computing the overall mean for the comparison of plan means. Equal weighting will be used as in previous versions of the macro. 1 = Use the sum of the weights to the plan level of the variable specified in the parameter wgtmean. This weight is used for weighting the overall and grand means used in the statistical comparisons of the plan means.
`* Id_resp`	If there is unique variable in the data set that identifies each individual respondent, then this variable name may be entered here—the default value is blank	Blank or the name of a variable in the data set. This variable is included in the residual data set when kp_resid = 1. The variable will be a character and have a maximum of 50 characters.

Option or argument	Description	Values
`* Subset`	Perform the case-mix adjustments and statistical test based on each subset of plans; the subset code is a column in the plan detail file—the default value is 1	1 = No subsetting done. Global case-mix model and centering. 2 = Global case-mix model with centered means for each subset before performing statistical tests. 3 = Subset case-mix model with centered means for each subset.
`* Splitflg`	The default value of 0 lets the macro run the data set as usual with every plan centered to the same mean and the case-mix being run once. If the flag is set to 1 then the data set must contain the variable SPLIT and the values of this variable in the data set must be 0 and 1 for each plan subset.	0 = Run macro with one case-mix model 1 = Run macro with two case-mix models Example: Managed Care plans split=0 and for the other plans, Fee for Service, split=1. If there are any missing values for this variable, then these records are dropped from the analysis; the default value is 0.
`* Smoothing`	Assign weight for pooled variance estimate in smoothing variances - the default value is 0.	Value greater than 0. For a detailed explanation of how this value is selected, see the section called ***Explanation of Statistical Calculations***.
`Dataset`	SAS data set name to be used in the analysis (variables recoded and renamed according to "Computing Requirements")	Data set name depends on how you called in the file.
`* Outregre`	A flag that indicates whether or not to include the regression output text created by SAS, in the final text report file—the default value is 0	0 = No regression output appears in the text report file. 1 = SAS printout from the PROC REG is included in the text report file.
`Outname`	Part of SAS data set name for output tables created for summary results	Maximum length is five characters if using SAS 6, can be longer for SAS 8 or later. If the user does not want SAS data sets created enter ' '. Remember, the results tables will still be created for the .out file.

Examples of using these arguments with the TEST data set are listed below.

* Executes CAHPS macro with global rating scale variable, no adjusters and combining strata.

```
%cahps(var = q38,
 vartype = 2,
 name  = Rating Health Plan,
 adjuster = ,
 adultkid = 3,
 visits = 1,
 wgtdata = 2,
 dataset = test,
 outname = rplan ) ;
```

* Executes CAHPS macro with "How Often" composite variables and the recode option = yes, item weighting option = yes and 2 adjusters;

```
%cahps(var = q15 q17 q19 q24,
 vartype = 3,
 recode = 1,
 name  = Getting Care Quickly,
 adjuster = age ghr,
 impute = 1,
 adultkid = 3,
 wgtdata = 2,
 dataset = test,
 outname = quick ) ;
```

* Executes CAHPS macro with global rating scale variable, age and ghr adjusters, combining strata, and smoothing variances. Note: smoothing = 25 is as an example. For a detailed explanation of how this value is determined, see the section called ***Explanation of Statistical Calculations*.**

```
%cahps(var = q38,
 vartype = 2,
 name  = Rating Health Plan,
 adjuster = age ghr,
 adultkid = 3,
 visits = 1,
 wgtdata = 2,
 smoothing = 25,
 dataset = test,
 outname = rplan ) ;
```

Interpreting the Results

The CAHPS Analysis Program prints the results of the analyses performed for each composite and global rating. The program produces entity-level estimates of missing data for the analysis items and case-mix adjusters, calculates the percentage of responses in each category, compares performance of entities on the reporting item, and flags entities with fewer than 100 responses. If adjusters are used, then the coefficients and the regression analysis are produced for each adjuster item. Examples of results tables from the test data set for global rating scales are reviewed below.

Please note that the results tables are also output to SAS data sets. These data sets implement the following naming conventions where &OUTNAME is the text assigned by the user to the variable "outname" in the CAHPS macro call.

Percentage Items Missing:	P_&OUTNAME
Percentage of Each Response:	N_&OUTNAME
(for Global Rating* aggregates to 0-6, 7-8, and 9-10, or 0-7, 8-9, 10 for "How Often" Scale aggregates to 1-2, 3, and 4)	
Regression Coefficients:	C_&OUTNAME
R-Squared Values	R2&OUTNAME
Residual Values (only if KP_RESID = 1)	Y_&OUTNAME
Overall Statistics for All Entities:	OA&OUTNAME
Star Ratings for All Entities:	SA&OUTNAME
Plans dropped by macro with only 0 or 1 record	DP&OUTNAME

* See the FAQs on the CAHPS Web site (https://www.cahps.ahrq.gov) to learn more about the cutpoints for this scale.

If the stratified weighting option = 2, then the following data sets will be created for each unstratified entity.

Percentage Items Missing:	PW&OUTNAME
Percentage of Each Response:	NW&OUTNAME
Overall Statistics for All Entities:	OW&OUTNAME
Star Ratings for All Entities	SW&OUTNAME

If the keep permanent data sets for case-mix adjusted frequencies option = 1 and the stratified weighting option = 1, (no stratified weighting), then the following data sets will be created.

Overall statistics for all entities for first bar/frequency:	F1&OUTNAME
Star rating details for all entities for first bar/frequency:	B1&OUTNAME
Overall statistics for all entities for second bar/frequency:	F2&OUTNAME
Star rating details for all entities for second bar/frequency:	B2&OUTNAME
Overall statistics for all entities for third bar/frequency (not for dichotomous variables):	F3&OUTNAME
Star rating details for all entities for third bar/frequency (not for dichotomous variables):	B3&OUTNAME

If the keep permanent data sets for case-mix adjusted frequencies option = 1 and the stratified weighting option = 2 (do stratified weighting), then the following additional data sets will be created.

Overall statistics for unstratified data for first bar/frequency:	FA&OUTNAME
Star rating details for unstratified data for first bar/frequency:	BA&OUTNAME
Overall statistics for unstratified data for second bar/frequency:	FB&OUTNAME
Star rating details for unstratified data for second bar/frequency:	BB&OUTNAME
Overall statistics for unstratified data for third bar/frequency (not for dichotomous variables):	FC&OUTNAME
Star rating details for unstratified data for third bar/frequency (not for dichotomous variables):	BC&OUTNAME

For a detailed description of the computation and statistical analyses used to develop these results, see the section ***Explanation of Statistical Calculations***. At this point, all data elements have been collected to perform the testing on the hypothesis.

Warnings and Parameter Info

The following page of the SAS text output shows each of the parameter settings for the Analysis Program. You can use this to identify items you may want to consider when interpreting the results produced by the program, such as entities with fewer than 100 responses to an item after considering missing adjusters and analysis items. No SAS data set is produced that contains all this information.

```
     Rating Scale (0 - 10): Rate Plan
Analysis = ADULTS ONLY - Visits = COMBINE LOW AND HIGH USERS

*----------------------------------------------*
  CAHPS SAS Analysis Program Version 4.1
  Report run on 20 May 2011 at 14:04:45
*----------------------------------------------*


   **********   WARNING NOTE   **********
      PLANS WITH FEWER THAN 100 CASES
--------------------------------------------

  Plan ID 2 - HMO_B_RURAL  - 95 Cases

  Plan ID 4 - HMO_C_RURAL  - 68 Cases


--------------------------------------------


 The Variable Item             = q38
 The Variable Type             = 2
 The 2 Adjuster Variables      = q40 q39

 Global Case Mix Model
 Global Centering of Means


 The RECODE   parameter        = 0
 The MIN_RESP parameter        = 0
 The MAX_RESP parameter        = 10
 The NAME     parameter        = Rating Health Plan
 The ADJ_BARS parameter        = 1
 The BAR_STAT parameter        = 0
 The IMPUTE   parameter        = 0
 The EVEN_WGT parameter        = 1
 The KP_RESID parameter        = 0
 The ADULTKID parameter        = 3
 The VISITS   parameter        = 1
 The PVALUE   parameter        = 0.05
 The CHANGE   parameter        = 0
 The MEANDIFF parameter        = 0
 The WGTDATA  parameter        = 2
 The WGTRESP  parameter        =
 The WGTMEAN  parameter        =
 The WGTPLAN  parameter        = 0
 The ID_RESP  parameter        =
 The SUBSET   parameter        = 1
 The SPLITFLG parameter        = 0
 The data set used             = test
 The OUTREGRE parameter        = 0
 The output data set suffix    = rplan
```

Percent of Items Missing by Health Plan

```
  Rating Scale (0 - 10): Rating Health Plan
Analysis = ADULTS ONLY - Visits = COMBINE LOW AND HIGH USERS
   PERCENT ITEMS MISSING BY HEALTH PLAN
```

Health Plan	Total # of Respondents	Global Rating of Plan	Age Range	General Health Rating
HMO_A_URBAN	345	30.43%	1.16%	0.58%
HMO_B_RURAL	134	29.10%	1.49%	0.75%
HMO_B_URBAN	530	24.15%	1.70%	0.94%
HMO_C_RURAL	90	23.33%	0.00%	2.22%
HMO_C_URBAN	874	29.18%	0.69%	0.92%

```
Report run on 20 May 2011 at 14:04:45
CAHPS SAS Analysis Program Version 4.1
Data Set out.p_rplan
```

Definition of column headings	
Column	**Description**
Health plan	Health plan names based on the variable PLAN recoded from PLANID.
Total number of respondents	Total number of respondents in the data set by health plan. For illustration purposes, an entity with less than 100 respondents has been left in the sample output.
Global rating of plan (Label for the variable item Q38)	Percent of cases or records with each item response missing. At least two items must have a valid response to be included in the computation. The only answers counted as nonmissing by the SAS program for the global rating scales are the responses with integer values from 0 to 10. All other values are converted to the SAS missing value '.'. Computation of percent missing is based on the total number of patients in each plan.
General health rating (Label for the adjuster Q50 – GHR)	Answers counted as nonmissing by the SAS program are the responses excellent (coded 1), very good (coded 2), good (coded 3), fair (coded 4), and poor (coded 5). Computation of percent missing is based on the total number of patients in each plan.
Age of adult (Label for the adjuster Q51 - AGE)	Answers for adult surveys counted as nonmissing by the SAS program are the responses 18-24 (coded 1), 25-34 (coded 2), 35-44 (coded 3), 45-54 (coded 4), 55-64 (coded 5), 65-74 (coded 6), and 75+ (coded 7). Answers for child surveys counted as nonmissing by the SAS program are the responses <1 (coded 0), 1-3 (coded 1), 4-7 (coded 2), 8-12 (coded 3), and 13-17 (coded 4). Computation of percent missing is based on the total number of patients in each plan.

Percent of Composite Responses by Category[2]

```
    Rating Scale (0 - 10): Rate Plan
Analysis = ADULTS ONLY - Visits = COMBINE LOW AND HIGH USERS
   PERCENT RESPONSE TYPE - NO IMPUTATIONS
```

Health Plan	Total # of Respondents	Number of Respondents Analyzed	% Rating 0 - 6	% Rating 7 - 8	% Rating 9 -10
HMO_A_URBAN	345	238	17.23%	37.39%	45.38%
HMO_B_RURAL	134	95	21.05%	38.95%	40.00%
HMO_B_URBAN	530	393	15.78%	34.35%	49.87%
HMO_C_RURAL	90	68	16.18%	33.82%	50.00%
HMO_C_URBAN	874	613	15.33%	38.99%	45.68%

Health Plan	Adjusted Bar 1	Adjusted Bar 2	Adjusted Bar 3
HMO_A_URBAN	17.81%	37.92%	44.27%
HMO_B_RURAL	20.84%	39.38%	39.78%
HMO_B_URBAN	15.92%	34.22%	49.87%
HMO_C_RURAL	15.51%	32.94%	51.55%
HMO_C_URBAN	15.49%	39.04%	45.47%

```
Report run on 20 May 2011 at 14:04:45
CAHPS SAS Analysis Program Version 4.1
Data Set out.n_rplan
```

Definition of column headings	
Column	**Description**
Health plan	Health plan names based on the variable PLAN recoded from PLANID.
Total number of respondents	Total number of respondents in the data set by health plan.
Number of respondents analyzed	Number of respondents with nonmissing values (as defined in Percent of Missing Items on page 44).
Percent Rating 0 – 6	Percent of Rating item with response values from 0 to 6, by health plan.
Percent Rating 7 – 8	Percent of Rating item with response values equal to 7 or 8, by health plan.
Percent Rating 9 – 10	Percent of Rating item with response values equal to 9 or 10, by health plan.
Adjusted Bar 1	Case-mix adjusted frequencies for bar 1, Percent of Rating item with response values from 0 to 6, by health plan.
Adjusted Bar 2	Case-mix adjusted frequencies for bar 2, Percent of Rating item with response values equal to 7 or 8, by health plan.
Adjusted Bar 3	Case-mix adjusted frequencies for bar 3, Percent of Rating item with response values equal to 9 or 10, by health plan.

[2] For a detailed explanation of how these calculations were performed, see the section called ***Explanation of Statistical Calculations***.

Case-mix Adjuster Regression Coefficients

```
     Rating Scale (0 - 10): Rate Plan
Analysis = ADULTS ONLY - Visits = COMBINE LOW AND HIGH USERS
  REGRESSION COEFFICIENTS FOR ADJUSTER VARIABLES

          Variable    Subset
   Name    split        Name    Q38

  Q40         0      GLOBAL          0.3113
  Q39         0      GLOBAL         -0.4191


      Report run on 20 May 2011 at 13:32:32
     CAHPS SAS Analysis Program Version 4.1
              Data Set out.c_rplan
```

Definition of column headings	
Column	**Description**
Variable name	Q39 = General health rating 1 – Excellent 2 – Very good 3 – Good 4 – Fair 5 – Poor Q40 = Age 1 - 18-24 2 - 25-34 3 - 35-44 4 - 45-54 5 - 55-64 6 - 65-74 7 - 75+
Split	If the SPLITFLG parameter equals 0, then there is no split for the case-mix regression and split is 0. If SPLITFLG equals 1, then the two regressions are run, one for each split, where split equals 0 and 1.
Subset name	If subsetting is used for case-mix adjustment, the subset name or code is found in this column. Otherwise it defaults to GLOBAL; that is, it used all records in the case-mix adjustment regression.
Q38	Regression coefficients for the Global Rating of Plan variable.

R-Squared Values for Dependent Variables

```
   Rating Scale (0 - 10): Rate Plan
Analysis = ADULTS ONLY - Visits = COMBINE LOW AND HIGH USERS
  R-SQUARED VALUES for DEPENDENT VARIABLES

          Subset    Dependent                        Adjusted
split      Name     variable        R-squared       r-squared

   0       GLOBAL       Q38             0.0768          0.0755

     Report run on 20 May 2011 at 13:32:32
      CAHPS SAS Analysis Program Version 4.1
               Data Set out.r2rplan
```

Definition of column headings	
Column	**Description**
Split	If the SPLITFLG parameter equals 0, then there is no split for the case-mix regression and split is 0. If SPLITFLG equals 1, then the two regressions are run, one for each split, where split equals 0 and 1.
Subset Name	If subsetting is used for case-mix adjustment, the subset name or code is found in this column. Otherwise it defaults to GLOBAL; that is, it used all records in the case-mix adjustment regression.
Dependent Variable	Global Rating of Plan variable.
R-squared	The R-squared value from the regression for the dependent variable.
Adjusted R-squared	The Adjusted R-squared value from the regression for the dependent variable.

Overall Statistical Test Results

```
   Rating Scale (0 - 10): Rate Plan
Analysis = ADULTS ONLY - Visits = COMBINE LOW AND HIGH USERS
 P-Value For Contrast = 0.05 - Change > 0 - Meandiff > 0
    Overall Statistics from t-test
     Ho: Plan Means All Equal

Subset     Overall
Name        Mean         DFR           DFE    F-Statistic     P-Value

GLOBAL     7.9671          4         1,400       1.4508       0.2150


      Report run on 20 May 2011 at 13:32:32
      CAHPS SAS Analysis Program Version 4.1
               Data Set out.oarplan
```

Definition of column headings	
Column	**Description**
Subset Name	If subsetting is used for case-mix adjustment, the subset name or code is found in this column. Otherwise it defaults to GLOBAL; that is, it used all records in the case-mix adjustment regression.
Overall Mean	The mean of all the plan means.
DFR	The numerator degrees of freedom.
DFE	The denominator degrees of freedom.
F-Statistic	The results of the F-test on the null hypothesis.
P-Value	One minus the probability of the F distribution.

NOTE: "Ho" presents a global test of the null hypothesis that all plans have the same adjusted mean rating.

Statistical Test Performance by Health Plan

Rating Scale (0 - 10): Rating Health Plan
Analysis = ADULTS ONLY - Visits = COMBINE LOW AND HIGH USERS
P-Value For Contrast = 0.05 - Change > 0 - Meandiff > 0
ALL PLANS

Plan Name	Total # of Respondents	# of Respondents Analyzed	Unweighted Unadjusted Plan Mean	Weighted Unadjusted Plan Mean	Adjusted Plan Mean	Plan Diff. From Overall Mean
HMO_A_URBAN	345	238	7.9622	7.9622	7.9189	-0.0482
HMO_B_RURAL	134	95	7.7053	7.7053	7.7120	-0.2551
HMO_B_URBAN	530	393	8.1501	8.1501	8.1434	0.1763
HMO_C_RURAL	90	68	7.9559	7.9559	8.0093	0.0422
HMO_C_URBAN	874	613	8.0620	8.0620	8.0519	0.0848

Plan Name	Std Error of Difference	+/- 95% Conf. Limit of Diff.	Variance of the Mean	Variance of the Mean - old	Rating	Plan Weight
HMO_A_URBAN	0.1356	0.2657	0.0125	0.0125	**	1.00
HMO_B_RURAL	0.1982	0.3884	0.0348	0.0360	**	1.00
HMO_B_URBAN	0.1155	0.2263	0.0071	0.0071	**	1.00
HMO_C_RURAL	0.2265	0.4440	0.0476	0.0497	**	1.00
HMO_C_URBAN	0.1042	0.2041	0.0045	0.0045	**	1.00

Report run on 20 May 2011 at 13:32:32
CAHPS SAS Analysis Program Version 4.1
Data Set out.sarplan

Definition of column headings	
Column	**Description**
Plan Name	Health plan names based on the variable PLAN recoded from PLANID.
Total Number of Respondents	Total number of patients in the data set by health plan.
Number of Respondents Analyzed	Number of respondents with nonmissing values (as defined above) for the composite.
Unadjusted Plan Mean	Average health plan composite not adjusted for age or health status.
Adjusted Plan Mean	Average health plan composite adjusted for age and health status.
Plan Diff. from Overall Mean	"Overall Mean" minus "Adjusted Plan Mean."
Std Error of Difference	Standard error of "Plan Difference From Mean."
Variance of the Mean	Variance of the plan means.
Rating	Star rating of plan performance for the global plan rating based on a comparison of plan's "Adjusted Plan Mean" to "Overall Mean" Identifies statistically meaningful differences. * = Health Plan was significantly below average ** = Health plan was not significantly above or below average *** = Health plan was significantly above average

Small Data Set Example

This section uses a small data set with ten records, two entities, two questions, two adjusters and one weight variable to walk through an example of what happens to the data set as it moves through the CAHPS macro. The periods (.) in the table below represent missing values. The observation numbers are not a part of the data set; they are used only for reference purposes later. The sample tables after this one use the shorter column headings (Obs, Plan, Q1, Q2, A1, A2).

Please note that a variable PLAN refers to your unit of analysis. Variables in tables below are used inside of the macro.

SAS data set SMALLTST						
Obs Observation	Plan	Q1 Question 1	Q2 Question 2	A1 Adj 1	A2 Adj 2	Weight
1	A	2	4	1	1	40
2	A	3	.	2	2	50
3	A	4	2	3	.	6
4	A	4	3	.	.	8
5	A	3	3	2	3	10
6	B	3	3	2	3	3
7	B	.	.	4	5	5
8	B	2	2	5	4	3
9	B	3	2	6	3	5
10	B	7	3	3	3	3

The plan detail file is created by the macro and looks like the data set below. This data set is used by the macro to identify the plans it needs to analyze and create sequential plan numbers for use within the macro. Plan A is assigned the value of 1 and Plan B is assigned the value of 2. The macro needs the numeric values to perform looping functions at various points. In the "ALLCASES" section of the macro, it merges this information with the SMALLTST data set. In this example, we do not perform any strata weighting or subsetting of the data, so the values for these fields are set to 1. For the remainder of this example, the plan names A and B are used and the internal macro identifications are ignored.

Plan details data				
Observation	Original plan	New plan	Population size (strata weight)	Subsetting code
1	A	A	1	1
2	B	B	1	1

This example follows two paths for the analysis of the composite measure Q1 and Q2. One uses no adjuster variables, **Run 1**, and the other uses the two adjusters, A1 and A2, without imputation of missing values of the adjuster's mean within plan, **Run 2**. The macro cleans (making sure the values are within the valid range for the given variable type) the items being analyzed, Q1 and Q2. In the macro call they were indicated as being a type 3 variable, which means the response values must be a 1, 2, 3, or 4. Any other response value is set to missing. In our small data set (observation 10), Q1 has a value of 7 so it is set to missing; all other values are fine. The adjuster values are not cleaned in the macro, so all values are accepted.

The first step in the macro that begins to prepare the data for the reports is the "USABLE" section of the macro. This checks for missing values in each observation and determines whether to keep the record based on the macro arguments. The results may differ depending on whether adjusters are used and whether missing adjusters get an imputed mean value. The observations that are dropped for Run 1 and Run 2 after the "USABLE" section are noted as follows:

Obs	Plan	Q1	Q2	A1	A2	Run 1 - No Adj	Run 2 - With Adj
1	A	2	4	1	1		
2	A	3	.	2	2		
3	A	4	2	3	.		A2 Missing
4	A	4	3	.	.		A1 & A2 Missing
5	A	3	3	2	3		
6	B	3	3	2	3		
7	B	.	.	4	5	Q1 & Q2 Missing	Q1 & Q2 Missing
8	B	2	2	5	4		
9	B	3	2	6	3		
10	B	.	3	3	3		

The next few sections of the macro use the records retained from the "USABLE" section, nine records for Run 1 and seven records for Run 2. These sections simply report and summarize that data for low number of respondents, percent missing for each variable, and the percent breakdown of the response categories. The next step is to standardize each analysis variable in the data to a mean of zero and perform the case-mix adjustment if there are adjusters, as in Run 2.

Next we must determine the case-mix adjusters for each plan and the residuals for each item to obtain the adjusted means and calculate the variance for each plan. In Run 1, we are not doing any case-mix adjustments, so the adjustment to the means is zero and the residuals are the standardized values in the above table. In Run 2, there are adjusters, so the macro performs the regression necessary to get the adjustments for the means and the residuals.

Adjustments (on mean = 0)		
Plan	Run 1	Run 2
A	0.00	-0.25
B	0.00	0.25

		Residuals			
		Run 1		Run 2	
Obs	Plan	Q1	Q2	Q1	Q2
1	A	-1.20	1.00	-0.67	0.50
2	A	-0.20	.	0.33	.
3	A	0.80	-1.00	NA	NA
4	A	0.80	0.00	NA	NA
5	A	-0.20	0.00	0.33	-0.50
6	B	0.33	0.50	0.33	0.50
7	B	NA	NA	NA	NA
8	B	-0.67	-0.50	-0.67	-0.50
9	B	0.33	-0.50	0.33	-0.50
10	B	.	0.50	.	0.50

The next step is to multiply each residual by the item's equal weight, 1/number items in the composite (0.50) and divide by the total number of valid responses within plan and item in the composite. Then we can sum the results of the weighted residuals for each composite record.

Obs	Plan	Run 1	Run 2
1	A	0.005	-0.09
2	A	-0.020	0.06
3	A	-0.045	NA
4	A	0.080	NA
5	A	-0.020	0.03
6	B	0.118	0.04
7	B	NA	NA
8	B	-0.174	-0.12
9	B	-0.007	0.06
10	B	0.063	-0.02

Next we sum the squared composite weighted residuals by plan.

Plan	Run 1	Run 2
A	0.009	0.01
B	0.048	0.02

The final step to prepare for the statistical test is to calculate the variance within the plan using the above results by plan, multiplied by the number of usable records, divided by the number of usable records minus one.

Plan	Run 1	Run 2
A	0.012	0.017
B	0.064	0.027

At this point, all data elements have been collected to perform the testing on the hypothesis.

Explanation of Statistical Calculations

The purpose of this section is to describe the CAHPS macros in sufficient detail so that a statistically sophisticated reader can understand what analyses have been applied and some of the essential details of the implementation.

Note: In this section, a plan represents an entity (e.g., health plans, provider groups, individual physicians).

General Description

The CAHPS macros are designed to carry out a series of standard analyses on cleaned CAHPS data sets.

Inputs to the macro. The input to the macro is a SAS data set with one observation per survey respondent. The data set may contain only child responses, only adult responses, or a mixture of the two. (If there is a mixture, an option must be selected to indicate which group(s) are included in the analysis.) See ***Step 2, Modifying CONTROL.SAS*** for more details on the preparation of the files.

Outputs and statistical tests. The macro produces the following printed outputs for each summary scale (0-10 single item scale) and composite evaluation item:

1. A summary of warnings and parameter information.
2. A summary of missing data (item nonresponse) rates by item and plan.
3. A summary table of responses by item and plan for each category (with "never" and "sometimes" combined for "How Often" Scale and 0-6, 7-8 and 9-10 combined for the Global Rating Scale).
4. If adjusters are used, a printout of the regression for each adjuster, a table of the adjuster coefficients and the dependent variable R-squared value.
5. Overall results of a summary hypothesis test indicating the strength of the evidence that the plan means are not all identical on the given item or composite.
6. A table summarizing the statistical analysis, which includes unadjusted and adjusted plan means (percent "yes" for "yes/no" items), differences from the overall mean and standard error of those differences, and star-ratings (one to three stars) indicating statistically significant differences.

Data Subsetting and Checking

Subsetting. If only adults or only children are being analyzed but the data file combines both groups, the appropriate records are selected. High users can be selected by specifying the appropriate macro argument.

Response option checking. Item responses are checked to make sure that they conform to the response options for that variable item. All other responses are converted to missing values. After this recoding operation, the percent of missing items by plan is calculated and printed for each item in the analysis.

Number of responses. The number of cases is calculated for each entity. If the number of cases in the analysis for an entity falls below a cutoff value (100 cases), a warning message is printed. This is only a warning and does not affect any further analyses.

Weighting Algorithm for Composites

Once the SAS program groups the questionnaire items, it then computes means for each entity for each composite and global rating. Because the composites include more than one item, a more elaborate computation is required to develop the mean. CAHPS uses item weights to compute the means of the composites for each entity. Two methods are available for computation of the item weights. First, the CAHPS macro now includes an option to use equal item weighting in the composite, even_wgt = 1, where the item weight equals one divided by the total number of items. So if equal weighting was chosen and there were four items in the composite, the item weight is 1/4 = 0.25 for each item. By default, if even_wgt is not specified in the macro argument, then the composite uses even weighting. An advantage of this approach is that the relative weights of the items in the composite are consistent among survey administrations. Furthermore, survey sponsors may regard each item as equally important even if some are answered more frequently than others. A disadvantage of this option is a possible loss of statistical precision if an item with few responses is combined, equally weighted, with an item with many responses.

The CAHPS 4.1 macro offers some options that solve this problem by downweighting of low-response items.

The first modification is motivated by the fact that responses to different items in the same composite often have different mean values for a variety of reasons, including how frequently problems arise in different kinds of interactions and services and how the questions are worded. If the items are weighted the same way for every unit to calculate the composite, the effect of these unequal means across units is minimal. However, if items are not weighted equally, this could give rise to variations unrelated to variations in quality. Thus, we first modify calculation of weighted composites to minimize the impact of such differences in item means on expected scores. To explain the need for this modification, suppose y_i is the mean score for item i at a given unit, and μ_i is the mean score for item i across all units. With weights w_i that sum to 1, the composite score is $\sum_i w_i y_i$ for a specific plan, and if that plan is at the average on all measures, its score is $\sum_i w_i \mu_i$. If the overall means μ_i differ, this last expression will depend on w_i; in other words even two plans that are average on every measure will receive different composite scores if the composites are calculated with different weights.

To remove this dependence, we center the scores at their means before combining them. Suppose now that w_i represents the weight for item i at a particular unit, and w_{0i} represents some standard weights common to the entire report. Now define a composite score as $\left[\sum_i w_i (y_i - \mu_i)\right] + \left[\sum_i w_{0i} \mu_i\right]$. Any unit that is average ($y_i = \mu_i$) on every item

will receive the same composite score $\sum_i w_{0i}\mu_i$ regardless of the weights w_i, so bias due purely to weighting is removed even if different units are scored with different weights. Note that the second term of this composite score expression is the same for every unit; it is included only to bring the average back to an interpretable level as an average score of overall means.

Given this modification, we can now consider modifying item weights for different units. The main requirement is that the weight must be zero (w_i=0) when there are no responses for item i; we also want the weights to be equal (or at least to approach equality) when there is "adequate" sample for every item.

One simple weighting mechanism meeting these requirements follows:

- Set $w_{i0}=1/I$, $i=1, \ldots, I$, where I is the number of items in the composite.
- Choose a cutoff number of observations K; weights will not be modified for items with at least K observations.
- Define unit-specific weights $w_i = \min(n_i, K) \Big/ \sum_{i'=1,\ldots,I} \min(n_{i'}, K)$, where n_i is the number of responses from the unit for item I, and $\min(n_i, K)$ is the lesser of n_i and K.
- Calculate composite scores as described above.

This procedure has the following desirable properties:

- For each unit, all items with at least K responses are given equal weight. Consequently there is no modification to equal item weighting for units with large samples.
- Items with no responses in a given unit are given no weight, so the composite score can still be calculated.
- Items with low numbers of responses ($<K$) are given reduced weight so their effect on variance is mitigated.
- The criterion for determining whether an item will be downweighted is very simple to describe.

The procedure can easily be modified for unequal baseline weights w_{i0}.

Examples

The following table illustrates the calculation of item weights for various scenarios in a composite with three items, assuming that the target minimum sample size K=20.

Sample sizes n_i	min(n_i,K)	Calculation of weights w_i	Weights w_i simplified	Interpretation
60, 70, 80	20, 20, 20	20/60, 20/60, 20/60	1/3, 1/3, 1/3	Every item has adequate sample so equal weighting is OK
0, 22, 24	0, 20, 20	0/40, 20/40, 20/40	0, 1/2, 1/2	Item with no responses gets no weight
10, 22, 34	10, 20, 20	10/50, 20/50, 20/50	1/5, 2/5, 2/5	One item has low response and is downweighted
2, 3, 5	2, 3, 5	2/10, 3/10, 5/10	2/10, 3/10, 5/10	If all samples are small, weight each item proportional to number of responses to improve efficiency of estimation

The following illustrates the calculation of the "centered" weighted average in a unit in which one item of the composite has few responses (third line of table above), again assuming K=20.

Description	Symbol	Item 1	Item 2	Item 3
Baseline equal weighting	w_{i0}	1/3	1/3	1/3
Overall (all units) mean	μ_i	3.45	2.75	2.65
Mean in a specific unit	y_i	3.55	2.80	2.75
Sample sizes in that unit	n_i	10	22	34
Weights in that unit	w_i	1/5	2/5	2/5
Centered unit means	$y_i - \mu_i$	0.10	0.05	0.10

The baseline weighting is assumed to be equal for the three items. Thus the overall mean composite score is (3.45+2.75+2.65)/3 = 2.95.

Because at the specific unit of interest there are only 10 responses for Item 1, it is given half the weight of each of the other items. The weighted mean for the unit is then $(1/5)\times3.55 + (2/5)\times2.80 + (2/5)\times2.75 = 2.93$. Note that this is below the overall mean composite score, despite the fact that the unit is above the mean on each item, because the item that generally has a high score is downweighted.

To calculate the score by the proposed method, we first calculate the centered means (last line of table), which are all positive. Their weighted mean is $(1/5)\times0.10 + (2/5)\times0.05 + (2/5)\times0.10 = 0.08$. We then add this mean deviation from mean and add it to the overall

mean, 0.08 + 2.95=3.03, which is the reported score. This correctly reflects the superiority of this unit across all the items.

A second approach weights items unequally, in proportion to the number of respondents. To use item weighting by number of responses, the argument even_wgt = 0 must be entered into the macro call. The following brief explanation of the rationale for selecting this method of computing composite means is followed by a description of how the SAS program carries out the computation. Because there are differences among respondent experiences and, consequently, differences in the rates at which respondents use various services (expressed in skip patterns for various items), there are often different numbers of responses for the items that make up a composite. To reflect the differing numbers of responses by item, each item may be assigned a different weight in the composite score (an item weight). Items that receive a greater number of respondent answers count more toward the composite score. Thus, an item such as "how often did you need to see your personal doctor or nurse" generally has greater weight than "how often did you need to see a specialist" because more respondents are likely to need to see their personal doctors than a specialist.

One not very satisfactory way to create item weights for the composites is to develop a different set of weights for each entity being evaluated, proportional to the number of item responses at that entity. This means, however, that entity composite scores depend not only on the means on each item at the entity but also on the number of valid responses on each item available for that entity. Therefore, an entity could actually have a lower composite score just because it had a high number of responses to an item with generally low scores. Because such effects could distort comparisons among entities, we do not recommend this approach and it is not implemented in the macro.

Another option accounts for the different numbers of valid responses for each item within a composite but does so across all entities by standardizing the item weight. We consider this a valid approach to item weighting for the composites and have incorporated it into the Analysis Program. This approach prevents entities from faring worse or better just because they have fewer available valid responses because of skip patterns.

For composites, the results for the several items must be weighted together. For each composite, a set of data-determined item weights is calculated and used for all entities. The number of valid responses obtained for each item determines these weights.

The following is a formal description of the calculation of unequal item weights. Let

n_{ip} = number of valid responses obtained from entity p on item i,

$n_i = \sum_p n_{ip}$ = total number of valid responses obtained on item i,

$w_i = n_i \Big/ \left(\sum_i n_i \right)$ = weight for item i, and

F_{ipr} = fraction of responses for item i for entity p that fall into response category r.

The sums are over items i that are part of this composite.

Then the weighted fraction for response category r in entity p is

$$\sum_i w_i F_{ipr}$$

Case-Mix Adjustment

Another important policy decision regarding the analysis of CAHPS data is whether and how to adjust the data for different case-mix patterns. The CAHPS Development Team has studied this issue, and its recommendations for the adjustment procedure are incorporated into the SAS code, although the user may choose what variables to adjust for. The following explains the importance of the adjustment and how it is implemented.

When comparing entities on the basis of the ratings by individuals covered by those plans, it is important to adjust the data for patient characteristics known to be related to systematic biases in the way people respond to survey questions. This is called case-mix adjustment. It is automatically performed by the CAHPS Analysis Program if adjusters are specified. For example, if you know that people of a particular age group are reluctant to report problems and persons of that group are disproportionately represented in certain entities, it may be desirable to account for that when comparing data among entities. However, it is important to recognize that differences in patterns of responses may reflect real differences in quality of care as well as systematic biases. There is no way to separate these two types of differences based purely on statistical analysis of satisfaction data. The most popular methods for adjusting the data to account for differences in patient characteristics related to systematic biases are regression, stratification, and propensity score analysis, with regression being by far the most commonly used method.

Health status and age are two patient characteristics frequently found to be associated with patient reports about the quality of their medical care. People in worse health tend to report lower satisfaction and more problems with care than do people in better health. Older patients tend to report more satisfaction and fewer problems than do younger patients, although this association is usually not as strong as the one between health status and ratings.

Results from numerous CAHPS surveys in entities of all types confirm these general findings. For example, consumer ratings about health care were consistently higher for those in better health. Health status may be related to ratings of care because sicker persons are more likely to give negative ratings in general (response tendency), because some people are likely to give negative ratings about anything, including their health and the medical care they receive (correlated error), or because they get worse care, perhaps their greater needs create more opportunities for failure. There is the same ambiguity with the age association. However, regardless of the reason, it is misleading to rate an entity worse simply because of the kind of patients its treats.

Case-mix adjustment is intended to minimize the effects of differences between entities in background characteristics. The weighting algorithm for composites contributes one kind of case-mix adjustment, because it causes the items of a composite to be weighted together in the same proportions, regardless of differences in the response rates to the different items at different plans.

Another kind of case-mix adjustment applies to all of the items and potentially affects all reported results. This part of the adjustment uses a regression methodology, also called covariance adjustment. The user of the software chooses the adjuster variables. For illustrative purposes and because this has been a common choice in CAHPS so far, we assume in this discussion that the adjuster variables are age (AGE) and health status (GHR). If both adult and child records are in the same analysis, there are three additional adjuster variables: child indicator (CHILD), age x child interaction, and health status x child interaction. The inclusion of these three additional variables has the effect of fitting separate regression coefficients for the adjuster variables among children and adults.

If data are missing for an adjuster variable, the program either (at the option of the user) deletes the case or imputes the entity mean for that variable. The latter procedure avoids losing observations because of missing data; it is acceptable in this setting because, typically, both the size of the adjustment and the amount of missing data on adjusters are small.

Let y_{ipj} represent the response to item *i* of respondent *j* from entity *p* (after recoding, if any, has been performed). The model for adjustment of a single item *i* is of the form;

$$y_{ipj} = \beta'_i x_{ipj} + \mu_{ip} + \varepsilon_{ipj}$$

where β_i is a regression coefficient vector, x_{ipj} is a covariate vector consisting of two or five adjuster covariates (as described above), μ_{ip} is an intercept parameter for entity *p*, and ε_{ipj} is the error term. The estimates are given by the following equation:

$$\left(\hat{\beta}_i' \; \hat{\mu}_i'\right)' = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}_i$$

where $\mu_i = \left(\mu_{i1}, \mu_{i2}, \ldots \mu_{ip}\right)'$ is the vector of intercepts, $\mathbf{y}_i$ is the vector of responses and the covariate matrix is

$$\mathbf{X} = \left(\mathbf{X}_a \quad u_1 \; u_2 \; \ldots \; u_p\right)$$

where the columns of $\mathbf{X}_a$ are the vectors of values of each of the adjuster covariates, and u_p is a vector of indicators for membership in entity *p*, *p* = 1, 2,...*P*, with entries equal to 1 for respondents in entity *p* and 0 for others.

Finally, the estimated intercepts are shifted by a constant amount to force their mean to equal the mean of the unadjusted entity means $\bar{y}_{ip}$ (to make it easier to compare adjusted and unadjusted means), giving adjusted entity means

$$\hat{a}_{ip} = \hat{\mu}_{ip} + \left(1/P\right)\sum_p \bar{y}_{ip} - \left(1/P\right)\sum_p \hat{\mu}_{ip}$$

For single-item responses, these adjusted means are reported. For composites, the several adjusted entity means are combined using the weighted mean

$$\hat{a}_p = \sum_i w_i \hat{a}_{ip}$$

Casemix adjustments for entities that you do not want to affect the casemix model and adjustments

Sometimes case-mix adjustments may be required for an entity but for some reason it is not be desirable for the ratings from that entity to affect the estimated casemix coefficients or the recentering of entity scores. An example would be where the purpose of the implementation is to make comparisons among HMO plans, but data were also collected for non-HMO units and the sponsor wants to include them for comparison without affecting the HMO scores. A quick way to implement case-mix adjustment in this instance is to use the case-weighting option. Data from the entities

designated not to affect the model are retained in the sample but assigned very small weights (such as 0.0000001, or 0.0000001 times their sampling weights if the data are already weighted). The case-mix model is then applied as usual, using the weights. This trick works because (1) the weights for the designated entities are so small that the associated data have essentially no influence on the fitted model and (2) case-mix adjustment is performed in full irrespective of the weights.

Variance Estimation

An approach to variance estimation is used which is applicable to both the single-item reports and the composites. Variances are calculated for the mean for each entity, conditional on the coefficients for the adjuster variables. Conditionally these means are independent (ignoring the recentering constant that is added to make the mean of the adjusted means equal to that of the unadjusted means for presentation purposes). Conditioning on the regression coefficients is a standard procedure in variance estimation in the analysis of surveys (see Cochran, ***Sampling Techniques***, 1977, Chapter 7). It is not difficult to allow for the covariance of the adjusted means due to uncertainty about the regression coefficients in the case of single-item reports, but it is difficult to do this in a general way for the multi-item composites, when the pattern of missing data varies by item. In the interests of consistency, we use the same procedure for both classes of reports.

We first calculate residuals from the regression model for every item response,

$$z_{ipj} = y_{ipj} - \beta_i x_{pj}$$

where β_i is the regression coefficient vector for item *I* and y_{ipj} is the response to item *i* from person *j* in entity *p*. The adjusted mean $\hat{\mu}_{ip}$ for entity *p*, item *i*, is the mean (across nonmissing observations) of z_{ipj} If we replace z_{ipj} with 0 for all missing responses and define $r_{ipj} = 1$ if there is a nonmissing response and 0 otherwise, then we can write this as

$$\mu_{ip} = \left(\sum_j z_{ipj}\right) \Big/ \left(\sum_j r_{ipj}\right)$$

and the composite score for the entity is

$$\mu_p = \sum_i w_i \left(\sum_j z_{ipj}\right) \Big/ \left(\sum_j r_{ipj}\right)$$

Linearizing this expression by taking derivatives with respect to each of the sums $\sum_j z_{ipj}$ and $\sum_j r_{ipj}$, we obtain the following approximation:

$$\mu_p \approx \sum_j \left(1/n_{ip}\right) \sum_i w_i \left(z_{ipj} - r_{ipj} m_{ip}\right) = \sum_j d_{pj}$$

where $n_{ip} = \sum_j r_{ipj}$ is the number of responses to item *i* from entity *p*, d_{pj} is defined by the summand, and m_{ip} is the mean of z_{ipj} for the item *i* in entity *p*. We now apply the standard formula for the variance of an estimated sum,

$$\hat{V}_p = \hat{Var}\left(\hat{\mu}_p\right) = \left(n_p / \left(n_p - 1\right)\right) \sum_j d_{pj}^2$$

where n_p is the number of respondents from entity *p*. This gives an estimate of a variance of the composite score for entity *p*. If the composite consists of a single item, or if there is no item nonresponse, these results correspond to the standard variance formula.

Note that we do not apply any finite population corrections in this variance calculation. The finite population correction is appropriate if the object of our inference is the mean rating from the population of members or patients who are in entity *p* at the present time. Our concern, however, is with predicting the mean rating that would represent the experiences of a new set of subscribers or patients joining or seeking care at the entity, because we are attempting to give guidance to those who are considering anew their choice of insurance or treatment site. Conceptually, we regard the present members as a sample from a super-population of potential users of the entity.

Combination of Strata

Versions 1.5 and higher of the CAHPS software permit stratified analyses. For these analyses, the adjusted means μ_s and variances V_s are first calculated separately for each stratum within each entity. These calculations are identical to those described above except that stratum within entity should be substituted every time there is a reference to entity. Then means are calculated by combining all the stratum means within each entity. Suppose that the stratum weights are W_s, where $\sum_{s \in p} W_s = 1$.

(Here the sum is over all strata *s* within entity *p*.) Typically these stratum weights are defined as the fraction of the entity's total enrollment that falls in each stratum. The entity mean is calculated as $\hat{\mu}_p^* = \sum_s W_s \hat{\mu}_s$. The corresponding variance is calculated

as $\hat{V}_p^* = \sum_{s \in p} W_s^2 \hat{V}_s$. The quantities $\hat{\mu}_p^*$ and $\hat{V}_p^*$ are then used in the remaining solutions in place of $\hat{\mu}_p$ and $\hat{V}_p$.

Smoothing Variances

In some CAHPS implementations, entities with very small sample sizes were reported as significantly above or below average, although this clearly could not be established from their small amount of data. Further investigation revealed that this occurred because the estimated standard error was implausibly small (often 0, if all of a small number of respondents gave identical answers) which made the reported score appear to be highly precise and significantly different from the mean. This finding (typically occurring in items with low item response rates, such as those pertaining to a service used by only a few patients) raised concerns about the accuracy of direct sampling variance estimates when the number of responses is small.

To overcome these problems arising with small but nonzero sample sizes, we derived a method that combines the sample variance for an entity with an aggregated (pooled) variance estimate evaluated over all of the entities. The procedure is motivated by the following model for the sample variance $\hat{S}_i$ of an item score or individual-level composite score, computed for m_i respondents in the *i*th of n entities:

$$\hat{S}_i = S_0 + \delta_i + \varepsilon_i$$

where $S_0 + \delta_i$ is the population variance for entity *i* with $\delta_i \in [0, \tau_\delta^2]$, and $\varepsilon_i \in [0, 2S_i^2/(m_i = -1)]$ with $\sigma_i^2 = S_i$ the sample variance for entity *i*. The expression for the sampling variance of $\hat{S}_i$ is implied by the assumption that the sampling distribution of entity sample variances is chi-square; this is the usual asymptotic assumption and appears to hold approximately in practice. This can be viewed as a small area estimation problem in which $\hat{S}_i$ is a sufficient statistic for the within area measurements and the objective is to estimate $S_0 + \delta_i$. Under the normal-normal distributional assumption (normality of the error terms in the above equation), the posterior mean estimate of $S_0 + \delta_i$ is given by $w_i \hat{S}_i + (1 - w_i) S_0$, where $w_i = (1 + 2S_i^2 / \{(m_i - 1)\tau_\delta^2\})^{-1}$. This can be re-expressed as

$$\text{Est}(S_0 + \delta_i) = \frac{(m_i - 1)\hat{S}_i + 2(S_i^2 / \tau_\delta^2) S_0}{m_i - 1 + 2(S_i^2 / \tau_\delta^2)} \qquad (1)$$

We take (1) as an approximation to the true estimate, a convenient expression that is a linear combination of the entity-specific and pooled variances.

The observations whose sample variances are considered here are simply individual responses in an analysis of a single item, but they are Taylor-linearized combinations

across items (as described in the section on variance estimation) for analyses of multi-item composites.

We use the method of moments to estimate the between-entity variance τ_δ^2 of the variance. Because ε and δ are independent,

$$E\sum_i(\hat{S}_i-\bar{S})^2 = E\sum_i\delta_i^2 + E\sum_i\varepsilon_i^2$$
$$= (n-1)\tau_\delta^2 + 2\sum_i S_i^2/(m_i-1)$$

where $\bar{S} = \dfrac{\sum_i (m_i-1)\hat{S}_i}{\sum_i (m_i-1)}$ estimates S_0. The CAHPS macro output contains (in the variable VP) the value $\hat{V}_{\text{macro},i} = \hat{S}_i / m_i$, the squared standard error as opposed to the sample variance. Hence, the between-entity component of the variance of the variance is estimated by

$$\tau_\delta^2 = \left(\sum_i(\hat{S}_i-\bar{S})^2 - 2\sum_i S_i^2/(m_i-1)\right)/(n-1)$$

and the square of the coefficient of variation is given by

$$CV^2 = \tau_\delta^2/\bar{S}^2 .$$

The square of the coefficient of variation of the chi-square distribution is $CV^2 = 2/A$, where A is the degrees of freedom of the distribution (which can be thought of as the inverse of a prior weight). Therefore, it makes sense that we use $A = 2/CV^2$ as the weight of the pooled variance across the entities in the expression for the usual precision-weighted estimator of the posterior mean of the variance of an individual entity's ratings. Substituting into (1), we obtain

$$\hat{S}_{\text{smoothed},i} = \frac{A\bar{S} + (m_i-1)\hat{S}_i}{A+(m_i-1)} .$$

We express this in terms of sampling variances (using the relationship $\hat{V}_i = \hat{S}_i / m_i$) to obtain:

$$\hat{V}_{\text{smoothed},i} = \hat{S}_{\text{smoothed},i}/m_i = \frac{A\bar{S}/m_i + (m_i-1)\hat{V}_{\text{macro},i}}{A+(m_i-1)} .$$

This ensures that $\hat{V}_{\text{smoothed},i} \approx \bar{S}/m_i$ when m_i is small (implying little information about the variance) and $\hat{V}_{\text{smoothed},i} \approx \hat{V}_{\text{macro},i}$ when $m_i \to \infty$ (large amount of information for

estimating the variance) or when $A \to 0$ (that is, when the differences in sample variances across entities are very small). In general $\hat{V}_{\text{smoothed},i}$ lies between these two extremes, smoothing the variances for small entities a greater amount than the variance for larger entities whose own estimate of variance is more precise.

We have added an optional module to the CAHPS macro that allows for smoothing of variance estimates in this way, requiring the user to specify only the weighting factor A. Based on the 2010 Medicare CAHPS survey, the values obtained for A were $A = 25$, $A = 20$, and $A = 15$ for the rating, composite, and report items respectively. We recommend these (or larger) values to users who prefer to minimize the deviation from previous procedures while obtaining adequate protection from the unreasonable results obtained without smoothing of variance estimates.

Hypothesis Tests and Assignment of Final Ratings

Global *F*-test. The first test calculated is intended to determine whether there is evidence for differences among entity means. If this test does not find significant differences, it is not necessarily appropriate to report results by entity on the corresponding item or composite.

The weighted mean is calculated as

$$\hat{\mu} = \left(\sum_p \hat{\mu}_p / \hat{V}_p \right) \Big/ \left(\sum_p 1 / \hat{V}_p \right)$$

Then the *F*-statistic is calculated as

$$F = \left(1/(P-1)\right) \sum_p \left(\hat{\mu}_p - \hat{\mu}\right)^2 / \hat{V}_p$$

This statistic has an approximate *F* distribution with (*P-1*, *q*) degrees of freedom; we have found in simulations that $q = n/P$ (the average sample size per entity) makes the *F*-test at worst slightly conservative with typical sample sizes and response distributions. In other words, reported *p*-values from the test are slightly larger than they should be, so significant differences are less likely to be declared.

***t*-tests for entity differences from mean.** We compare each entity mean to the mean of the entity means using a *t*-test. The corresponding contrast is

$$\Delta_p = \hat{\mu}_p - (1/P) \sum_{p'} \hat{\mu}_{p'} = ((P-1)/P)\hat{\mu}_p - \sum_{p'}^{*} (1/P)\hat{\mu}_{p'}$$

where Σ^* represents a sum over all entities except entity *p*. Note that the last expression is simply (*P*-1)/*P* times the difference of $\hat{\mu}_p$ from the mean of all entities

except entity p; therefore the two formulations (mean vs. mean of all, or mean vs. mean of all others) are equivalent. The variance of Δ_p is

$$\hat{V}(\Delta_p) = [(P-1)/P]^2 \hat{V}_p + 1/P^2 \sum_{p'} \hat{V}_p$$

and the t-statistic is calculated as $\Delta_p / \hat{V}(\Delta_p)^{1/2}$, and referred to a t distribution with $(n_p - 1)$ degrees of freedom, which again is usually slightly conservative.

Reliability of CAHPS measures. It is often of interest to evaluate the precision with which CAHPS measures distinguish among the entities being compared in a given implementation. The reliability statistic R summarizes the fraction (on a 0 to 1 scale) of the variation among entity scores (based on samples) that is due to real variation in quality in the population At one extreme, R=0 means that there is no population variation across entities and all the observed variation is due to sampling variation, so the measure is essentially useless for distinguishing quality among entities. At the other extreme, R=1 means that all the entity scores are free of sampling error. Reliability will be high when there is good agreement among respondents in the same entity, large differences among entities, and large sample sizes.

There are two ways of computing reliability using the CAHPS macro results, both requiring additional analysis outside the current release of the macro. The first bases reliability of CAHPS measurements about entities on the F-statistic for testing differences among entities on an item or composite. The numerator of the F-statistic summarizes the amount of variation among the means for different entities on the measure and thus measures between-entity variation. The denominator summarizes the amount of random variation expected in these means due to sampling of individuals. If there were no real differences among entities, so that all the differences were due to random variations in the reports of the patients sampled for the survey, the F-statistic would be about 1. The greater the real differences among entities, relative to random variation, the larger the F-statistic is expected to be. A summary measure of reliability is obtained by the formula $R=1-1/F$. When F=1 (only random variation), R = 0 (no reliability), while for large F, R approaches 1 (best possible reliability). Because the CAHPS macro routinely outputs F, R is trivial to compute. In CAHPS, $R<0.70$ is commonly considered poor reliability, and R>0.90 is considered high.[3]

The above calculation of R pools information across all entities into a single survey-wide scalar summary for each item or composite. The number of respondents will vary across entities, however, giving them different sampling variances; hence an alternative is to estimate reliability for each entity. Another possible objective is to predict the reliability of measurements made on future entities given their numbers of respondents. These calculations use the adjusted entity mean and its associated

[3] Keller S, O'Malley AJ, Hays RD, Matthew RA, Zaslavsky AM, Hepner KA, Cleary PD. Methods Used to Streamline the CAHPS® Hospital Survey. Health Services Research, 2005, 40, 2057-2077. PMID: 16316438.

variance (squared standard error), which are standard CAHPS macro outputs. Let the entity mean, standard error, and total number of responders to the item or composite entity $i = 1, \ldots, D$ be denoted m_i, s_i and n_i respectively. Then use the following procedure:

1) Compute the total number of respondents across all entities: $N = \sum_{i=1}^{D} n_i$.
2) Compute the overall mean rating: $\bar{m} = N^{-1} \sum_{i=1}^{D} n_i m_i$
3) Compute the sample variance estimate for each entity: $v_i = n_i s_i^2$.
4) Compute the within-entity variance: $\bar{v} = N^{-1} \sum_{i=1}^{D} n_i v_i$.
5) Compute the between-entity variance:
$$b = \max\left\{ (N(1 - 1/D))^{-1} \sum_{i=1}^{D} n_i (m_i - \bar{m})^2 - \bar{v}, 0 \right\}.$$
6) Reliability for a specific entity is calculated as $R_i = b / (b + s_i^2)$
7) For projections for a future survey with *r* respondents per entity compute reliability as $R = b / (b + \bar{v} / r)$.

Assignment of star ratings. An "average" entity is assigned two stars. One or three stars are assigned on the basis of simultaneously satisfying a criterion of statistical significance and one of substantive significance (if specified). The difference of an entity from the mean is deemed to be statistically significant if the two-sided *p*-value of the *t*-test described above is smaller than a predetermined level. The comparison value for determination of substantive significance is a minimum difference determined as $\Delta_0 = K \min\left(\hat{\mu} - \mu_{\text{low}}, \mu_{\text{high}} - \hat{\mu}\right)$, where $\hat{\mu}$ is the overall mean, μ_{low} and μ_{high} are the lowest and highest possible values of the score and *K* is a fixed constant chosen by the user. A star is given or removed (relative to two stars) only if the difference is statistically significant and also $\left|\hat{\mu}_p - \hat{\mu}\right| > \Delta_0$. (If *K* = 0 then $\Delta_0 = 0$ and the second part of the criterion has no effect.)

Examining Sample Size Issues for CAHPS Surveys

To examine the effect of small sample sizes, the CAHPS Team looked at the data from the demonstration sites for the CAHPS Health Plan Survey 3.0. In each site the range of number of plans was 2 to 27, with a mean of 10 plans. The minimum sample size per reportable measure ranged from 17 to 418 (mean = 155) and the average sample size ranged from 46 to 468 (mean = 238). The percentage of two star plans per reportable measure ranged from 0 to 100 percent with an average of 71 percent.[4]

The observed site data were used to estimate power for different combinations of number of plans and sample size per plan. Table 3 provides effect sizes (difference

[4] Dichotomous items were excluded in these calculations because ***CAHPS Health Plan Survey 3.0*** does not include them.

between the mean of one plan and the mean of all other plans/SQRT(MSE) that can be detected with 80 percent power and alpha of 0.05 (two-tailed) for 2, 5 and 15 health plans with health plan sample sizes varying from 40 to 300. Note that the effect size that can be detected with a sample size per plan of 300 for two health plans is similar to the effect that can be detected for a sample size of 200 when there are 5 health plans (effect sizes of 0.23 and 0.22, respectively). Similarly, the effect size detectable is comparable for two plans with n = 100 per plan versus five plans with n = 60 per plan (0.40 and 0.41, respectively).

Table 3. Effect size detected with 80 percent power (alpha = 0.05) by number of plans and sample size per plan

Sample size per plan	2 plans	5 plans	15 plans
300	0.23	0.18	0.17
200	0.28	0.22	0.21
100	0.40	0.32	0.29
80	0.45	0.35	0.33
60	0.51	0.41	0.38
40	0.63	0.50	0.46

Table 4 provides information about the effect size detectable for one plan when the sample size of all plans except that one is fixed at 300. These effect sizes are similar to those reported in Table 5-1 with a few exceptions (lower left corner of tables), revealing how small sample size for one plan has a major impact on the power to detect a difference between the plan and the other plans.

Table 4. Effect size detected with 80 percent power (alpha = 0.05) by number of plans and sample size for one plan (n = 300 for all other plans)

Sample size for plan	2 plans	5 plans	15 plans
300	0.23	0.18	0.17
200	0.26	0.22	0.21
100	0.33	0.29	0.29
80	0.35	0.33	0.32
60	0.40	0.37	0.37
40	0.47	0.45	0.45

As an illustration of how effect sizes translate into CAHPS scale points, adult data from one of the sites were examined for the rating of specialist care. Sample sizes ranged from 32 to 104 per plan (mean = 67). Seven of the 10 plans received two stars, one plan received a single star, and two plans received three stars. The overall mean on the item was 8.30 and the smallest plan difference from the mean of other plans

that was statistically significant was 0.24 (standard error of difference = 0.12) for an entity with 57 completes. The estimated square root of MSE (SE_d/SQRT($1/n_1$ + $1/(N-n_1)$) is 0.83. Thus, the observed effect size for this difference was 0.28 (a relatively small effect). Assuming a similar MSE, we have 80 percent power (alpha = .05) to detect differences between one plan and the mean of the other plans of 0.33, 0.27 and 0.24 on the 0-10 scale for sites with 2, 5 and 15 plans, respectively, when the sample size of each plan is 100. Note that these differences correspond to 0.19, 0.16 and 0.14 on the CAHPS "meaningful differences" scale (proportion of the distance from the mean to the nearest extreme).

Implications of this Analysis

The power to detect a difference between one plan and the mean of the other plans depends upon the sample sizes for all plans (although the sample size of the plan in question has an especially important influence) and the number of plans. With the current CAHPS recommendation of a sample of 100 per health plan per reportable item, there is adequate power to detect an effect size of 0.29 (15 plans) to 0.40 (2 plans) if every plan has the minimum sample size. A similar effect size (0.38) can be detected for a sample size of 60 per plan if there are a total of 15 plans. It is important to note, however, that these power calculations pertain only to the determination of the number of stars an entity receives. The CAHPS bar charts provide an opportunity for pair-wise comparisons and the sample size requirements for a given power are therefore larger.

Appendix

Summary of Features Included in Each Version of the CAHPS Analysis Program

Version 1.0 of the CAHPS SAS Analysis Program offered the following features:

- An assessment of significance using practical and statistical (*p*-value) criteria;
- An option to analyze data based on outpatient utilization groupings;
- An option to analyze child and adult data together or separately;
- Comparisons of health plan performance; and
- Case-mix adjustments.

Version 1.5 of the CAHPS SAS Analysis Program added the following enhancements:

- **Weighting and stratification.** The SAS program performs the correct analyses for disproportionate stratified sampling designs. One way such designs might appear is when two plans that were surveyed separately have subsequently merged their operations into a single business entity, and their results will be reported as a single plan. They also may appear when the sponsor decides to collect additional surveys by using larger sample sizes for a certain subset of people (based on geographic area, gender, age groups, etc.) beyond what would appear there by proportionate allocation. To use this feature, the user must specify which strata are combined and the number of members in each stratum out of the entire population (the weights).

- **Plan name flexibility.** Plan identifiers for programming and output purposes are no longer required to be numeric. Text or numeric names are allowed to facilitate programming and interpretation of results.

- **Case-mix adjusters.** The program no longer requires two case-mix adjusters (age and health status) to be used in the analyses. The user can now specify an unlimited number of adjuster variables or choose not to adjust the data.

- **Substantive differences.** A new method of specifying an absolute difference that must be achieved before a difference is meaningful has been added to the program. While the previous method of determining a meaningful difference is still available, the user can now simply choose an absolute difference that must exist between means for a difference to be flagged as significant.

- **Results tables.** Version 1.5 has an additional feature that creates SAS data sets of the results tables the program produces. This allows users to perform additional analyses on the aggregate results or to create summary reports. Linear regression coefficients for the adjuster variables are now output as part of the results tables and reports.

- **Missing data for adjusters.** In the initial version of the Analysis Program, missing data for the case-mix adjustment variables was imputed at each item's health plan mean. Version 1.5 allows the user to specify whether or not the analysis is conducted with imputation for the adjuster variables.

Version 2.0 and 2.1 of the CAHPS SAS Analysis Program added the following enhancements and changes:

- The SAS code has been converted to **require only Base SAS and the SAS/STAT** module, eliminating the need for SAS/IML. If adjuster variables are excluded, then the REG procedure in the SAS/STAT module is not needed. The code has been modularized into macros to aid in maintaining the macro and understanding what the macro is doing.

- The macro now has two **additional ways in which to subset the data** being run through the Analysis Program without having to create separate calls of the Analysis Program. With SUBSET = 2, the Analysis Program runs the case-mix model on the entire data set but does the plan/unit comparisons at the subset levels specified in the fourth column of the plan detail file created by the user. With SUBSET = 3, the Analysis Program does both the case-mix and the plan/unit comparisons at the subset levels.

- **Data sets are now created for the output** of the case-mix and hypothesis test calculations. This allows for easy export to Excel or other programs for report generation.

- The **composites are no longer restricted to the "How Often" (1-4) question responses.** The variable type is indicated in the macro call and the macro runs a composite calculation if the number of variables is greater than one. This change was made to accommodate the need to create composites from questions with dichotomous and trichotomous variables. The program can now create composites using all variable types used in the survey

- The **weighting of the composite items** now has the option of doing **equal weighting** across items as well as **weighting based on the number of responses** in each item divided by the total number of responses in all items. The default option for the macro is to use the equal weighting.

- An option is available for **recoding the global rating scales** from 0 – 10 to 1 – 3 and the **"How Often" scales** from 1 – 4 to 1 – 3 using the new parameter RECODE. The primary rationale for the recoding into three categories is to make the data entering into the hypothesis tests entirely consistent with the information presented in the "Bar Graph" reports.

  A secondary rationale for recoding is that it may improve the statistical properties of the tests. On general statistical principles, it would not be surprising if the analysis of very skewed data were improved by a transformation that reduced the skewness. In the CAHPS survey, it is plausible that the difference between 0 and 2, both indicating strong dissatisfaction, carries with it less information than the difference between 8 and 10, reflecting average and maximum satisfaction, respectively. Therefore, combining categories at the low end of the scale may remove some meaningless variation from the data. Statistical improvement would be reflected in larger values of the *F*-statistic in the recoded data compared to the original data.

  The recoding is defined as:

Rating scale		How often scale	
Response value	**Recode**	**Response value**	**Recode**
Option 1:			
0 – 6	1	1 – 2	1
7 – 8	2	3	2
9 – 10	3	4	3
Option 2:			
0 – 7	1		
8 – 9	2		
10	3		

- A new parameter, **KP_RESID**, has been added to the macro call to allow **the residual values from the regression to be saved** as a permanent SAS data set. By default, these values are only saved temporarily while the macro is running.

Version 3.0-3.3 of the CAHPS SAS Analysis Program adds the following enhancements and changes:

- The **plan detail file**, plandtal.dat, and the filename statement that assigns **PLAN_DAT are now optional**. If the plan detail file does not exist, then the macro uses the PLAN variable in the dataset called by the CAHPS macro. If used, the plan detail file must have a unique record for each plan name or code. Only the first column is required; if the second column is missing, then the macro creates dummy values for the new plan name equivalent to the first column. If the third and fourth columns have missing values, then they are all set to the value of 1. Each column must be separated by spaces.

- The Analysis Program **now removes any plans that are to be analyzed that have only zero or one usable records.** These changes were made in the submacro USABLE. The plans that are dropped by the macro are saved in a permanent SAS data set labeled dp&outname.

- **The CHILD variable is now optional.** If it does not exist, then the macro creates the variable CHILD. If the ADULTKID parameter is set to 2, then the macro assumes all records in the analysis data set are child records and sets CHILD = 1, otherwise CHILD will be set to 0, indicating there are no child records. If there is a mix of child and adult records in the data set, the user must set up a variable named CHILD and set it equal to 1 for child records and some other value, usually 0 for adult records. Version 3.3 of the CAHPS macro corrects a logic error found in version 3.2 of the macro.

- **The EVEN_WGT parameter now can apply individual level weights to the composite items.** This third option is activated by setting EVEN_WGT=2 and uses the weight variable, referenced by WGTRESP.

- The **variance of the mean** variable, vp, has been added to the text output of the adjusted mean report.

- A **CAHPS version label** has been added to the permanent data sets to indicate which version of the CAHPS Analysis Program created the data set. The version number has also been added to the text output.

- Users can now **case-mix the triple-stacked bar frequencies,** using the **ADJ_BARS** parameter, and include both the non-case-mixed frequencies with the case-mixed frequencies in the final frequency output data set, n_*. For variables of type 5 (vartype = 5), these cannot have case-mixed bars since the frequencies for the response values are not aggregated into three bars. To make this work for nonstandard variable types, it is best to

do some recoding first to make the three desired ranges and then run the new variable through as a vartype = 4.

The following parameters have been added:

- The **parameter ID_RESP stores the original respondent ID value, if one exists, in the permanent data sets.** If there is a unique variable in the data set that identifies each respondent, then enter the variable name in this parameter. The macro carries it through the individual data sets and attaches it to the residual data set if KP_RESID = 1 so the data set can be easily linked to the original if needed. If no ID variable is entered, then the ID_RESP variable in the macro is set to '.z'. The variable will be a character and have a maximum of 50 characters.

- The parameter flag **OUTREGRE** indicates **whether or not the regression output should appear in the text output file.** If set to 0, the default, then the SAS printed output from the regressions in the case-mix procedure is not printed out into the output file. If set to 1, then the regression output appears.

- The parameter **WGTRESP** accepts the **variable name that contains the weights for individual respondents.** This weight is used in the case-mix adjustment regression procedure.

- The parameter **WGTMEAN** accepts a **variable that contains the weights to be applied to the means of the plans** before the case-mix adjustments are applied.

- The parameter **SPLITFLG** allows the data set to be **split into two groups for the purpose of centering the means differently and running two case-mix models through the macro.** This was done to deal with the Medicare Managed Care and Fee-for-Service analysis. By default, the parameter is 0 and is not used but, if set to 1, then the data set must contain a variable with the name SPLIT and must have the values of 0 and 1. Any record with a missing value is dropped from the analysis.

- The parameter **BAR_STAT stores the results of the case-mixed bars in permanent data** sets with the same format as the case-mixed survey question results. The new data sets created have the format B#&outname and F#&outname where the B* files hold the stars and statistics by plan and the F* files hold the overall means and statistics. The # has the values 1-3 for a normal macro run, where 1 = the first bar frequency, 2 = the second bar frequency, and 3 = the third bar frequency if it is not dichotomous. &outname is the value given in the macro call parameter OUTNAME. If the data are stratified and stratification weights are used by having the macro parameter WGTDATA = 2, up to six additional files are

created with # having the values A-C, where A = the first bar frequency of the combined strata, B = the second bar frequency of the combined strata, and C = the third bar frequency of the combined strata.

- Version 3.3 **corrects a logic error**, contained within version 3.2, that occurred when the parameter **SUBSET = 3**, which runs the macro multiple times based on the subsetting variable in the plan detail file referenced by the FILENAME PLANDTAL statement.

- The text output on the **Warnings and Parameter Info page contains more accurate information about the adjusters** when there are child interactions, when **ADULTKID = 1**. The number of adjusters will reflect the original adjuster variables times 2 plus 1, so if there are originally 2 adjusters, the total number of adjusters with child interactions will be 5, ADJ#1, ADJ#2, ADJ#1 * CHILD, ADJ#2 * CHILD, and CHILD.

- Added two **flag lines to the log file** that will indicate if the macro **finds the CHILD and PLAN variables** in the original analysis data set. If there is no child variable, the flag will indicate how the macro created a new CHILD variable.

Version 3.4 (May-June 2003) of the CAHPS SAS Analysis Program adds the following enhancements and changes:

- Added three **additional variables to the sa* data set and the output text** of the statistical tests. The **unweighted, unadjusted plan mean** was added to help clarify what the unadjusted mean actually is. Only when the wgtmean parameter is used will the unweighted, unadjusted mean be different from the weighted unadjusted mean. The other variable added is the **95% Confidence Limits for the Difference of the Mean**. This is computed as 1.96 * the standard error of the difference. When wgtplan = 1, then a third column containing the **summed weights for each plan** will also be added to the sa* data set, the b* dataset if frequency bars are to be stored (bar_stat = 1) and the output text.

- Added in the **weighted, unadjusted frequencies** to the frequency table n_* data set and the output text, when the frequency bars are also case mix adjusted.

- The **wgtmean** parameter purpose has been **expanded** to allow for the use of the **sum of the weights to the plan level to be used in the comparison of the plan means**. If a variable exists for the wgtmean parameter, then the individual record level weight is used to compute the weighted, unadjusted plan means. In addition, if the new parameter wgtplan = 1, then the sum of the individual weights to the plan level will be used in weighting the plan mean comparisons. The wgtplan parameter can have

the value of 0, default, or 1. When 0, the macro will use equal weights when comparing the plan means. When 1, and the wgtmean parameter has a variable listed, then the sum of the weights to the plan level will be used computing the overall and grand means which are used in the statistical comparisons of the plan means.

- Added checks on the DATASET parameter to make sure it exists or that the value in the DATASET parameter is a valid SAS data set. If there is an error, the macro will stop processing and print an error message to the log file.

- Added error checking on the merging of the plan detail file with the analysis dataset. If there are no records matching, then the macro will print out the frequencies of the unique PLAN values for both the plan detail file and the analysis data set to the output text file and also print out and error to the log file.

Version 3.5 (September 2005) of the CAHPS SAS Analysis Program added the following enhancements and changes:

- A disclaimer and copyright statement were added.

- If weights are being used for the individual or plan means, records with weights that are less than zero or missing are removed.

- When macro converts the numeric plan in allcases to character, it left justifies and trims trailing blanks.

- The macro checks that there are plans in all subcodes after the usable data set is made. If some subcodes have all missing plans, it recomputes how the subcodes are used in the looping in the star macro.

- The log comment for when child variable is not found in the original data set was changed.

- A bug was identified in the CAHPS 3.4b macro: Two lines that have length planname $ 20 when it should be $ 40 causing a merge problem with the N_* data sets. $ 20 was changed to $ 40.

Version 3.6 (April 2006) of the CAHPS SAS Analysis Program added the following enhancements and changes:

- This new version corrects an error in some previous versions affecting calculation of the variances for the comparison of an entity mean to the mean of all other plan means, when the plans were weighted. This error only affects analyses with parameter wgtplan=1 using CAHPS macro versions 3.4b (released May 2003) and 3.5 (released September 2005). By

default, the macro sets wgtplan=0 so the error does not affect unweighted plan analysis.

The error caused significance tests to be calculated incorrectly when determining whether an entity's mean was significantly above or below the average. This could cause some plans to be declared 1- or 3-star plans when they were respectively below or above average, but not by a statistically significant amount.

- (July 2006) Modified formula for special case of using only one plan unit and a division by zero error may occur. This case used to work in prior versions. Modified code for checking if SE may be missing to set T=0 in that case. Also, VO can now have a zero denominator, in the case where there is only one unit being analyzed, modified code to catch that error.

(3.6b as of June 2007) This modification to Version 3.6 puts the _wgtmean variable in the strata data step in order to address a problem with a missing line that was not keeping the _WGTPLAN variable in the data step that created wstemp. Because of the missing line, the use of wgtdata=2 for combining strata generated a SAS error.

Version 4.0 (September 2011) and 4.1 (April 2012) of the CAHPS SAS Analysis Program added the following enhancements and changes:

- One part of the code that creates plandtal data set (it is in usable macro program) was modified. This only affects when subset = 3.

- The calculation of weights for the composite items was modified. The sum of weights based on the number of responses from each item is used as the weight of the composite case. Also, the calculation of item weights for even_wgt = 1 was modified. For more details about how the weights are computed, please see the *Explanation of Statistical Calculation* section.

- A new warning note was added in the macro output (it is in . mkreport macro program). The note lists plan IDs when they have zero responses in measured items. A new option of assigning smoothing variances was added. Users can assign a weight parameter called *smoothing* on the variances as option. The default is *smoothing* = 0. This provides the original variances. If *smoothing* is greater than zero, the value that users input will be used as the weight for the variances. If *smoothing* is less than zero, the weight will be computed inside of the macro automatically. For more details about how that weight is computed inside of the macro, please see the *Explanation of Statistical Calculation* section.

- A SAS procedure PROC STANDARD was replaced with PROC STDIZE. The macro centers all adjusters before it runs regression procedure if adjusters are required. PROC STANDARD was not applicable when some adjusters contain only the same values. As a result, it did not standardize the value correctly. PROC STDIZE is able to handle the situation.

- (April 2012) Modified codes for computing adjusted composite means when composite even weight option (even_wgt = 1) is selected. The macro computes the weights for all entities regardless of whether they get dropped out for the analysis due to a lack of the sample size. In the prior version, this caused incorrect adjusted means when some entities did not make it to the final analysis. Also, the macro did not handle correctly on computing adjusted means when some entities have different weights from the even weights. Version 4.1 is able to handle the case and provide appropriate adjusted means.