Version 07 HHS-HCC Risk Adjustment Modeling "Statistical Analysis System (SAS)" Software Documentation for the 2022 Benefit Year
April 11, 2023[1]

Section 1343 of the Patient Protection and Affordable Care Act (ACA) provides for a permanent risk adjustment program. To protect against potential effects of adverse selection and help stabilize premiums in the individual and small group (including merged) markets, the risk adjustment program transfers funds from plans with relatively low-risk enrollees to plans with relatively high-risk enrollees. It generally applies to non-grandfathered individual and small group (including merged market) plans inside and outside Exchanges.

The HHS risk adjustment methodology was first described in the HHS Notice of Benefit and Payment Parameters for 2014 final rule (78 FR 15410), which was published in the *Federal Register* on March 11, 2013. Modifications to the HHS risk adjustment methodology for the 2022 benefit year are described in the HHS Notice of Benefit and Payment Parameters for 2022 final rule (86 FR 24140) (2022 Payment Notice final rule), which was published in the *Federal Register* on May 5, 2021.[2] The 2022 benefit year uses the Version 07 (V07) HHS-HCC classification, first implemented in the 2021 benefit year. The 2022 benefit year risk adjustment model was recalibrated using blended coefficients from the 2016, 2017, and 2018[3] enrollee-level External Data Gathering Environment (EDGE) data. The high-cost risk pool calculation incorporated into the HHS risk adjustment methodology since the 2018 benefit year continued for the 2022 benefit year.

Consistent with 45 C.F.R. § 153.320(b)(1)(i), the Centers for Medicare & Medicaid Services (CMS) released a document on July 19, 2021 that detailed a further update to the 2022 benefit year final risk adjustment model coefficients (or factors), which reflected an additional adjustment to apply a constraint to the enrollment duration factors (EDFs) that was missing in the 2022 Payment Notice final rule's risk adjustment model coefficients. [4] The 2022 benefit year risk adjustment model coefficients used in this software correspond to the final set of model coefficients posted July 19, 2021.[5]

The methodology that HHS will use when operating a risk adjustment program on behalf of a State for the 2022 benefit year[6] will calculate a plan average risk score for each covered plan based upon the relative risk of the plan's enrollees and apply a state payment transfer formula to

[1] This document corresponds to software for the HHS risk adjustment models for the 2022 benefit year, with revisions from the previous 2022 benefit year software instructions posted on the CCIIO website on December 29, 2022, available at: https://www.cms.gov/files/zip/hhs-hcc-software-v0722141b2.zip.
[2] See the Patient Protection and Affordable Care Act; HHS Notice of Benefit and Payment Parameters for 2022; Final Rule; 86 FR 24140 (May 5, 2021), available at: https://www.govinfo.gov/content/pkg/FR-2021-05-05/pdf/2021-09102.pdf.
[3] The coefficients for RXC 09 *Immune Suppressants and Immunomodulators* and its associated HCCs and interactions were developed using blended coefficients from only 2016 and 2017 enrollee-level EDGE data. See the 2022 Payment Notice final rule, 86 FR at 24180.
[4] See "Updated 2022 Benefit Year Final HHS Risk Adjustment Model Coefficients," July 19, 2021, available at: https://www.cms.gov/files/document/updated-2022-benefit-year-final-hhs-risk-adjustment-model-coefficients-clean-version-508.pdf.
[5] Ibid.
[6] HHS will operate risk adjustment for the 2022 benefit year in all 50 states and the District of Columbia.

determine risk adjustment payments and charges for plans within a State market risk pool. The HHS risk adjustment methodology addresses three considerations: (1) adverse selection in the individual and small group (including merged) markets; (2) plan metal level differences and permissible rating variation; and (3) the need for risk adjustment transfers that net to zero. The Federally certified risk adjustment methodology developed by HHS for the 2022 benefit year:

- Is developed on enrollee-level EDGE data, which directly reflects claims data for ACA individual and small group (including merged) market enrollees;
- Employs the hierarchical condition category (HCC) grouping logic used in the Medicare risk adjustment program, but with HCCs refined and selected to reflect the expected population in risk adjustment covered plans[7];
- Includes a selected number of Prescription Drug Categories (RXCs) and RXC interactions in the adult models;
- Establishes concurrent risk adjustment models, one for each combination of metal level (platinum, gold, silver, bronze, catastrophic) and age group (adult, child, infant);
- Pools catastrophically high-cost enrollees nationally with a portion of the costs funded by a percent of premium charge to issuers of risk adjustment covered plans in each national market[8];
- Results in state transfers that net to zero within a State market risk pool;
- Adjusts transfers for plan metal level, geographic rating area, induced demand, premium assistance Medicaid alternative plans, and age rating, so that transfers reflect health risk and not other cost differences; and
- Transfers funds between plans within a State market risk pool based on differences in relative actual risk.[9]

Key Revisions for Benefit Year 2022:

- (July 2022 Revisions) Updated Table 2 to add 2022 CPT/HCPCS codes used for diagnosis filtering, as described in Section III. Table 2 includes review of 2022 quarterly updates with effective dates as of April 1, 2022. Replaced the 2020 column of code information with 2021 codes (used for historical data purposes).
- (July 2022 Revisions) Updated software to include Fiscal Year (FY) 2022 ICD-10 diagnosis code assignments and FY2022 Medicare Code Editor (MCE) edits and to remove FY2021 ICD-10 assignments and FY2021 MCE edits. Revised fiscal year validity checks for ICD-10 diagnosis codes and corresponding service dates.
- (July 2022 Revisions) Revised Table 3 ICD-10 to HHS-Condition Categories (CC) Crosswalk to remove Fiscal Year (FY) 2021 and CY2021 Medicare Code Editor (MCE) columns. Revised explanatory text in Sections II and V to clarify that

[7] See 45 CFR 153.20 for the definition of "Risk Adjustment Covered Plan".

[8] The high-cost risk pool calculations under the HHS risk adjustment methodology involve two national markets – one for the individual market (including catastrophic and non-catastrophic plans, and merged market plans), and another for the small group market. See, for example, the Patient Protection and Affordable Care Act; HHS Notice of Benefit and Payment Parameters for 2018; Amendments to Special Enrollment Periods and the Consumer Operated and Oriented Plan Program; Final Rule, 81 FR 94058 at 94080 – 94082 (Dec. 22, 2016).

[9] For the 2022 benefit year, statewide average premiums are reduced (i.e., adjusted) by 14 percent in the state payment transfer formula to account for the proportion of administrative costs that do not vary with claims. See the 2018 Payment Notice, 81 Fed. Reg. at 94099 - 94100. Also, see the 2019 Payment Notice, 83 Fed. Reg. at 16955; 2020 Payment Notice, 84 Fed. Reg. at 17485 - 17486; and, 2021 Payment Notice, 85 Fed. Reg. at 29192.

FY2022 ICD-10 diagnosis codes and FY2022 MCE edits should be used in 2022 benefit year risk adjustment and that FY2021 code valid information is retained for historical purposes.

- (July 2022 Revisions) Updated Tables 10a and 10b to contain NDCs and HCPCS codes in the National Library of Medicine's RxNorm dataset as of April 1, 2022.
- (December 2022 Revisions) Updated Table 2 to include review of 2022 quarterly updates of CPT/HCPCS codes with effective dates as of October 1, 2022.[10] In addition, anesthesiology, echocardiography, and therapeutic radiology diagnostic procedure codes have been removed from RA eligibility and will no longer be accepted for risk adjustment eligible diagnosis filtering.[11] The Calendar Year (CY) 2022 column in Table 2 has been changed to a value of "no" for these codes.
- (July 2022 Revisions) Updated coefficients and denominators for the 2022 benefit year using 2016, 2017, and 2018 EDGE data (Sections II and VIII).
- (July 2022 Revisions) Added facility bill type code 731 and 737 (clinic – freestanding admit through discharge and replacement of prior claim) to the list of acceptable outpatient facility sources of diagnosis.
- (December 2022 Revisions) Updated software to account for new FY2023 ICD-10 codes, new CY2022 (FY22/FY23) MCE edits, and updated CC assignments. Revised Table 3 ICD-10 to HHS-Condition Categories (CC) Crosswalk to include FY2022 and FY2023 ICD-10 diagnosis codes and FY2022 and FY2023 MCE age and sex conditions. Updated ICD-10 code labels to reflect changes in FY2023. Updated CC assignments to account for new FY2023 ICD-10 codes. Updated the combined set of MCE age and sex conditions to be used for CY2022 that covers both fiscal years (FY2022 and FY2023). Revised explanatory text in Section III to clarify the use of FY2022 and FY2023 ICD-10 diagnosis codes and MCE edits.
- (December 2022 Revisions) Updated software to account for most recent 2022 NDC and HCPCS codes used in RXC crosswalks. Updated Tables 10a and 10b to include NDCs and HCPCS codes in the National Library of Medicine's RxNorm dataset as of October 1, 2022. (Tables 10a and 10b will be updated as EDGE reference data updates and posted on REGTAP in April 2023 to be used as the final set of NDCs and HCPCS codes for the 2022 benefit year.)
- (April 2023 Revisions) Updated software to account for most recent NDC and HCPCS codes used in RXC crosswalks. Updated Tables 10a and 10b to contain NDCs and

[10] As a result of the Coronavirus (COVID-19) public health emergency (PHE), the CPT/HCPCS list in Table 2 was previously expanded in Calendar Year (CY) 2020 to include certain telehealth and telephonic service codes that were not previously accepted for the HHS-operated risk adjustment program. The allowable telehealth services on Table 2 will continue to be accepted for risk adjustment eligible diagnosis filtering for the HHS-operated risk adjustment program applicable for the individual and small group (including merged) markets for the 2022 benefit year, if the services are otherwise allowable under applicable state law. For more information on the telehealth and telephonic services that are valid for the HHS-operated risk adjustment program, please refer to the Risk Adjustment Telehealth and Telephone Services During COVID-19 FAQs (https://www.cms.gov/files/document/RA-Telehealth-FAQ.pdf), updated April 29, 2022.

[11] These diagnoses could not be validated in HHS risk adjustment data validation (HHS-RADV) since stand-alone diagnostic reports are not able to substantiate a diagnosis and must be submitted in conjunction with a valid medical record. Diagnoses that are more reliably accurate may be found on corresponding evaluation and management (E&M) claims or surgery claims. For more information on this filtering update, please refer to the BY 2022 EDGE Server Webinar Series Announcements (https://regtap.cms.gov/reg_library_openfile.php?id=4185&type=l) (Slide 19), posted November 29, 2022.

HCPCS codes in the National Library of Medicine's RxNorm dataset as of December 1, 2022 to be used as the final set of NDC and HCPCS codes for the 2022 benefit year.

The state payment transfer formula that is part of the HHS risk adjustment methodology consists of concurrent risk adjustment models, one for each combination of metal level (platinum, gold, silver, bronze, and catastrophic) and age group (adult, child, infant). This document provides the detailed information needed to calculate risk scores given individual diagnoses. Please direct questions regarding these instructions to CMS at RARIPaymentOperations@cms.hhs.gov.

CMS has created two versions of software (SAS software and HHS-developed risk adjustment model algorithm "Do It Yourself [DIY]" software) and software instructions for issuers to use with their enrollment data to simulate their enrollee populations' 2022 benefit year risk scores within the HHS-HCC risk adjustment models. **This software is being issued only as a supplemental tool for issuers of risk adjustment covered plans to better understand and simulate the calculation of plan liability risk scores for their enrollees.**

This software is not a required prerequisite to submitting claims data to the EDGE server for risk adjustment, nor is it a requirement of the HHS-operated risk adjustment program. Furthermore, issuers should not use this software to filter their own claims prior to submitting claims data to the EDGE server. The EDGE server software may have several additional layers of operational rules. This software merely provides a simulation tool for issuers to calculate enrollees' risk scores. Because risk adjustment transfers are dependent on the data submitted by other issuers within the State market risk pool, an issuer that wishes to use this information to assist with estimating its 2022 benefit year transfer(s) should do so with caution and in combination with other data.

This document describes software for HHS-HCC risk adjustment modeling (Version 07). The software requires SAS® version 9.

This software (V0722 141 B3) is designed to be used only with 2022 dates of service and with ICD-10 diagnosis codes. If the user will be using historical data (i.e., 2021 or earlier service dates), the user should refer to earlier versions of the software for HHS-HCC risk adjustment modeling also posted on the CCIIO website.

List of Tables[12]

[12] HCCs described in the tables that have splits, such as HCC035_1 and HCC035_2, are identified in the EDGE reference data as 351 and 352.

Terminology: The abbreviations ICD-10 and ICD-10-CM are used interchangeably in this document to refer to International Classification of Diseases, 10th Revision, Clinical Modification. The abbreviations CC and HCC used in these instructions refer to the HHS-HCC risk adjustment models. These are different HCCs from those used in the CMS-HCC risk adjustment model for Medicare Part C.

I. Software description

The software reads four user-provided input SAS® data sets (Section IV); constructs demographic variables for each enrollee; crosswalks ICD-10 diagnoses to Condition Categories (CCs) using SAS® formats which are stored in a FORMAT library; creates Hierarchical Condition Categories (HCCs) by imposing hierarchies on the CCs; creates Prescription Drug Categories (RXCs) based on National Drug Codes (NDCs) and Healthcare Common Procedure Coding System (HCPCS) codes, and imposes hierarchies on RXCs.

The software uses the demographic variables, adult enrollment duration variables, HCCs, and RXCs to compute risk scores for three models (adult, child, infant); cost sharing reduction (CSR)-adjusted scores for each model including adjustment for enrollment in premium assistance Medicaid alternative plans; and final scores based on the enrollee's age and plan benefit design. Scores for enrollees without diagnoses, NDCs, or HCPCS codes are computed from person-level variables; i.e., zeros are assigned to all CCs, HCCs, and RXCs.

The software's main program (CY22P07C) calls primary macro CY22M07C and passes a set of user-specified parameters (a macro is a subroutine that performs a specific task). Macro CY22M07C calls five external macros (provided as separate files):

- AGESEXV6 – creates age/sex variables;
- I0V07ED1 – performs edits on ICD-10 codes based on age and/or sex;
- V07141L1 – assigns labels to HCCs and RXCs;
- V07141H1 – sets selected HCCs to zero based on hierarchical rules;
- SCOREV4 – calculates risk score variables.

Identical program files with .SAS and .TXT extensions are provided. The .TXT versions are easier to view with some programs. The user must use the files with extension .SAS when installing the software. File names are case sensitive on some computing platforms, so software modules assume that file names are upper case (e.g., I0V07ED1.SAS).

The software:

Step 1: Includes external macros; these are most likely to vary among releases.

Step 2: Defines internal macro variables, formats, and internal macros.

Step 3: Merges the PERSON, NDC, HCPCS, and DIAGNOSIS SAS® data sets, and outputs one record for each enrollee record in the PERSON data set. Input records must be fully compliant with validity rules (e.g., SEX must be M/m/F/f/1/2), and all data sets must be sorted

by the common person identifier variable. The name of the common person identifier variable is set in the macro variable &IDVAR (e.g., &IDVAR = *ID, or EnrolleeID*).

Step 3.1: Declares variable lengths, retained variables, and arrays.

Step 3.2: Appends calibration coefficients for all models.

Step 3.3: Merges the PERSON, NDC, HCPCS and DIAGNOSIS data sets by the person identifier variable named in &IDVAR. Each enrollee must have exactly one PERSON record, and may have zero or more NDC, HCPCS, or DIAGNOSIS records.

Step 3.4: Performs tasks when the enrollee's first record is detected.

Step 3.5: If the enrollee has at least one NDC or HCPCS code, this step: creates RXCs using the crosswalk formats specified in parameter &RXCFMTN and &RXCFMTH (see Section II for details regarding the format library and formats specific to this version of software).

Step 3.6: If the enrollee has at least one diagnosis, this step: creates CCs using the crosswalk formats specified in parameter &CCFMT0Y1 and &CCFMT0Y2 (see Section II for details regarding the format library and formats specific to this version of software); performs ICD-10 edits using macro I0V07ED1; and creates additional CCs for some ICD-10 diagnoses.

Step 3.7: When the enrollee's last record is detected, this step: creates demographic variables using macro AGESEXV6; creates HCCs by applying hierarchy rules to CCs using macro V07141H1; sets HCCs to zero if the enrollee has no diagnoses; applies hierarchy rules to RXCs; sets RXCs to zero if the enrollee has no NDC or HCPCS codes; applies validity filters to various input variables; creates additional model-specific variables (e.g., severe illness indicators, HCC groups, interaction terms, adult enrollment duration indicators, RXC and RXC*HCC interactions); creates unadjusted and CSR-adjusted scores for each plan level for each enrollee including enrollment in premium assistance Medicaid alternative plans; and defines output formats and labels for variables.

Step 4: The software uses SAS® CONTENTS and PRINT procedures to document the output data set.

II. Files included with the software

The following programs and files are included:

- **CY22P07C –** main program containing all user-provided parameters (see below for the parameter and variable list). The program calls primary macro CY22M07C.
- **CY22M07C –** primary macro that merges input files, crosswalks NDCs and HCPCS to RXCs, crosswalks ICD-10 codes to CCs, creates HCC and risk score variables by calling various external and internal macros. Table 3, ICD-10 to Condition Categories (CC) Crosswalk, summarizes the ICD-10 to CC assignments. Only ICD-10 codes assigned to HCCs in the risk adjustment models are included in this crosswalk. All other ICD-10

codes will be ignored by the software. Table 10a, NDC to RXC Crosswalk, and Table 10b, HCPCS to RXC Crosswalk, summarize the NDC and HCPCS assignments to RXCs. NDC and HCPCS not listed in the tables will be ignored by the software.

- **AGESEXV6 –** creates age/sex variables.
- **I0V07ED1 –** performs edits on ICD-10 codes based on age and/or sex. The Medicare Code Edits (MCEs) and further specified CC age and sex splits are performed by this macro.[13] If the enrollee has an invalid age and/or sex for a particular ICD-10 code, then the ICD-10 code will be ignored. Table 3, ICD-10 to Condition Categories (CC) Crosswalk, summarizes the ICD-10 code edits; it describes the ICD-10 Medicare Code Edits (MCEs) for age and sex, and additional edits for CC age and sex splits.
- **V07141L1 –** assigns labels to HCCs and RXCs. Table 4, HHS-Hierarchical Condition Categories (HCC) Hierarchies, lists the HCC labels.
- **V07141H1 –** copies CCs into HCCs and sets selected HCCs to zero based on hierarchical rules. Table 4, HHS-Hierarchical Condition Categories (HCC) Hierarchies, summarizes the hierarchy assignments.
- **SCOREV4 –** calculates risk score variables.
- **CY22F07C FY2022 ICD10.TXT** – is a text version of the format that crosswalks ICD-10 codes to CC categories (and is provided for reference). The format includes ICD-10 codes valid in FY2022.
- **CY22F07C_FY2023 ICD10.TXT** – is a text version of the format that crosswalks ICD-10 codes to CC categories (and is provided for reference). The format includes ICD-10 codes valid in FY2023.
- **CY22F07C_ICD10_MCE_AGE.TXT** – is a text version of the format that crosswalks ICD-10 codes to an acceptable age range if MCE edits on ICD-10 codes are to be performed (provided for reference only).
- **CY22F07C_ICD10_MCE_SEX.TXT** – is a text version of the format that crosswalks ICD-10 codes to an acceptable sex value if MCE edits on ICD-10 codes are to be performed (provided for reference only).
- **CY22F07C_ICD10_BUNDLED_MOTHER.TXT** – is a text version of the format that contains FY2022/FY2023 completed pregnancy diagnoses for use in detecting mother-infant bundled claims (provided for reference only).
- **CY22F07C_ICD10_BUNDLED_INFANT.TXT** – is a text version of the format that contains FY2022/FY2023 newborn diagnoses for use in detecting mother-infant bundled claims (provided for reference only).
- **CY22F07C_NDC8_4_22_12.TXT** – is a text version of the format that contains Table 10a RXC to National Drug Code (NDC) Crosswalk.
- **CY22F07C_HCPCS8_4_22_12.TXT** – is a text version of the format that contains Table 10b RXC to Healthcare Common Procedure Coding System (HCPCS) Crosswalk.
- **CY22F07C.TRN –** a SAS® transport file containing one format library with all requisite formats. Format name suffixes must be specified as macro parameters in the main program as follows:

[13] The diagnosis-code edits used are based on the Definitions of Medicare Code Edits (MCEs), which are updated and published each year to correspond with ICD-10 code updates. The MCEs detect inconsistencies based on a person's age and diagnosis or sex and diagnosis.

    - HHS_V07FY22R141C – crosswalks ICD-10 codes to CC categories that are transformed to HCC categories, and contains ICD-10 codes used in the risk adjustment models that are valid in FY2022. This suffix must be specified in macro parameter **CCFMT0Y1** for this version of the software.
    - HHS_V07FY23R141C – crosswalks ICD-10 codes to CC categories that are transformed to HCC categories, and contains ICD-10 codes used in the risk adjustment models that are valid in FY2023. This suffix must be specified in macro parameter **CCFMT0Y2** for this version of the software.
    - NDCV2212_RXCV8_4F – crosswalks NDC codes to RXC categories for codes valid in calendar year 2022. This format must be specified in macro parameter **RXCFMTN**.
    - HCPC2212_RXCV8_4F – crosswalks HCPCS codes to RXC categories for codes valid in calendar year 2022. This format must be specified in macro parameter **RXCFMTH**.
    - AGECY22MCE – crosswalks ICD-10 codes to an acceptable age range if MCE edits on ICD-10 codes are to be performed. This suffix must be specified in macro parameter **AGEFMT0**.
    - SEXCY22MCE – crosswalks ICD-10 codes to an acceptable sex value if MCE edits on ICD-10 codes are to be performed. This suffix must be specified in macro parameter **SEXFMT0**.
- **CY22C07A.TRN** – a SAS® transport file containing relative coefficients for regression models, created using CY2016, CY2017, and CY2018 data and corresponding denominators (defined as the weighted average plan liability for the full modeling sample for the given year).

The two SAS® transport files (with filename extension .TRN) contain the SAS® format library and model coefficients data set. They may be used on any SAS® version 9 platform after uploading them and converting them using SAS® PROC CIMPORT.

If your computing platform is z/OS, both transport files should be uploaded using the following attributes: RECFM(F or FB) LRECL(80) BLKSIZE(8000).

The two transport files should be converted (imported) as follows:

- Model coefficients:

```
FILENAME INC     "user defined location of transport file CY22C07A.TRN";
LIBNAME  INCOEF  "user defined location for creation of coefficient file";

proc cimport infile=INC data=INCOEF.Coefficients; run;
```

- Format library:

```
FILENAME INF     "user defined location of transport file CY22F07C.TRN";
LIBNAME LIBRARY  "user defined location for creation of format library";

proc cimport infile=INF library=LIBRARY; run;
```

III. Creation of a diagnosis data set, NDC data set, and HCPCS data set

A. Diagnosis-level data set. The diagnosis input SAS® data set (DIAGNOSIS) must include ICD-10-CM diagnosis codes used for risk adjustment, listed in Table 3, ICD-10 to Condition Categories (CC) Crosswalk. The user must evaluate each claim or encounter record to determine whether its diagnoses are included in the DIAGNOSIS data set. Encounter records normally report dates, provider or bill types, diagnoses and procedures, and other information, though they may not have payment information.

This section explains how each record is evaluated to determine whether the record's diagnoses are to be used in CC/HCC creation. It is the user's responsibility to create the DIAGNOSIS data set according to the filtering logic below. This document provides filtering instructions and a list of the CY2021 (for historical data purposes) and CY2022 CPT/HCPCS codes that define service or procedure types that identify acceptable sources of diagnoses for risk adjustment.[14] However, the user must create the DIAGNOSIS data set for input to the risk adjustment algorithm; the data set is not created by the software.

NOTE: Supplemental diagnosis codes may be submitted in certain circumstances. These instructions and the software do not address the addition of supplemental diagnosis codes. Therefore, risk score output from this software will not account for inclusion of supplemental diagnoses.

Only ICD-10-CM diagnosis codes from sources allowable for risk adjustment should be included in the DIAGNOSIS data set. ICD-10 codes that are not listed in Table 3 may be included in the DIAGNOSIS data set but are ignored by the software.[15] The steps below provide logic to determine which diagnoses are allowable. Note that Steps 1 and 3 refer to Table 2, CPT/HCPCS Included List for Diagnosis Code Filtering, which provides the 2021 (for historical data purposes) and 2022 CPT/HCPCS codes used to define service or procedure types that are acceptable sources of diagnoses for risk adjustment.

- The CPT/HCPCS codes identifying services with diagnoses allowable for risk adjustment are listed in column A of Table 2.
- Column B lists the short descriptions of the CPT/HCPCS codes.
- Columns C and D, respectively, indicate whether a CPT/HCPCS code is acceptable in 2021 or 2022.
- Column E identifies applicable footnotes on the CPT/HCPCS codes.
- Notes begin on row 6,666 of the Excel table with the line "Notes:" and should not be imported by any program.

The DIAGNOSIS data set should include diagnoses from claims/encounter records with **discharge dates or through dates** within the benefit year. Though the term "claim" is used in the steps below, the steps apply equally to encounter records. For the EDGE server, only claims with discharge diagnoses are used for HHS risk adjustment.

[14] Definitions taken directly from the Current Procedural Terminology (CPT®) codes and the Healthcare Common Procedure Coding System (HCPCS) code set. Note that although CY2021 codes are provided for historical purposes, this software is designed to be used only with CY2022 data.

[15] If the user conducts fiscal year code validity checks described later in this section before using the software, only codes valid for risk adjustment will be included in the final diagnosis-level file.

1. Professional source of diagnosis
    a. For professional records, use diagnoses from records that have at least one line item with an acceptable CPT/HCPCS code (Table 2). If there is at least one acceptable line on the record, use all the header diagnoses. There are three possible values for CPT/HCPCS codes in columns C and D:
        i. yes = code is acceptable in that calendar year
        ii. no = code is not acceptable in that calendar year
        iii. N/A = code is not in existence in that calendar year
    b. For professional records, if a line item has an acceptable CPT/HCPCS code, use all diagnoses from the line item.
    c. If there are no acceptable service lines on the record, do not use any of the diagnoses for risk adjustment.
2. Inpatient facility source of diagnosis
    a. Use all header diagnoses from records where facility bill type code equals one of the following:
        i. 111 (inpatient admit through discharge); or
        ii. 117 (inpatient replacement of prior claim).
    b. There is no procedure screen for inpatient facility record types.
3. Outpatient facility source of diagnosis
    a. Restrict records to those with facility bill type code equal to:
        i. 131 (hospital outpatient admit through discharge); or
        ii. 137 (hospital outpatient replacement of prior claim); or
        iii. 711 (rural health clinic admit through discharge); or
        iv. 717 (rural health clinic replacement of prior claim); or
        v. 731 (clinic – freestanding admit through discharge); or
        vi. 737 (clinic – freestanding replacement of prior claim); or
        vii. 761 (community mental health center admit through discharge); or
        viii. 767 (community mental health center replacement of prior claim); or
        ix. 771 (federally qualified health center admit through discharge); or
        x. 777 (federally qualified health center replacement of prior claim).
        xi. 851 (critical access hospital admit through discharge); or
        xii. 857 (critical access hospital replacement of prior claim); or
        xiii. 871 (freestanding non-residential opioid treatment programs [OTPs]); or
        xiv. 877 (OTPs replacement of prior claim).
    b. For records with at least one acceptable CPT/HCPCS code (Table 2) on a service line, use all header diagnoses. Otherwise, do not use the diagnoses for risk adjustment.

Fiscal year code validity: Section IV further describes the diagnosis input data set. After creating that data set, the user will have the variables needed to conduct fiscal year validity checks before using the software if desired. Table 3 identifies the fiscal year(s) in which the diagnosis codes used for risk adjustment are valid. The user should check that for a given diagnosis (variable DIAG) and service date (variable DIAGNOSIS_SERVICE_DATE), the diagnosis code has a Y in the corresponding Table 3 Code Valid column. ICD-10 diagnosis codes with service dates of January 1, 2022 – September 30, 2022 should have a Y in the Code Valid in FY2022 column; otherwise, the user should exclude them. ICD-10 diagnosis

codes with service dates of October 1, 2022 – December 31, 2022 should have a Y in the Code Valid in FY 2023 column; otherwise, the user should exclude them. As noted, this software can detect that an ICD-10 diagnosis code is not valid for a given fiscal year and will optionally flag the enrollee record in the "Errors/warnings/notes log" (see Section VIII.5, message 16).

Note on bundled claims for mother and newborn infant: In practice, some hospital claims for childbirth include both the mother's record and the newborn infant's record on the same claim (diagnoses and procedure codes). Because there are separate adult, child, and infant risk adjustment models and some of the diagnosis codes may not be distinguishable between mother and infant on bundled claims, **any bundled claims should be redefined as two separate records whenever possible (mother and infant, <u>each with a separate ID, sex, and age</u>) in order for the diagnoses to be appropriately included in the input data set and used for appropriately calculating risk scores.**

The user will need to independently create a program to detect any bundled claims and redefine them as two separate claims (i.e., it is not part of these instructions). For example, a bundled claim detection program would need to identify enrollees with a claim containing the following elements:

Mother is the enrollee:

- AGE_LAST >= 2 (an age corresponding to the child or adult models; more specifically age should be appropriate for a maternity diagnosis [i.e., 9-64)[16] and
- ICD-10 diagnoses corresponding to a completed pregnancy HCC (HCC 207 or 208 or 209) and
- ICD-10 diagnoses corresponding to a newborn HCC (HCC 242 or 243 or 244 or 245 or 246 or 247 or 248 or 249).

Infant is the enrollee:

- AGE_LAST = 0 (an age corresponding to the infant model; more specifically age is appropriate for a newborn diagnosis at birth) and
- ICD-10 diagnoses corresponding to a completed pregnancy HCC (HCC 207 or 208 or 209) and
- ICD-10 diagnoses corresponding to a newborn HCC (HCC 242 or 243 or 244 or 245 or 246 or 247 or 248 or 249).

See CY22F07C_ICD10_BUNDLED_MOTHER.TXT and CY22F07C_ICD10_BUNDLED_INFANT.TXT or Table 3, ICD-10 to Condition Category (CC) Crosswalk, for diagnosis codes corresponding to the completed pregnancy and newborn HCCs.

As noted, this software can detect that an enrollee might have bundled claims and will optionally flag the enrollee record in the "Errors/warnings/notes log," but it cannot redefine them as separate mother/infant claims (see Section VIII.5, message 25).

[16] Section IV of this document identifies the two age variables used in the software and specifies when each is used. MCE edits specifying appropriate age for maternity diagnoses are included in Table 3.

Infants with a record in the person-level file that cannot be matched with a claim or who do not have claims will have no diagnoses in the diagnosis data set. Infants without diagnoses will be assigned to the lowest severity category and the Age 1 maturity category for infants. Age 0 infants with diagnoses but who lack a newborn HCC will be assigned to the corresponding severity category and the Age 1 maturity category for infants. Male infants will also have the male demographic factor assigned. Age 0 male infants who lack a newborn HCC will have their demographic factor reassigned to Age 1.

B. NDC-level data set. The National Drug Code input SAS® data set (NDC) must include NDCs used for risk adjustment, listed in Table 10a RXC to NDC Crosswalk. Only pharmacy claims (not medical claims) are the acceptable sources for NDCs. The user must evaluate each claim to determine whether the claim's NDCs are included in the NDC data set.

The NDCs are to be used for RXC creation. It is the user's responsibility to create the NDC data set for input to the risk adjustment software; the data set is not created by the software. The inclusion of RXCs in the 2022 benefit year HHS operated risk adjustment methodology is limited to the adult risk adjustment models. Users should not include information for child or infant enrollees in the NDC data set.

The NDC data set should include NDCs from pharmacy claims with **prescription filled dates** within the benefit year. NDC codes should be in the 11-digit, no dashes, HIPAA format to match the format required for EDGE submission.[17] (Note: Table 10a in the Excel file contains the NDC codes formatted as text, not numbers, to retain any leading zeroes needed for 11-digit codes.) NDC codes that are not listed in Table 10a may be included in the NDC data set but are ignored by the software and are not included in RXCs for the adult risk adjustment models' risk score calculations. Section IV further describes the NDC data set.

C. HCPCS-level data set. The Healthcare Common Procedure Coding System (HCPCS) input SAS® data set must include HCPCS codes used for risk adjustment RXCs, listed in Table 10b RXC to HCPCS Crosswalk. Inpatient, outpatient, and professional medical claims are acceptable sources for HCPCS codes. Inpatient and outpatient claims should be restricted to the same facility bill type codes used for the diagnosis data set (see Section III. 2a and 3a). HCPCS should only be used for medications when an NDC is not available from a pharmacy claim. The user must evaluate each claim to determine whether the claim's HCPCS codes are included in the HCPCS data set.

The HCPCS codes in the HCPCS input data set are to be used for RXC creation. It is the user's responsibility to create the HCPCS data set for input to the risk adjustment software; the data set is not created by the software. The inclusion of RXCs in the 2022 benefit year HHS operated risk

[17] The source for the NDC codes is the U.S. Food and Drug Administration's Comprehensive NDC SPL Data Elements File: https://www.fda.gov/ForIndustry/DataStandards/StructuredProductLabeling/ucm240580.htm. The NDCs are validated as current prescriptions through the U.S National Library of Medicine's RxNorm data set: https://www.nlm.nih.gov/research/umls/rxnorm/. The RxNorm Technical Documentation includes an algorithm the user can access to normalize NDC codes to the 11-digit, no dashes, HIPAA format. The source for the NDC start/end dates is the U.S. Food and Drug Administration's Orange Book: https://www.accessdata.fda.gov/scripts/cder/ob/index.cfm.

adjustment methodology is limited to the adult risk adjustment models. Users should not include information for child or infant enrollees in the HCPCS data set.

The HCPCS data set should include HCPCS codes from inpatient, outpatient, and professional medical claims with **discharge dates or through dates** within the benefit year for adult enrollees. HCPCS codes that are not listed in Table 10b may be included in the HCPCS data set but are ignored by the software and are not included in RXCs for the adult risk adjustment models' risk score calculations. Section IV further describes the HCPCS data set.

IV. SAS® data sets supplied by the user

This section describes the four input SAS® data sets required to create CC and HCC groupings, RXC and RXC*HCC interactions, enrollment duration variables, demographic variables, and risk score variables—a person-level data set (PERSON), a diagnosis data set (DIAGNOSIS), an NDC data set (NDC), and a HCPCS data set (HCPCS). It is the responsibility of the user to create these input data sets with the variables listed in this section. All input data sets must be ordered in ascending order by the person identifier variable.

Note on CSR_INDICATOR

In operations, cost-sharing reduction (CSR) plan variations and premium assistance Medicaid Alternative plans (i.e., private options) will be identified by the Health Insurance Oversight System (HIOS) variant ID. Listed below are the codes that will be used to identify the plan variation.[18] Please note that unlike the risk adjustment software person-level CSR indicator, the HIOS variant ID is a plan-level indicator.

Cost-Sharing Reduction (CSR) Level	HIOS Variant ID	CSR RA Factor	RA Software Person-level CSR Indicator
CSR: 94% AV Silver Plan Variation	06	1.12	1
CSR: 87% AV Silver Plan Variation	05	1.12	2
CSR: 73% AV Silver Plan Variation	04	1.00	3
CSR: Zero Cost Sharing – Platinum	02	1.00	4
CSR: Zero Cost Sharing – Gold	02	1.07	5
CSR: Zero Cost Sharing – Silver	02	1.12	6
CSR: Zero Cost Sharing – Bronze	02	1.15	7
CSR: Limited Cost Sharing – Platinum	03	1.00	8
CSR: Limited Cost Sharing – Gold	03	1.07	9
CSR: Limited Cost Sharing – Silver	03	1.12	10
CSR: Limited Cost Sharing – Bronze	03	1.15	11
CSR: Premium Assistance Medicaid Alternative Plan w/94% AV Silver Plan	36	1.12	12

[18] We note that Massachusetts CSR variant plans have a state-specific CSR factor table, as discussed in the 2022 Payment Notice final rule, 86 at 24180). In addition to the CSR variants listed above with factors of 1.12, plan variants of 04 are also 1.12 in Massachusetts only.

Cost-Sharing Reduction (CSR) Level	HIOS Variant ID	CSR RA Factor	RA Software Person-level CSR Indicator
CSR: Premium Assistance Medicaid Alternative Plan w/Zero Cost Sharing – Silver	32	1.12	13
Non-CSR/unknown CSR	00	1.00	0
Non-CSR/unknown CSR	01	1.00	0

Note on Enrollment Duration

The adult models include enrollment duration factors for months enrolled when an enrollee's enrollment period in an issuer's plans is less than 12 months. There are two steps involved in creating the enrollment duration indicator variables:

STEP 1: For the PERSON file, the user should create an ENROLDURATION variable for each enrollee with 12 possible values corresponding to 1-12 months based on an enrollee's total number of days enrolled in the plan in the benefit year as described below. Although ENROLDURATION will only be used to create variables needed for the adult models, this software was designed for ENROLDURATION to be constructed for *all* enrollees to maintain consistency in the variables present in the PERSON file. Thus, enrollees missing ENROLDURATION will receive this Error message: *WARNING: [Msg33] Invalid ENROLDURATION, enrollee rejected.* Once created, the ENROLDURATION variable will be ignored for enrollees in the child or infant models.

STEP 2: The monthly enrollment duration indicator variables (ED_1–ED_11) are created by the software for adult enrollees as specified in Section VI.

The variable names must be spelled as written; SAS® variable names are case-insensitive (i.e., SEX and Sex and sex and SeX designate the same variable), but are illustrated in upper case.

1. PERSON data set
    a. &IDVAR (Person identification code). As noted, &IDVAR is the name of the common person identifier variable (e.g., ID).
        i. Character or numeric type, any length, not missing.
        ii. Unique to an individual, and unique in the data set (i.e., no duplicates).
    b. SEX.
        i. Character type, 1 byte, 1/M=male, 2/F=female, not missing.
        ii. Converted to upper case by the software.
    c. DOB.
        i. Numeric type, 8-digit numeric field (YYYYMMDD), valid calendar date, not missing, provides the enrollee's date of birth.
        ii. Used to calculate AGE_AT_DIAGNOSIS for MCE diagnosis code age edits.
    d. AGE_LAST (Age as of last day of enrollment in benefit year).

i. Numeric type, integer, 0 or greater, not missing.
ii. Used for all risk adjustment purposes except MCE diagnosis code age edits.
iii. For infants born in the previous year but not discharged until the benefit year, users should substitute Age 0 for Age 1 in AGE_LAST.

e. METAL (Enrollee's plan level – platinum, gold, silver, bronze, catastrophic).
   i. Character type, 1 byte, P/G/S/B/C (only 1 of these values), not missing.[19]
   ii. Converted to upper case by the software.

f. CSR_INDICATOR (Person-level indicator. Enrollees who qualify for cost-sharing reductions or those enrolled in premium assistance Medicaid alternative plans must be assigned CSR_INDICATOR = 1-13. Non-CSR recipients must be assigned CSR_INDICATOR = 0).
   i. Numeric type, integer, 0-13, not missing.
   ii. Values are:
      - 1 = Enrollees in 94% AV Silver Plan Variation.
      - 2 = Enrollees in 87% AV Silver Plan Variation.
      - 3 = Enrollees in 73% AV Silver Plan Variation.
      - 4 = Enrollees in Zero Cost Sharing Plan Variation of Platinum Level QHP.
      - 5 = Enrollees in Zero Cost Sharing Plan Variation of Gold Level QHP.
      - 6 = Enrollees in Zero Cost Sharing Plan Variation of Silver Level QHP.
      - 7 = Enrollees in Zero Cost Sharing Plan Variation of Bronze Level QHP.
      - 8 = Enrollees in Limited Cost Sharing Plan Variation of Platinum Level QHP.
      - 9 = Enrollees in Limited Cost Sharing Plan Variation of Gold Level QHP.
      - 10 = Enrollees in Limited Cost Sharing Plan Variation of Silver Level QHP.
      - 11 = Enrollees in Limited Cost Sharing Plan Variation of Bronze Level QHP.
      - 12 = Enrollees in a Premium Assistance Medicaid Alternative Plan with 94% AV Silver Plan Variation.
      - 13 = Enrollees in a Premium Assistance Medicaid Alternative Plan with Zero Cost Sharing Plan Variation of Silver Level QHP.
      - 0 = Non-CSR recipient, and enrollees with unknown CSR.

g. ENROLDURATION
   i. Numeric type, integer, 1-12, not missing.
   ii. Person-level enrollment duration variable. Although ENROLDURATION is for use in adult models only, user should create it for all enrollees for

[19] Although the user is required to select a single metal level for the enrollee, the software produces score variables for all levels. The final unadjusted and CSR-adjusted score variables correspond to the single metal level selected, as is noted in Section VI.

consistency in PERSON file preparation. Values will be ignored for enrollees in child or infant models.

iii. Allowable values are 1-12 based on months enrolled in plan in benefit year as defined by days:
- 1 = 1–31 days enrolled
- 2 = 32–62 days enrolled
- 3 = 63–92 days enrolled
- 4 = 93–123 days enrolled
- 5 = 124–153 days enrolled
- 6 = 154–184 days enrolled
- 7 = 185–214 days enrolled
- 8 = 215–245 days enrolled
- 9 = 246–275 days enrolled
- 10 = 276–306 days enrolled
- 11 = 307–335 days enrolled
- 12 = 336–366 days enrolled

2. DIAGNOSIS data set
   a. &IDVAR (Person identification code). As noted, &IDVAR is the name of the common person identifier variable (e.g., ID).[20]
      i. Character or numeric type, any length, not missing.
      ii. Unique to an individual.
   b. DIAG (ICD-10-CM diagnosis codes).
      i. Character type, 7-byte field, no periods or embedded blanks, left justified.
      ii. Converted to upper case by the software.
      iii. Codes should be to the greatest level of available specificity.
      iv. Age and sex edits for diagnoses are performed in macro I0V07ED1 to ensure diagnoses are appropriate for the age and sex of the enrollee.
      v. Only diagnoses from allowable sources should be included in the DIAGNOSIS data set.
      vi. Invalid diagnoses are ignored; warning messages are optional.[21]
      vii. A valid ICD-10 diagnosis must have a valid DIAGNOSIS_SERVICE_DATE.
   c. DIAGNOSIS_SERVICE_DATE
      i. Numeric type, 8-digit numeric field (YYYYMMDD), valid calendar date, not missing, provides the diagnosis's service date.[22]

[20] Please note that in operation, this information can not include personally identifiable information.
[21] In the context of this software's instructions, valid refers to "included" in the HHS-HCC risk adjustment model and invalid refers to "not included."
[22] Valid diagnosis service date in this version of software (V0722 141 B3): year is 2022, month is 01-12, and day is 01-31 and appropriate for the given month (i.e., cannot be February 30). The service date cannot occur before the date of birth.

ii. As described in Section III, this variable can be used with DIAG and Table 3 to precheck that a diagnosis code is valid for a given fiscal year.[23]

3. NDC data set
    a. &IDVAR (Person identification code). As noted, &IDVAR is the name of the common person identifier variable (e.g., ID).[24]
        i. Character or numeric type, any length, not missing.
        ii. Unique to an individual.
    b. NDC
        i. NDC normalized drug code (11-digit, no dashes, HIPAA standard format), 11-character field, left justified. This format matches the format required for submission to the EDGE server. Only NDCs from pharmacy claims filled in the benefit year or inpatient claims with discharges in the benefit year (Section III B) for adult enrollees should be included in the NDC data set.

4. HCPCS data set
    a. &IDVAR (Person identification code). As noted, &IDVAR is the name of the common person identifier variable (e.g., ID).[25]
        i. Character or numeric type, any length, not missing.
        ii. Unique to an individual.
    b. HCPCS
        i. HCPCS drug code, 5-character field, left justified. Only HCPCS drug codes from inpatient, outpatient, or professional medical claims from allowable sources (Section III) with discharge or through dates in the benefit year (Section III C) for adult enrollees should be included in the HCPCS data set.

AGE_AT_DIAGNOSIS, the age as of the diagnosis service date, is calculated by the software using DOB from the PERSON data set and DIAGNOSIS_SERVICE_DATE from the DIAGNOSIS data set. It is used only for MCE diagnosis code age edits.

The four user-provided data sets (PERSON, DIAGNOSIS, NDC, HCPC) are illustrated below. These examples are not based on actual data.

- Person-level data set example (PERSON) containing seven variables; we use ID as the person identifier variable to illustrate:

ID	SEX	DOB	AGE_LAST	METAL	CSR_INDICATOR	ENROLDURATION
201	M	19591201	63	P	0	12
202	F	20090315	13	C	0	12

[23] The software has a fiscal year validity check. If an ICD-10 code is not valid for a given DIAGNOSIS_SERVICE_DATE (e.g., a deleted in FY2022 code with a FY2022 service date), the optional software warning message will be Message 16 *Diagnosis lookup failed, diagnosis ignored.*
[24] Please note that in operation, this information can not include personally identifiable information.
[25] Please note that in operation, this information can not include personally identifiable information.

```
301    F    19670414    55        G          5              7
302    M    19730101    49        B          11             12
304    X    19710132              R          16             3
305    M    19820101    40        S          0              12
```

- Diagnosis data set example (DIAGNOSIS) containing three variables; we use ID as the person identifier variable and ICD-10 diagnoses to illustrate:

```
ID     DIAG     DIAGNOSIS_SERVICE_DATE

201    E118             20220113
201    H9319            20220113
201    M532X9           20220629
201    M25461           20220630
201    M25569           20220706
201    M25579           20220706
201    209              20220835
202    J4530            20220219
302    J200             20220317
302    Z430             20220504
303    E890             20220929
304    Z0000            20220617
305    B20              20220302
```

- NDC data set example (NDC) containing two variables; we use ID as the person identifier variable and NDC normalized drug code, 11-digits, HIPAA standard format, character field, left justified, to illustrate:

```
ID     NDC

201    00002751001
202
303    42291018920
304    13411019102
305    0003-1964-1
```

- HCPCS data set example (HCPCS) containing two variables; we use ID as the person identifier variable and HCPCS code, 5 digits, left justified, to illustrate:

```
ID     HCPCS

302    C9482
303    J1324
304    Q3028
305    J87
```

- ID 301 has no diagnoses; the other IDs in PERSON have one or more diagnoses.
- ID 303 in DIAGNOSIS, NDC, and HCPCS will be ignored because there is no ID 303 in PERSON.
- Missing or invalid information in any PERSON variable will cause that enrollee and all his/her diagnoses, NDCs, and HCPCS codes to be ignored (e.g., ID 304).

- Missing or invalid information in DIAGNOSIS will cause that diagnosis to be ignored (e.g., ID 201 DIAG 209).
- Missing or invalid information in NDC will cause that NDC to be ignored (e.g., ID 202, and ID 305).
- Missing or invalid information in HCPCS will cause that HCPCS to be ignored (e.g., ID 305).
- Risk scores for enrollees without diagnoses, NDCs, and HCPCS codes are calculated using only PERSON-level information (e.g., ID 301).

If an enrollee has N different diagnoses, the enrollee will have N records in DIAGNOSIS and 1 record in PERSON. If an enrollee has no diagnoses, the enrollee will have zero records in DIAGNOSIS and 1 record in PERSON.

V. Parameters supplied by the user

The user must set the following parameters when calling macro CY22M07C:

- **INP** – input PERSON SAS® data set name (e.g., *IN1.Person*).
- **IND** – input DIAGNOSIS SAS® data set name (e.g., *IN2.Diagnosis*).
- **INN** – input NDC SAS® data set name (e.g., *IN3.NDC*).
- **INH** – input HCPCS SAS® data set name (e.g., *IN4.HCPCS*).
- **OUTDATA** – output SAS® data set name (e.g., *OUT.OutputScores*).
- **IDVAR** – name of the person identifier variable (e.g., *ID, or EnrolleeID*). This variable can be either character or numeric type, and any length.
- **KEEPVAR** – variables written to the output data set. There is a list of KEEP variables in the program, but the user can alter the list (e.g., *DOB, AGE_LAST, SEX, METAL, CSR_INDICATOR, SCORE_:, CSR_ADJ_SCR_:, or _ALL_*).
- **CCFMT0Y1 –** format name suffix for formats that crosswalk ICD-10 codes to HHS-CCs for fiscal year 2022. For this version of the software it is *HHS_V07FY22R141C*.
- **CCFMT0Y2 –** format name suffix for formats that crosswalk ICD-10 codes to HHS-CCs for fiscal year 2023. For this version of the software it is *HHS_V07FY23R141C.*
- **RXCFMTN –** format name for format that crosswalks NDC codes to RXC for calendar year 2022. For this version of software, it is *NDCV2212_RXCV8_4F.*
- **RXCFMTH –** format name for format that crosswalks HCPCS codes to RXC for calendar year 2022. For this version of software, it is *HCPC2212_RXCV8_4F.*
- **AGEFMT0 –** format name suffix for formats that crosswalk ICD-10 codes to an acceptable age range when MCE edits on ICD-10 codes are performed. For this version of the software it is *AGECY22MCE.*
- **SEXFMT0 –** format name suffix for formats that crosswalk ICD-10 codes to an acceptable sex value when MCE edits on ICD-10 codes are performed. For this version of the software it is *SEXCY22MCE.*

VI. Variables output by the software

The software generates a person-level output SAS® data set. As noted, the user can specify variables to KEEP in the **KEEPVAR** parameter of the macro CY22M07C call.

The following variables can be specified:

1. Any person-level variable from the original PERSON data set.
2. Demographic age/sex variables created by the software:

```
AGE0_MALE       AGE1_MALE

MAGE_LAST_2_4   MAGE_LAST_5_9   MAGE_LAST_10_14 MAGE_LAST_15_20
MAGE_LAST_21_24 MAGE_LAST_25_29 MAGE_LAST_30_34 MAGE_LAST_35_39
MAGE_LAST_40_44 MAGE_LAST_45_49 MAGE_LAST_50_54 MAGE_LAST_55_59
MAGE_LAST_60_GT

FAGE_LAST_2_4   FAGE_LAST_5_9   FAGE_LAST_10_14 FAGE_LAST_15_20
FAGE_LAST_21_24 FAGE_LAST_25_29 FAGE_LAST_30_34 FAGE_LAST_35_39
FAGE_LAST_40_44 FAGE_LAST_45_49 FAGE_LAST_50_54 FAGE_LAST_55_59
FAGE_LAST_60_GT
```

3. CCs created by the software (before hierarchies are applied).
4. HCCs created by the software (after hierarchies are applied).
5. HCC groups and HCC interactions created by the software.
6. RXCs created by the software (after hierarchies are applied).
7. RXC*HCC interactions created by the software.
8. Adult models enrollment duration indicators (ED_1–ED_11) created by the software.
9. Infant model maturity categories, severity level categories, and maturity by severity level interactions created by the software.
10. Score variables created by the software:
    a. Adult Models
        i. SCORE_ADULT_PLATINUM
        ii. SCORE_ADULT_GOLD
        iii. SCORE_ADULT_SILVER
        iv. SCORE_ADULT_BRONZE
        v. SCORE_ADULT_CATASTROPHIC
    b. Child Models
        i. SCORE_CHILD_PLATINUM
        ii. SCORE_CHILD_GOLD
        iii. SCORE_CHILD_SILVER
        iv. SCORE_CHILD_BRONZE
        v. SCORE_CHILD_CATASTROPHIC
    c. Infant Models
        i. SCORE_INFANT_PLATINUM
        ii. SCORE_INFANT_GOLD
        iii. SCORE_INFANT_SILVER
        iv. SCORE_INFANT_BRONZE
        v. SCORE_INFANT_CATASTROPHIC
11. CSR-adjusted score variables:
    a. Adult model
        i. CSR_ADJ_SCR_ADULT_PLATINUM
        ii. CSR_ADJ_SCR_ADULT_GOLD
        iii. CSR_ADJ_SCR_ADULT_SILVER

iv. CSR_ADJ_SCR_ADULT_BRONZE
v. CSR_ADJ_SCR_ADULT_CATASTROPHIC

b. Child model
i. CSR_ADJ_SCR_CHILD_PLATINUM
ii. CSR_ADJ_SCR_CHILD_GOLD
iii. CSR_ADJ_SCR_CHILD_SILVER
iv. CSR_ADJ_SCR_CHILD_BRONZE
v. CSR_ADJ_SCR_CHILD_CATASTROPHIC

c. Infant model
i. CSR_ADJ_SCR_INFANT_PLATINUM
ii. CSR_ADJ_SCR_INFANT_GOLD
iii. CSR_ADJ_SCR_INFANT_SILVER
iv. CSR_ADJ_SCR_INFANT_BRONZE
v. CSR_ADJ_SCR_INFANT_CATASTROPHIC

12. Final unadjusted and CSR-adjusted score variables depending on the enrollee's metal (plan benefit) level and CSR indicator, including enrollment in premium assistance Medicaid alternative plans, created by the software.
    a. Adult scores
        i. SCORE_ADULT
        ii. CSR_ADJ_SCR_ADULT
    b. Child scores
        i. SCORE_CHILD
        ii. CSR_ADJ_SCR_CHILD
    c. Infant scores
        i. SCORE_INFANT
        ii. CSR_ADJ_SCR_INFANT

The user must determine which of the scores is appropriate for the enrollee, depending upon the enrollee's age and plan benefit design of that enrollee.

VII. Computing platforms

The software has been tested using SAS® v9 on this platform:

- Linux (server)

VIII. Steps

1. Install software:
    - Copy files to the computing platform on which the risk scores will be calculated. If the platform is z/OS, upload the two transport files (.TRN) using RECFM(F or FB) LRECL(80) BLKSIZE(8000).
    - Use files with .SAS extensions. Files with .TXT extensions are identical, but might be more easily viewed by the user. File names are case sensitive on some computing platforms; software modules assume that file names are upper case (e.g., I0V07ED1.SAS).
2. Prepare software-provided SAS® input format library and coefficients data set:

- Convert both .TRN files (containing the SAS® format library and model coefficients data set) using SAS® PROC CIMPORT on the computing platform on which the risk scores will be calculated as described in Section II.
- The format library and coefficients data set are provided with the software, but must be imported by the user; they are not imported by the risk adjustment modeling software.

3. Prepare user-provided SAS® input data sets:
   - Create PERSON, DIAGNOSIS, NDC, and HCPCS data sets using the guidelines in Section III and data set descriptions in Section IV.
   - These data sets are created by the user; they are not created by the risk adjustment modeling software.
4. Generate scores:
   - Set parameters as described in Section V.
   - Execute SAS® program CY22P07C and generate variables described in Section VI.
5. Review errors/warnings, notes: the software prints messages in the "Errors/warnings/notes log" for various situations. The user may print (or suppress printing) any of them. To print messages of type `nn`, set macro variable MSGnn to blank; e.g., `%let MSG01= ; .` To suppress printing messages of type `nn`, set macro variable MSGnn to *; e.g., `%let MSG01=*; .`

We recommend the following be printed because they indicate possible errors in data sets, variables or variable values:

```
ERROR  : [Msg01] Variable --- is not in --- file
ERROR  : [Msg02] User-provided variable --- in --- file must be --- type
ERROR  : [Msg03] Duplicate IDVARs in PERSON file
ERROR  : [Msg04] Program halted due to duplicate IDVARs in PERSON file
OK     : [Msg05] PERSON file is free of duplicate IDVARs
ERROR  : [Msg06] Program halted due to non-existent variable(s) in PERSON file
OK     : [Msg07] PERSON file contains all requisite variables
ERROR  : [Msg08] Program halted due to incorrect user-provided variable type(s) in PERSON file
OK     : [Msg09] PERSON file`s variables have the correct type
ERROR  : [Msg10] Program halted due to non-existent variable(s) in DIAG file
OK     : [Msg11] DIAG file contains all requisite variables
ERROR  : [Msg12] Program halted due to incorrect user-provided variable type(s) in DIAG file
OK     : [Msg13] DIAG file`s variables have the correct type
WARNING: [Msg14] Diagnosis matches no enrollee, diagnosis ignored
WARNING: [Msg15] Blank diagnosis code, diagnosis ignored
WARNING: [Msg18] Missing IDVAR, enrollee rejected
WARNING: [Msg19] Invalid SEX, enrollee rejected
WARNING: [Msg20] Invalid DOB, enrollee rejected
WARNING: [Msg21] Invalid AGE_LAST, enrollee rejected
WARNING: [Msg22] Invalid METAL, enrollee rejected
WARNING: [Msg23] Invalid CSR_INDICATOR, enrollee rejected
WARNING: [Msg24] Failed HHS HCC filter, enrollee rejected
WARNING: [Msg27] Invalid DIAGNOSIS_SERVICE_DATE, diagnosis ignored
WARNING: [Msg28] Invalid AGE_AT_DIAGNOSIS, diagnosis ignored
WARNING: [Msg29] AGE_AT_DIAGNOSIS > AGE_LAST, diagnosis ignored
ERROR  : [Msg30] Program halted, file --- does not exist
WARNING: [Msg31] AGE_LAST minus AGE_AT_DIAGNOSIS > 1, diagnosis ignored
WARNING: [Msg32] DOB > DIAGNOSIS_SERVICE_DATE, diagnosis ignored
WARNING: [Msg33] Invalid ENROLDURATION, enrollee rejected
ERROR  : [Msg34] Program halted due to non-existent variable(s) in NDC file
OK     : [Msg35] NDC file contains all requisite variables
ERROR  : [Msg36] Program halted due to incorrect user-provided variable type(s) in NDC file
OK     : [Msg37] NDC file`s variables have the correct type
WARNING: [Msg38] NDC matches no enrollee, NDC ignored
```

```
WARNING: [Msg39] Blank NDC code, NDC ignored
ERROR  : [Msg41] Program halted due to non-existent variable(s) in HCPCS file
OK     : [Msg42] HCPCS file contains all requisite variables
ERROR  : [Msg43] Program halted due to incorrect user-provided variable type(s) in HCPCS file
OK     : [Msg44] HCPCS file`s variables have the correct type
WARNING: [Msg45] HCPCS matches no enrollee, HCPCS ignored
WARNING: [Msg46] Blank HCPCS code, HCPCS ignored
```

We recommend the following be printed during testing with small data sets. The user may choose to suppress printing the messages during production runs with large data sets as these conditions tend to generate many messages.

```
WARNING: [Msg16] Diagnosis lookup failed, diagnosis ignored
NOTE   : [Msg17] Enrollee has no diagnoses, risk score based on remaining information
WARNING: [Msg25] Possible bundled mother/infant claim(s) -- ---
WARNING: [Msg40] NDC lookup failed, NDC ignored
WARNING: [Msg47] HCPCS lookup failed, HCPCS ignored
```

Suppressing printed output for type `nn` does not affect whether an enrollee record or diagnosis is rejected. I.e., diagnosis code ZZZZZ will be ignored by the software even if `%let MSG16=*;` is set.

End of Document