

**HHS-Developed Risk Adjustment Model Algorithm “Do It Yourself (DIY)”**  
**Software Instructions for the 2021 Benefit Year**  
**August 3, 2021 Update<sup>1</sup>**

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 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 plans inside and outside Exchanges.

The HHS risk adjustment methodology is 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. The HHS risk adjustment methodology for the 2021 benefit year is described in the HHS Notice of Benefit and Payment Parameters for 2021 final rule (85 FR 29164) (2021 Payment Notice final rule), which was published in the *Federal Register* on May 14, 2020.<sup>2</sup> The 2021 benefit year implements a new Version 07 (V07) HHS-HCC classification. Changes to the previous Version 05 (V05) classification, referred to as Version 06a, were initially detailed in a technical paper posted June 17, 2019: Potential Updates to HHS-HCCs for the HHS-operated Risk Adjustment Program.<sup>3</sup> HHS conducted further analyses and applied additional classification changes, referred to as Version 06b, which were described in the HHS Notice of Benefit and Payment Parameters for 2021 proposed rule (85 FR 7088).<sup>4</sup> Following input from the public on proposed changes and internal review and analysis, HHS finalized the V07 HHS-HCC classification with certain modifications from Version 06b in the 2021 Payment Notice final rule. A draft V07 ICD-10 to HHS-HCC Crosswalk for benefit year 2021 was posted May 12, 2020,<sup>5</sup> and the final V07 classification changes were presented in the 2021 Payment Notice final rule.<sup>2</sup> In addition to classification changes, the 2021 benefit year risk adjustment model was recalibrated using blended coefficients from the 2016, 2017, and 2018 enrollee-level External Data Gathering Environment (EDGE) data. As described in the 2021 Payment Notice final rule, the HHS risk adjustment methodology gradually transitioned from using MarketScan<sup>®</sup> commercial database, and the 2021 benefit year is the first year in which no

---

<sup>1</sup> This document provides instructions for the HHS risk adjustment models for the 2021 benefit year, with revisions from the software instructions posted on the CCIIO website on April 16, 2021, available at: <https://www.cms.gov/files/document/cy2020-diy-instructions04132021.pdf>.

<sup>2</sup> The final 2021 benefit year risk adjustment model coefficients were provided in the final rule (see the Patient Protection and Affordable Care Act; HHS Notice of Benefit and Payment Parameters for 2021; Final Rule; 85 FR 29164 [May 14, 2020], available at: <https://www.govinfo.gov/content/pkg/FR-2020-05-14/pdf/2020-10045.pdf>).

<sup>3</sup> Potential Updates to HHS-HCCs for the HHS-operated Risk Adjustment Program, posted on the CMS website on June 17, 2019, available at: <https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Potential-Updates-to-HHS-HCCs-HHS-operated-Risk-Adjustment-Program.pdf>.

<sup>4</sup> See the Patient Protection and Affordable Care Act; HHS Notice of Benefit and Payment Parameters for 2021; Proposed Rule; 85 FR 7088 (February 6, 2020), available at: <https://www.govinfo.gov/content/pkg/FR-2020-02-06/pdf/2020-02021.pdf>.

<sup>5</sup> The draft V07 crosswalk included ICD-10 diagnosis codes through the most recent fiscal year (2020) at time of posting, May 12, 2020, available at: <https://www.cms.gov/files/document/draft-2021-update-icd-10-crosswalk-hhs-hcc-risk-adjustment-model.xlsx>.

MarketScan<sup>®</sup> data is used in the blended coefficients.<sup>6</sup> The high-cost risk pool calculation incorporated into the HHS risk adjustment methodology since the 2018 benefit year continued with the same parameters for the 2021 benefit year.

The methodology that HHS will use when operating a risk adjustment program on behalf of a State for the 2021 benefit year<sup>7</sup> 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 in order to 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 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 2021 benefit year:

- Is developed on enrollee-level EDGE data, which directly reflects claims data for ACA individual and small group 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 risk adjustment population;
- Includes a selected number of Prescription Drug Categories (RXC) 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 market;
- Results in state transfers that net to zero within a State market risk pool;
- Adjusts state 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 actuarial risk.

#### Key Revisions in 2021:

- (August 2021 Revisions) Updated Table 1 to include revised references to row numbers in Table 8 (see below), corresponding with edits made to address changes from V05 to V07 of the HHS-HCC infant risk adjustment models.
- (August 2021 Revisions) Updated Table 2 to add 2021 CPT/HCPCS codes used for diagnosis filtering, as described in Section II. Table 2 includes review of 2021 quarterly updates with effective dates as of April 1, 2021. Replaced the 2019 column of code information with 2020 codes (used for historical data purposes). (Table 2 will be updated again in late 2021 to

---

<sup>6</sup> Updates to data used for risk adjustment model recalibration are described in the 2021 Payment Notice final rule, 85 at 29173, available at: <https://www.govinfo.gov/content/pkg/FR-2020-05-14/pdf/2020-10045.pdf>.

<sup>7</sup> HHS will operate risk adjustment for the 2021 benefit year in all 50 states and the District of Columbia.

- include CPT/HCPCS quarterly updates through October 1, 2021.)<sup>8</sup>
- (August 2021 Revisions) Revised Table 3 ICD-10 to HHS-Condition Categories (CC) Crosswalk to reflect new V07 classification with updated Fiscal Year (FY) 2021 ICD-10 diagnosis codes effective January 1, 2021. Revised FY2021 Medicare Code Editor (MCE) code edits to correspond to V07. Revised explanatory text in Sections II and V to clarify that FY2021 ICD-10 diagnosis codes and FY2021 MCE edits should be used in 2021 benefit year risk adjustment. Although codes and MCE edits from the previous fiscal year are usually included for historical purposes, this table does not include FY2020 information due to the change to the V07 classification. Calendar Year (CY) 2021 is the initial implementation year for the V07 HHS-HCC risk adjustment model. Removed the Additional CC column and added two new columns, Second CC and Third CC, to account for ICD-10 code mappings to more than 2 HCCs. (Table 3 will be updated in late 2021 to include FY2022 ICD-10 diagnosis codes and FY2022 and CY2021 MCE code edits.)
  - (August 2021 Revisions) Updated Table 4 to include hierarchies and HCC numbers and labels consistent with changes from the V05 to the V07 HHS-HCC risk adjustment models.
  - (August 2021 Revisions) Revised Table 6 to include HCC groups and the severe illness indicator, interactions, and group for the V07 HHS-HCC adult risk adjustment models.
  - (August 2021 Revisions) Revised Table 7 to include HCC groups for the V07 HHS-HCC child risk adjustment models.
  - (August 2021 Revisions) Updated Table 8 to include changes in HCCs used in infant severity categories for the V07 HHS-HCC infant risk adjustment models.
  - (August 2021 Revisions) Updated Table 9 Model Factors for the 2021 benefit year (using 2016, 2017, and 2018 EDGE data).
  - (August 2021 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, 2021.
  - (August 2021 Revisions) Added a new table, Table 12, Summary of V07 HHS-HCC Model Exclusions—Adult, Child, Infant, to show the payment HCCs excluded from each of the V07 HHS-HCC risk adjustment models. Information provided in Table 12 is not required to compute predicted risk scores and does not impact the algorithm; it is included for reference purposes only.

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 simulate the calculation of risk scores given individual diagnoses. Please direct questions regarding these instructions to HHS HCC Risk Adjustment Models at

---

<sup>8</sup> As a result of the Coronavirus (COVID-19) public health emergency (PHE), the CPT/HCPCS list in Table 2 was expanded in Calendar Year (CY) 2020 to include certain telehealth and telephonic service codes that were not previously accepted for HHS-operated risk adjustment. The allowable telehealth services on Table 2 will be accepted for risk adjustment eligible diagnosis filtering for the HHS-operated risk adjustment program applicable for the individual, merged and small group markets for the 2021 benefit year, if the services are otherwise allowable under applicable state law. For more information on the use of telehealth and telephonic services in 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) (<https://www.cms.gov/files/document/RA-Telehealth-FAQ.pdf>), updated May 6, 2021.

[hshccraops@cms.hhs.gov](mailto:hshccraops@cms.hhs.gov). This mailbox will be used only to answer questions pertaining to operations of the HHS risk adjustment models. We look forward to assisting with inquiries pertaining to your risk adjustment program operations using the HHS-HCC risk adjustment models for the 2021 benefit year.

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’ 2021 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. The algorithm software merely provides a simulation tool for issuers to calculate enrollees’ risk scores. Because risk adjustment transfers under the state payment transfer formula 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 2021 benefit year state transfer(s) should do so with caution and in combination with other data.**

### **List of Tables<sup>2</sup>**

Table 1. Model Membership

Table 2. CPT/HCPCS Included List for Diagnosis Code Filtering

Table 3. ICD-10 to HHS-Condition Categories (CC) Crosswalk

Table 4. HHS-Hierarchical Condition Categories (HCC) Hierarchies

Table 5. Age-Sex Variables

Table 6. Additional Adult Variables

Table 7. Additional Child Variables

Table 8. Additional Infant Variables

Table 9. Model Factors

Table 10a. Prescription Drug Categories (RXC) to National Drug Code (NDC) Crosswalk

Table 10b. Prescription Drug Categories (RXC) to Healthcare Common Procedure Coding System (HCPCS) Crosswalk

Table 11. Prescription Drug Categories (RXC) Hierarchies

Table 12. Summary of HHS-HCC Model Exclusions—Adult, Child, Infant

---

<sup>9</sup> 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 the 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. Introduction**

The risk adjustment program established under Section 1343 of ACA transfers funds from lower than average risk plans to higher than average risk plans using a specified methodology. The HHS risk adjustment methodology generally includes risk adjustment models, a determination of plan average actuarial risk, and a calculation of payments and charges.<sup>10</sup> The latter two elements are referred to by HHS as the state payment transfer formula. This document provides a detailed description of the HHS risk adjustment models and enables interested parties to develop software to determine risk scores given individual diagnoses.

This document should be viewed as a supplemental tool to the 2021 benefit year methodology finalized in the HHS Notice of Benefit and Payment Parameters for 2021 final rule (85 FR 29164). Please note that the models described herein were finalized in this final rule.

The following is a description of the HHS risk adjustment models for the 2021 benefit year. This description includes a narrative and accompanying tables for calculating the risk scores from diagnoses in the HHS risk adjustment models. The HHS risk adjustment models calculate risk scores by summing an enrollee's factors (age/sex, HCCs, RXCs, enrollment duration, and interaction terms).<sup>11</sup> This description shows, in detail, how individual diagnoses are assigned to HCCs and NDCs/HCPCS codes are assigned to RXCs, and then allows the user to build individual risk scores from those inputs.

These instructions outline the steps needed to produce individual risk scores from the HHS risk adjustment models.

The HHS risk adjustment methodology includes risk adjustment models based on age, as detailed in Table 1 Model Membership. Adult scores are calculated for enrollees whose age is equal to or greater than 21 using the adult model set. Child scores are calculated for enrollees age 2 – 20 using the child model set. Infant scores are calculated for enrollees age 0 – 1 using the infant model set. Because HHS models predict plan liability, there are 5 models within each age group model set – one each for platinum, gold, silver, bronze, and for the individual market only, catastrophic plans.

The following five sections describe the algorithm for creating risk scores. Any algorithm developed to calculate risk scores should prepare the diagnosis-level file and use it to assign HCCs for each person. This is described in Section II and includes the list of acceptable

---

<sup>10</sup> The HHS-operated risk adjustment program also includes a data collection approach, data validation procedures, and a schedule for program operation. These elements are not discussed in this document.

<sup>11</sup> The RXC and enrollment duration factors only apply to the adult models.

CPT/HCPCS codes listed in Table 2.<sup>12</sup> Additionally, Section II describes how to prepare NDC-level and HCPCS-level files for use in creating RXCs. Section III and Section IV, respectively, identify data inputted to and outputted from the applicable risk adjustment model. Finally, Section V describes the algorithm to produce risk scores (i.e., outputs). First, the algorithm should crosswalk acceptable diagnoses to Condition Categories (CCs). Then the algorithm should create Hierarchical Condition Categories (HCCs) by imposing hierarchies on the CCs. The algorithm should also crosswalk acceptable NDCs and HCPCS codes to RXCs. For persons without claims or encounter records, zeros are assigned to all CCs, HCCs, and RXCs. After HCCs and RXCs are created, the algorithm should compute predicted scores using the factors from the adult, child, and infant HHS risk adjustment regression models. This information is referenced in Tables 3-11. Information provided in Table 12 is not required to compute predicted risk scores and does not impact the algorithm; this table is included for reference purposes only.

The factors from the following regression models are used to produce risk scores:

- Adult regressions
  - Adult platinum
  - Adult gold
  - Adult silver
  - Adult bronze
  - Adult catastrophic
- Child regressions
  - Child platinum
  - Child gold
  - Child silver
  - Child bronze
  - Child catastrophic
- Infant regressions
  - Infant platinum
  - Infant gold
  - Infant silver
  - Infant bronze
  - Infant catastrophic

## **II. Creation of diagnosis-level file, NDC-level file, and HCPCS-level file**

A. Diagnosis-level file. The diagnosis-level input file (Section III) should 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 the claim or record's diagnoses are included in the diagnosis file to be inputted to the algorithm or software. Encounter records normally report dates, provider or bill types, diagnoses and procedures, and other information, though they may not have payment information.

This section on the diagnosis file preparation explains how each record is evaluated to

---

<sup>12</sup> Definitions taken directly from the Current Procedural Terminology (CPT®) codes and the Healthcare Common Procedure Coding System (HCPCS) code set.

determine whether the record's diagnoses are to be used for CC/HCC creation. It is the user's responsibility to create the diagnosis-level file according to the filtering logic below. This document provides filtering instructions and a list of the CY2020 (for historical data purposes) and CY2021 CPT/HCPCS codes that define service or procedure types that identify acceptable sources of diagnoses for risk adjustment. However, the user must create the diagnosis-level file for input to the risk adjustment algorithm; the data set is not created by the algorithm.

**NOTE: CMS stated that 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 diagnosis codes from sources allowable for risk adjustment should be included in the diagnosis-level file. ICD-10 codes that are not listed in Table 3 may be included in the diagnosis-level file but are ignored by the algorithm.<sup>13</sup> 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 2020 (for historical data purposes) and 2021 CPT/HCPCS codes used to define service or procedure types that are acceptable sources of diagnoses for risk adjustment.

- The CPT/HCPCS codes marking services with diagnoses allowable for risk adjustment are listed in column A of Table 2.
- Column B lists the short description of the CPT/HCPCS codes.
- Columns C and D, respectively, indicate whether a CPT/HCPCS code is acceptable in 2020 or 2021.
- Column E identifies applicable footnotes on the CPT/HCPCS codes.
- Notes begin on row 6,590 of the Excel table with the line "Notes:" and should not be imported by any program.

The diagnosis-level input file 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.

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.

---

<sup>13</sup> If the user conducts fiscal year code validity checks described later in this section, only codes valid for risk adjustment will be included in the final diagnosis-level file.

- 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. 761 (community mental health center admit through discharge); or
    - vi. 767 (community mental health center replacement of prior claim); or
    - vii. 771 (federally qualified health center admit through discharge); or
    - viii. 777 (federally qualified health center replacement of prior claim); or
    - ix. 851 (critical access hospital admit through discharge); or
    - x. 857 (critical access hospital 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 III further describes the diagnosis-level input data file. After creating that file, the user will have the variables needed to conduct fiscal year validity checks. 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. For this preliminary 2021 software, the user should use information from the FY2021 columns. ICD-10 diagnosis codes with service dates of January 1, 2021 – September 30, 2021 should have a Y in the Code Valid in FY2021 column; otherwise, the user should exclude them. Although codes from the previous fiscal year are usually included for historical purposes, this table does not include FY2020 codes due to the change to the V07 classification. CY2021 is the initial implementation year for the V07 HHS-HCC risk adjustment model. The 2021 software will be updated later this year to include FY2022 code information corresponding to service dates of October 1, 2021 – December 31, 2021.

**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, each with a separate ID, sex, and age) in order for the diagnoses to be appropriately included in the input dataset and used for appropriately calculating risk scores.**



The user will need to independently create a program to detect any bundled claims and redefine them (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  $\geq$  2 (an age corresponding to the child or adult models; more specifically age should be appropriate for a maternity diagnosis [i.e., 9-64])<sup>14</sup> 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 Table 3, ICD-10 to Condition Categories (CC) Crosswalk, for diagnosis codes corresponding to the completed pregnancy and newborn HCCs.

For purposes of creating the diagnosis-level input file (refer to Section III), if bundled claims will not be detected and redefined, enrollees with bundled claims should be included. If the enrollee is a female with an age appropriate for a maternity diagnosis, then the diagnoses on the bundled claim are assumed to correspond to the mother's enrollee ID. If the enrollee is age 0, then the diagnoses on the bundled claim are assumed to correspond to the infant's enrollee ID. All diagnoses on a bundled claim that could apply to either the mother or the infant (i.e., diagnoses that do not require passing an age/sex edit) are assumed to apply to the enrollee's ID. All HCCs that are valid for the mother or infant (based on diagnoses that pass or do not require age/sex edits) will be used to compute the enrollee's risk score; and all diagnoses that do not pass age/sex edits (i.e., newborn diagnosis codes for the mother; pregnancy diagnosis codes for the infant) will be ignored.<sup>15</sup> Table 1, rows 11-20, provides an algorithm for including these enrollees.

Infants with a record in the person-level file (refer to Section III) that cannot be matched with a claim or who do not have claims will have no diagnoses in the diagnosis-level file. Infants without diagnoses will be assigned to the lowest severity category and the Age 1 maturity category for infants as detailed in Table 8, rows 130, 132, and 137. Age 0 infants

---

<sup>14</sup> Section III of this document identifies the two age variables used in the software and specifies when each is used. Section V includes the detailed algorithm for creation of HCCs. MCE edits specifying appropriate age for maternity diagnoses are included in Table 3.

<sup>15</sup> This software does not flag or otherwise identify diagnoses that do not pass age/sex edits. Software users may want to create a diagnosis-level flag for bundled claims that identifies rejected diagnoses, or create a person-level flag for bundled claims that identifies enrollees with at least one rejected diagnosis submitted.

with diagnoses but who lack a newborn HCC will be assigned to the corresponding severity category and the Age 1 maturity category for infants as detailed in Table 8, rows 5-130 and 137. 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, as detailed in Table 8, row 142.

B. NDC-level file. The National Drug Code (NDC)-level input file (Section III) should include NDCs used for risk adjustment, listed in Table 10a RXC to NDC Crosswalk. Only pharmacy claims (not medical claims) are the acceptable source for NDCs. The user must evaluate each claim to determine whether the claim's NDCs are included in the NDC file to be inputted to the algorithm or software.

The NDCs are to be used for RXC creation. It is the user's responsibility to create the NDC-level file for input to the risk adjustment algorithm; the data set is not created by the algorithm.

The NDC-level input file 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.<sup>16</sup> (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-level file but are ignored by the algorithm and are not included in RXCs for the adult risk adjustment models' risk score calculations.

C. HCPCS-level file. The Healthcare Common Procedure Coding System (HCPCS)-level input file (Section III) should 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-level file (see Part II.A.2.a. and 3.a). 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 file to be inputted in the algorithm or software.

The HCPCS codes in the HCPCS-level input file are to be used for RXC creation. It is the user's responsibility to create the HCPCS-level file for input to the risk adjustment algorithm; the data set is not created by the algorithm.

The HCPCS-level input file should include HCPCS codes from inpatient, outpatient, and professional medical claims with **discharge dates or through dates** within the benefit

---

<sup>16</sup> 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 dataset: <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 dates is the U.S. Food and Drug Administration's Orange Book: <https://www.accessdata.fda.gov/scripts/cder/ob/index.cfm>.

year. HCPCS codes that are not listed in Table 10b may be included in the HCPCS-level file but are ignored by the algorithm and are not included in RXCs for the adult risk adjustment models' risk score calculations.

### III. Input data files for CC/HCC and RXC variable creation

This section describes the four input data files needed to create CC and HCC grouping and regression variables, RXC and RXC\*HCC interaction variables, and all other regression variables: a person-level file, a diagnosis-level file, an NDC-level file, and a HCPCS-level file. It is the responsibility of the user to create these input data files with the variables listed in this section.

#### 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 and described in Table 1, rows 29-42, are the codes that will be used to identify the plan variation.<sup>17</sup> 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
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

<sup>17</sup> We note that Massachusetts CSR variant plans have a state-specific CSR factor table, as discussed in the 2021 Payment Notice final rule, 85 at 29190. In addition to the CSR variants listed with factors of 1.12, plan variants of 04 are also 1.12 in Massachusetts only.

### 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 and in Table 1, rows 22-23. Although ENROLDURATION will only be used to create variables needed for the adult models, we recommend it be constructed for *all* enrollees to maintain consistency in the variables present in the PERSON file. The created ENROLDURATION variable can be ignored for enrollees in the child or infant models.

STEP 2: The monthly enrollment duration indicator variables (ED\_1–ED\_11) will be created only for adult enrollees by the algorithm as described in Section V and in Table 6, rows 92-102.

1. PERSON file – a person-level file of demographic and enrollment information
  - a. ENROLID (or another person identification variable)
    - i. character or numeric type and unique to an individual
  - b. SEX
    - i. one character, 1 = male; 2 = female
  - c. DOB
    - i. 8-character numeric field, provides the enrollee's date of birth
    - ii. Used to calculate AGE\_AT\_DIAGNOSIS for MCE diagnosis code age edits
  - d. AGE\_LAST
    - i. Age as of last day of enrollment in benefit year
    - 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
    - i. Enrollee's metal level
    - ii. Allowable values: platinum, gold, silver, bronze, catastrophic (only 1 of these values)
  - f. CSR\_INDICATOR
    - i. Person-level indicator. Enrollees who qualify for cost-sharing reductions or those enrolled in premium assistance Medicaid alternative plans will be assigned CSR\_INDICATOR = 1-13. Non-CSR recipients will be assigned CSR\_INDICATOR = 0.
      - 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 = Enrollee in Zero Cost Sharing Plan Variation of Platinum Level QHP.
- 5 = Enrollee in Zero Cost Sharing Plan Variation of Gold Level QHP.
- 6 = Enrollee in Zero Cost Sharing Plan Variation of Silver Level QHP.
- 7 = Enrollee in Zero Cost Sharing Plan Variation of Bronze Level QHP.
- 8 = Enrollee in Limited Cost Sharing Plan Variation of Platinum Level QHP.
- 9 = Enrollee in Limited Cost Sharing Plan Variation of Gold Level QHP.
- 10 = Enrollee in Limited Cost Sharing Plan Variation of Silver Level QHP.
- 11 = Enrollee in Limited Cost Sharing Plan Variation of Bronze Level QHP.
- 12 = Enrollee in a Premium Assistance Medicaid Alternative Plan with 94% AV Silver Plan Variation.
- 13 = Enrollee 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. Person-level enrollment duration variable. Although it is for use in adult models only, it is recommended the user create it for all enrollees for consistency in PERSON file preparation. Values will be ignored for enrollees in child or infant models.
- ii. 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. DIAG file – a diagnosis-level input file of diagnoses

- a. ENROLID (or another person identification variable that must be the same as in PERSON file).<sup>18</sup>
  - i. person identifier of character or numeric type and unique to an

---

<sup>18</sup> Please note that in operation, this information cannot include personally identifiable information.

- individual
  - b. DIAG
    - i. ICD-10-CM diagnosis code, 7-character field, no periods, left justified. Codes should be to the greatest level of available specificity. (Age and sex edits for diagnoses are applied in Part V, step 1 to ensure diagnoses are appropriate for the age and sex of the enrollee.) Only diagnoses from allowable sources (Part II) should be included in the diagnosis-level file.
  - c. DIAGNOSIS\_SERVICE\_DATE
    - i. 8-character numeric field, provides the date of the diagnosis
    - ii. As described in Part II, this variable can be used with DIAG and Table 3 to check that a diagnosis code is valid for a given fiscal year.
  - d. AGE\_AT\_DIAGNOSIS
    - i. Age as of the diagnosis service date, calculated using DOB from PERSON file and DIAGNOSIS\_SERVICE\_DATE
    - ii. Used only for MCE diagnosis code age edits
- 3. NDC file – an NDC-level input file of drug codes
  - a. ENROLID (or another person identification variable that must be the same as in PERSON file)
    - i. person identifier of character or numeric type and 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 (Part II) should be included in the NDC-level file.
- 4. HCPCS file – a HCPCS-level input file of drug codes
  - a. ENROLID (or another person identification variable that must be the same as in PERSON file)
    - i. Person identifier of character or numeric type and 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 (Part II) and with discharge or through dates in the benefit year (Part II) should be included in the HCPCS-level file.

#### **IV. Variables output by the algorithm**

This section describes the variables that are outputted by the HHS risk adjustment model. Any software or algorithms developed from the tables output a person-level file and need to be capable of adding any person-level variables that the user wants to keep.

The following variables can be specified:

1. Any person-level variables from the original person-level file
2. Demographic age/sex variables created by the algorithm
3. HCCs created by the algorithm
4. CCs (condition categories assigned before hierarchies are applied)
5. HCC groups and HCC interactions created by the algorithm
6. RXCs and RXC interactions created by the algorithm
7. Infant model maturity categories, reassigned age/sex variables, severity level categories, and maturity by severity level interactions created by the algorithm
8. Adult model enrollment duration variables created by the algorithm
9. Score variables:
  - 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
10. CSR-adjusted score variables:
  - a. Adult Models
    - i. CSR\_ADJUSTED\_SCORE\_ADULT\_PLATINUM
    - ii. CSR\_ADJUSTED\_SCORE\_ADULT\_GOLD
    - iii. CSR\_ADJUSTED\_SCORE\_ADULT\_SILVER
    - iv. CSR\_ADJUSTED\_SCORE\_ADULT\_BRONZE
    - v. CSR\_ADJUSTED\_SCORE\_ADULT\_CATASTROPHIC
  - b. Child Models
    - i. CSR\_ADJUSTED\_SCORE\_CHILD\_PLATINUM
    - ii. CSR\_ADJUSTED\_SCORE\_CHILD\_GOLD
    - iii. CSR\_ADJUSTED\_SCORE\_CHILD\_SILVER
    - iv. CSR\_ADJUSTED\_SCORE\_CHILD\_BRONZE
    - v. CSR\_ADJUSTED\_SCORE\_CHILD\_CATASTROPHIC
  - c. Infant Models
    - i. CSR\_ADJUSTED\_SCORE\_INFANT\_PLATINUM
    - ii. CSR\_ADJUSTED\_SCORE\_INFANT\_GOLD
    - iii. CSR\_ADJUSTED\_SCORE\_INFANT\_SILVER
    - iv. CSR\_ADJUSTED\_SCORE\_INFANT\_BRONZE

- v. CSR\_ADJUSTED\_SCORE\_INFANT\_CATASTROPHIC
- 11. Final unadjusted and CSR-adjusted score variables depending on the enrollee’s metal level and CSR indicator, including enrollment in premium assistance Medicaid alternative plans.
  - a. Adult scores
    - i. SCORE\_ADULT
    - ii. CSR\_ADJUSTED\_SCORE\_ADULT
  - b. Child scores
    - i. SCORE\_CHILD
    - ii. CSR\_ADJUSTED\_SCORE\_CHILD
  - c. Infant scores
    - i. SCORE\_INFANT
    - ii. CSR\_ADJUSTED\_SCORE\_INFANT

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

#### V. Algorithm to produce output

This section describes the process by which the outputs identified in Section IV are created. The algorithm to produce risk scores can be segmented into five parts—CC, HCC, and RXC creation model; variable creation; access factors table; unadjusted score calculation; and CSR-adjusted score calculation.

Step 1A: CC and HCC creation. This first step should use Table 3 for CC creation and Table 4 for HCC creation.

Table 3 crosswalks ICD-10 codes to the Condition Categories (CCs) in the risk adjustment models. This table incorporates the ICD-10 Medicare Code Edits (MCEs) for age and sex, the further specified CC age and sex splits, and the creation of CCs and additional CCs.<sup>19</sup> Explanations of the fields in Table 3 are provided below:

- OBS (column A): Observations, numeric count from 1 to 10,816 of the code list.
- ICD10 (column B): Only ICD-10 codes assigned to HCCs in the risk adjustment models are included in this crosswalk. All other ICD-10 codes correspond to HCCs not included in the risk adjustment models and are not used to calculate risk scores.
- ICD10 Label (column C): Full ICD-10 code labels.
- Code Valid column (column D): Identifies the fiscal year (column D FY2021) in which the ICD-10 codes are valid. There are two possible values in column D:
  - Y = Yes, code is valid ICD-10 code in that fiscal year
  - N = No, code is not valid ICD-10 code in that fiscal year; code may not yet be in existence or code was deleted in that fiscal year; the code should not be used for risk adjustment in that fiscal year period
- FY2021 MCE Age Condition [use AGE\_AT\_DIAGNOSIS] (column E): Checks

---

<sup>19</sup> The CC age and sex splits relate to how specific conditions are reclassified by age or sex within the full HCC classification. These reassignments are not related to age or sex coding guidelines.



that a person with a diagnosis code has the appropriate age for the diagnosis. (Table 3 will be updated later in 2021 to include an FY2022 MCE Age Condition column and a CY2021 [FY21/FY22] MCE Age Condition column.) Conditions listed in this column are required for the CC to be assigned.

- Newborn diagnosis: Age of 0 years
- Pediatric diagnosis: Age range is 0-17 years inclusive
- Maternity diagnosis: Age range is 9-64 years inclusive
- Adult diagnosis: Age range is 15 years or older
- FY2021 MCE Sex Condition (column F): Checks that a person with a diagnosis code has the appropriate sex for the diagnosis. (Table 3 will be updated later in 2021 to include an FY2022 MCE Sex Condition column and a CY2021 [FY21/FY22] MCE Sex Condition column.) Conditions listed in this column are required for the CC to be assigned.
  - Female indicates the diagnosis is only appropriate for females.
  - Male indicates the diagnosis is only appropriate for males.
- CC Age Split [use AGE\_LAST] (column G): Identifies diagnosis codes whose assignments to CCs in the risk adjustment models change depending on the age of the enrollee. ICD-10 codes assigned by age to more than one CC in the risk adjustment models are listed more than once. In those cases, the age splits are mutually exclusive and enrollees are assigned to only one CC for the given diagnosis.
- CC Sex Split (column H): Identifies diagnosis codes whose assignments to CCs in the risk adjustment models change depending on the sex of the enrollee. ICD-10 codes assigned by sex to more than one CC in the risk adjustment models are listed more than once. Enrollees are assigned to only one CC for the given diagnosis.
- CC (column I), Second CC (column J), and Third CC (column K): CC (column I) identifies the CC in the risk adjustment models to which a person with the ICD-10 code who meets the MCE conditions and CC splits is assigned. Additional CCs (columns J and K) are assigned when the ICD-10 code is assigned to more than one HCC in the risk adjustment models because it contains more than one clinical concept (such as an ICD-10 diagnosis code that encompasses both diabetes and a complication of diabetic peripheral angiopathy with gangrene). Second CC (column J) identifies the second CC, and Third CC (column K) identifies the third CC, in the risk adjustment model to which a person with the ICD-10 code who meets the MCE conditions and CC Age or Sex splits is assigned.
- Footnote (column L): Identifies applicable footnotes for the ICD-10 code. Notes begin on row 10,822 of the Excel table with the line “Notes:” and should not be imported by any program.

Table 4 provides the hierarchy rules to apply to the CCs to create HCCs. Explanations of the fields in Table 4 are provided below:

- OBS (column A): Observations, numeric count from 1 to 141 of HCCs in the HHS risk adjustment models.
- V07 HCC (column B): Only HCCs in the risk adjustment models are identified. HCCs are created when hierarchies are applied to CCs.

- Set to 0 HCCs (column C): Hierarchy exclusions for each HCC are listed. Users should set the HCCs in this column to 0 when a person has the HCC listed in column B.
- HCC Label (column D): Full HCC labels.

Use Table 3 and Table 4 to create CCs and HCCs:

1. Create CC and HCC 0/1 indicator variables for each HCC in the risk adjustment models listed in Table 4 (Example: CC001, CC002, CC003, CC004, CC006, CC008 ... CC254; HHS\_HCC001, HHS\_HCC002 ... HHS\_HCC254). Initially, set CCs and HCCs to 0 for each person.
2. If there are any acceptable diagnoses for a person, then:
  - a. Perform ICD-10 Medicare Code Edits (MCEs) for age and sex (use FY2021 MCE Age Condition and MCE Sex Condition columns in Table 3). Use AGE\_AT\_DIAGNOSIS for applying MCE age edits.
  - b. If a person meets the MCE conditions, apply CC Age Split and CC Sex Split to create CC (set to 1) using crosswalk of ICD-10 to CCs (use CC Age Split, CC Sex Split, and CC columns in Table 3). Use AGE\_LAST for CC age splits.
  - c. Create additional CCs (set to 1) using Second CC and Third CC columns.
3. Create HCCs after all diagnoses for a person are processed and CC variables are created using hierarchies provided in Table 4.
  - a. Set HCC to 1 when corresponding CC is set to 1.
  - b. Set the HCCs in column C to 0 when a person has the HCC listed in column B set to 1.

Step 1B: RXC creation. This step should use Tables 10a/10b and Table 11 for RXC creation.

Table 10a crosswalks National Drug Codes (NDCs) to the Prescription Drug Categories (RXC) in the adult risk adjustment models. This table identifies the subset of NDCs used in the creation of RXCs. Explanations of the fields in Table 10a are provided below:

- RXC (column A): Identifies the RXC number in the risk adjustment models that a person with the NDC is assigned.
  - RXCs 1-10 are included in the adult risk adjustment models as stand-alone variables and as interactions, each with risk factors included in a person's risk score.
- RXC\_Label (column B): Provides the label of each RXC.
- NDC (column C): Only NDCs assigned to RXCs in the adult risk adjustment models are included in this crosswalk. All other NDCs do not map to an RXC and are not used to calculate risk scores. NDCs are stored in string (character) to avoid dropping of leading zeroes.

Table 10b crosswalks Healthcare Common Procedure Coding System (HCPCS) codes to the Prescription Drug Categories (RXC) in the adult risk adjustment models. This table identifies the subset of HCPCS codes used in the creation of RXCs. Explanations of the fields in Table 10b are provided below:

- RXC (column A): Identifies the RXC number in the risk adjustment models

that a person with the NDC is assigned. (See Table 10a description above for information about RXCs 1-10.)

- RXC\_Label (column B): Provides the label of each RXC.
- HCPCS (column C): Only HCPCS codes assigned to RXCs in the adult risk adjustment models are included in this crosswalk. All other HCPCS drug codes do not map to an RXC and are not used to calculate risk scores.

Table 11 provides the hierarchy rules to apply to the RXCs. Explanations of the fields in Table 11 are provided below:

- RXC (column A): RXCs in the risk adjustment models are identified.
- Set to 0 RXC (column B): Hierarchy exclusions for each RXC are listed. Users should set the RXCs in column B to 0 when a person has the RXC listed in column A.
- RXC\_Label (column C): Provides the label of each RXC in column A.

Use Tables 10a/10b and Table 11 to create RXCs:

1. Create 10 RXC 0/1 indicator variables for each RXC in the risk adjustment models listed in Table 10a (Example: RXC\_01, RXC\_02, RXC\_03 ... RXC\_09, RXC\_10). Initially, set RXCs to 0 for each person.
2. Using Table 10a, if there are any acceptable NDCs for a person, then set to 1 the corresponding RXC(s) using the RXC and NDC columns in the crosswalk.
3. Using Table 10b, if there are any acceptable HCPCS codes for a person, then set to 1 the corresponding RXC(s) using the RXC and HCPCS columns in the crosswalk.
4. Using hierarchies provided in Table 11, set the RXCs in column B to 0 when a person has the RXC listed in column A set to 1.

The RXC interactions will be created in a later step.

### Step 2: Model variable creation.

This step should use Table 5 for demographic variable creation, Table 6 for additional adult model variables including enrollment duration, Table 7 for additional child model variables, and Table 8 for additional infant model variables.

Explanations of the fields in Tables 5-8 are provided below:

- Column A identifies the model set in which the variable is used (adult, child or infant).
- Column B identifies the variable name.
- Column C gives a variable description.
- Column D indicates if the variable is used in the model.
- Column E defines the variable creation.

Variables must be assigned values in the same order as defined in Tables 5, 6, 7, and 8. For example, in Table 6, assign a value to variable SEVERE\_V3 first, then the HCC group variables, then the severe illness interactions, then the severe illness interaction group.

#### Variable creation:

1. Create demographic variables needed for score calculation using Table 5.
  - a. A person's age and sex are used to create 1 of 28 mutually exclusive age-sex variables (18 for adult models, 8 for child models, 2 for infant models). Use AGE\_LAST.
  - b. Because the infant model is computed with females as the reference group for the age-sex demographic variables, there are no separate age-sex demographic variables for females age 0 or 1. Note that if male age 0 infants are assigned to Age 1 because they lack a newborn HCC, their age/sex demographic variable will be reassigned to Age 1 using Table 8.
  - c. Adult age-sex variables are listed in rows 8-25.
  - d. Child age-sex variables are listed in rows 27-34.
  - e. Infant age-sex variables are listed in rows 36 and 37.
2. For the adult model, create the severe illness indicator variable using Table 6.
  - a. The severe illness indicator variable, SEVERE\_V3, is created when an adult has at least 1 of 8 HCCs that indicate a severe illness. Rows 5-12 of column E in Table 6 identify these 8 HCCs.
    - i. SEVERE\_V3 is initially set to 0 for each adult and is set equal to 1 if the adult has at least 1 of the 8 HCCs.
3. For the adult and child models – create HCC groups using Tables 6 and 7.
  - a. HCC groups are effectively treated as single HCCs in the models. The HCC groups are assigned variables GXX, where XX is the group number.
  - b. For the adult models:
    - i. There are 18 adult HCC groups. The group variable names are listed in rows 14-69 of column B of Table 6.
    - ii. Initially create the 18 HCC groups for adults and set to 0 for each person.
    - iii. The HCCs in an HCC group are listed in column E of Table 6 starting in row 14.
    - iv. A variable for an HCC group is set to equal 1 if an adult has at least one of the HCCs in a given adult HCC group.
  - c. For the child models:
    - i. There are 21 child HCC groups. The HCC group variable names are listed in rows 5-67 of column B of Table 7.
    - ii. Initially create the 21 HCC groups for children and set to 0 for each person.
    - iii. The HCCs in an HCC group are found in column E of Table 7 starting in row 5.
    - iv. A variable for an HCC group is set to equal 1 if a child has at least one of the HCCs in a given child HCC group.
4. For the adult models – create HCC interaction variables using Table 6.
  - a. The variable SEVERE\_V3 is used to create the severe illness indicator in HCC or HCC group interactions. There are 9 interactions identified in rows 72-80 of Table 6.
    - i. Interaction variables are initially set to 0 for each adult.
    - ii. The variable for an interaction is set to 1 if an adult has SEVERE\_V3 = 1 and the HCC or HCC group in the interaction =

- 1.
- b. The interactions are aggregated to create a severe illness interaction group based on the predicted expenditures of the interactions that are in the models.
  - i. Variable INT\_GROUP\_H is set to 0 for each adult.
  - ii. INT\_GROUP\_H is set to 1 if an adult has at least 1 of the 9 high-cost interactions. The variable and interactions are found in rows 82-90 of Table 6.
5. For the adult models – create enrollment duration variables using Table 6.
  - a. The variable ENROLDURATION is used to create the 11 mutually exclusive (0/1) enrollment duration indicators for adult enrollees who have partial year enrollment. The variables are identified in rows 92-102 of Table 6.
    - i. The enrollment duration variables ED\_1–ED\_11 are initially set to 0 for each adult.
    - ii. An enrollment duration variable (ED\_X) is set to 1 if an adult has ENROLDURATION equal to the corresponding number (X). For example, ED\_1 = 1 if ENROLDURATION = 1, ED\_2 = 1 if ENROLDURATION = 2, and so on.
    - iii. Enrollees who have ENROLDURATION = 12 (indicating 12-month enrollment) will have all 11 ED variables set to 0. All other adult enrollees will have exactly 1 ED variable set to 1.
6. For the adult models – create RXC interaction variables using Table 6.
  - a. The HCCs and the RXC variables (RXC\_01-RXC\_10) created earlier in Step 1B are used to create the 13 (0/1) RXC interaction variables. The RXC interaction variables are identified in rows 115-127 of Table 6.
    - i. Create the 13 RXC interaction variables listed in rows 115-127 in column B (e.g., RXC\_01\_X\_HCC001) and initially set to 0 for each adult. These interactions are based on RXCs listed in rows 104-113 of Table 6.
    - ii. An RXC interaction variable (e.g., RXC\_01\_X\_HCC001) is set to 1 if an adult has both the corresponding RXC set to 1 and the HCC(s) set to 1 as listed in column E of Table 6 starting in row 115.
    - iii. There are 5 RXC interactions that include a single HCC (rows 115, 117, 122, 124-125).
    - iv. There are 7 RXC interactions that include a set of HCCs in which only 1 HCC of the set must be set to 1 (rows 116, 118-121, 126-127).
    - v. There is 1 RXC interaction (RXC\_09\_X\_HCC056\_057\_AND\_048\_041 in row 123) that includes two sets of HCCs in which at least 1 HCC in each set must be set to 1.
7. For the infant models – create maturity and severity categories, reassign age-sex variable if needed, and create maturity by severity level interaction variables using Table 8.
  - a. Table notation: Severity category and maturity category variables (which are not included in the regression, but are used to create regression interaction variables) use the “IHCC\_” prefix designating them as Infant variables. Examples: IHHC\_Severity5; IHCC\_Extremely\_Immature.
  - b. Create 5 severity level category 0/1 indicator variables based on HCC flags (See Table 8, rows 5-124). Set to 1 if any of the corresponding HCC variables

is set to 1.

- i. Severity Level 5 (Highest)
  - ii. Severity Level 4
  - iii. Severity Level 3
  - iv. Severity Level 2
  - v. Severity level 1 (Lowest)
- c. Apply hierarchies so that each age 0 or age 1 infant has only a single severity level variable, the highest qualifying level, set to 1 (see Table 8, rows 126-130). Note that infants lacking any of the HCCs corresponding to severity levels are assigned to Severity Level 1 (Lowest).
- d. Create 5 maturity category 0/1 indicator variables set to 1 based on age only if age 1, and on age as well as newborn HCC flags if age 0 (see Table 8, rows 132-137). Note that age 0 infants lacking a newborn HCC (242-249) are assigned to Age 1.
- i. Age 1
  - ii. Extremely Immature
  - iii. Immature
  - iv. Premature/Multiples
  - v. Term
- e. Apply hierarchies so that each age 0 or age 1 infant has only a single maturity variable set to 1 (see Table 8, rows 138-140).
- f. Reassign male age 0 infants lacking a newborn HCC to the Age 1 age-sex variable (see Table 8, row 142).
- g. Create 25 maturity by severity level interactions (0/1) using the 5 maturity category variables and 5 severity level category variables (see Table 8, rows 144-168). These are mutually exclusive interactions—each infant will have only one maturity by severity level interaction set to 1.
- i. Extremely\_Immature\_x\_Severity5
  - ii. Extremely\_Immature\_x\_Severity4
  - iii. Extremely\_Immature\_x\_Severity3
  - iv. Extremely\_Immature\_x\_Severity2
  - v. Extremely\_Immature\_x\_Severity1
  - vi. Immature\_x\_Severity5
  - vii. Immature\_x\_Severity4
  - viii. Immature\_x\_Severity3
  - ix. Immature\_x\_Severity2
  - x. Immature\_x\_Severity1
  - xi. Premature\_Multiples\_x\_Severity5
  - xii. Premature\_Multiples\_x\_Severity4
  - xiii. Premature\_Multiples\_x\_Severity3
  - xiv. Premature\_Multiples\_x\_Severity2
  - xv. Premature\_Multiples\_x\_Severity1
  - xvi. Term\_x\_Severity5
  - xvii. Term\_x\_Severity4
  - xviii. Term\_x\_Severity3
  - xix. Term\_x\_Severity2
  - xx. Term\_x\_Severity1

- xxi. Age1\_x\_Severity5
- xxii. Age1\_x\_Severity4
- xxiii. Age1\_x\_Severity3
- xxiv. Age1\_x\_Severity2
- xxv. Age1\_x\_Severity1

Step 3: Access factors in Table 9.

1. For the adult models, use adult factors in rows 4-205 for each of the metal levels (columns D-H).
2. For the child models, use child factors in rows 207-365 for each of the metal levels (columns D-H)
3. For the infant models, use infant factors in rows 367-393 for each of the metal levels (columns D-H)

Step 4: Unadjusted score calculation (using directions in Tables 6-8, variables created in Tables 5-8, and factors from Table 9).

1. For each adult, create an adult model score for the metal level in which that adult is enrolled (corresponding to scores in Part IV).
  - a. Create scores for each metal level using directions in rows 129-137 in Table 6.
    - i. SCORE\_ADULT\_PLATINUM
    - ii. SCORE\_ADULT\_GOLD
    - iii. SCORE\_ADULT\_SILVER
    - iv. SCORE\_ADULT\_BRONZE
    - v. SCORE\_ADULT\_CATASTROPHIC
  - b. The score for each of the above variables is the sum of the factors of the corresponding metal level in Table 9 for all variables set to 1 for that adult: demographic variable, any HCCs, any HCC groups, a severe illness interaction group if applicable, an enrollment duration variable if applicable, RXCs 1-10, and RXC interactions.
    - i. SCORE\_ADULT\_PLATINUM uses factors from column D
    - ii. SCORE\_ADULT\_GOLD uses factors from column E
    - iii. SCORE\_ADULT\_SILVER uses factors from column F
    - iv. SCORE\_ADULT\_BRONZE uses factors from column G
    - v. SCORE\_ADULT\_CATASTROPHIC uses factors from column H
  - c. Identify the metal level in which the adult is enrolled (Table 1 row 27).
  - d. Create an adult model score SCORE\_ADULT for the metal level in which the adult is enrolled using directions in rows 160-161 in Table 6.
2. For each child create a child model score for the metal level in which that child is enrolled (corresponding to scores in Part IV).
  - a. Create scores for each metal level using directions in rows 70-74 in Table 7.
    - i. SCORE\_CHILD\_PLATINUM
    - ii. SCORE\_CHILD\_GOLD
    - iii. SCORE\_CHILD\_SILVER
    - iv. SCORE\_CHILD\_BRONZE

- v. SCORE\_CHILD\_CATASTROPHIC
  - b. The score for each of the above variables is the sum of the factors of the corresponding metal level in Table 9 for all variables set to 1 for that child: demographic variable, any HCCs, and any HCC groups.
    - i. SCORE\_CHILD\_PLATINUM uses factors from column D
    - ii. SCORE\_CHILD\_GOLD uses factors from column E
    - iii. SCORE\_CHILD\_SILVER uses factors from column F
    - iv. SCORE\_CHILD\_BRONZE uses factors from column G
    - v. SCORE\_CHILD\_CATASTROPHIC uses factors from column H
  - c. Identify the metal level in which the child is enrolled (Table 1 row 27).
  - d. Create a child model score SCORE\_CHILD for the metal level in which the child is enrolled using directions in rows 97-98 in Table 7.
- 3. For each infant create an infant model score for the metal level in which that infant is enrolled (corresponding to scores in Part IV) using directions starting at row 170 of Table 8, infant variables created in Tables 5 and 8, and factors from Table 9.
  - a. Unadjusted infant risk score is the sum of (0/1 variable x factor) over all the following:
    - i. 0/1 age-sex categories for INFANT models
    - ii. 0/1 maturity x severity level interactions for INFANT models
  - b. The unadjusted score is the sum of the factors of the corresponding metal level in Table 9 for all variables set to 1 for that infant.
    - i. SCORE\_INFANT\_PLATINUM uses factors from column D
    - ii. SCORE\_INFANT\_GOLD uses factors from column E
    - iii. SCORE\_INFANT\_SILVER uses factors from column F
    - iv. SCORE\_INFANT\_BRONZE uses factors from column G
    - v. SCORE\_INFANT\_CATASTROPHIC uses factors from column H
  - c. Females will have a single maturity level x severity level interaction set to 1 and that factor will be their score. Males will have a single maturity level x severity level interaction set to 1 and a single age-sex variable set to 1; the sum of the factors for those two variables will be their score.
  - d. Identify the metal level in which the infant is enrolled (Table 1 row 27).
  - e. Create an infant model score SCORE\_INFANT for the metal level in which the infant is enrolled using directions in rows 197-198 in Table 8.

Step 5: CSR-adjusted score calculation (using instructions in Tables 6-8).

- 1. For each adult, create an adult model CSR-adjusted score for the metal level in which that adult is enrolled (corresponding to CSR-adjusted scores in Part IV).
  - a. Each adult will have one of the following CSR-adjusted risk scores that corresponds to his/her metal level. CSR-adjusted scores are created using directions in rows 139-158 in Table 6.
    - i. CSR\_ADJUSTED\_SCORE\_ADULT\_PLATINUM
    - ii. CSR\_ADJUSTED\_SCORE\_ADULT\_GOLD
    - iii. CSR\_ADJUSTED\_SCORE\_ADULT\_SILVER
    - iv. CSR\_ADJUSTED\_SCORE\_ADULT\_BRONZE
    - v. CSR\_ADJUSTED\_SCORE\_ADULT\_CATASTROPHIC



- b. Identify the metal level in which the adult is enrolled (Table 1 row 27).
  - c. Create an adult model score CSR\_ADJUSTED\_SCORE\_ADULT for the metal level in which the adult is enrolled using directions in rows 163-165 in Table 6.
- 2. For each child, create a child model CSR-adjusted score for the metal level in which that child is enrolled (corresponding to CSR-adjusted scores in Part IV).
  - a. Each child will have one of the following CSR-adjusted risk scores that corresponds to his/her metal level. CSR-adjusted scores are created using directions in rows 76-95 in Table 7.
    - i. CSR\_ADJUSTED\_SCORE\_CHILD\_PLATINUM
    - ii. CSR\_ADJUSTED\_SCORE\_CHILD\_GOLD
    - iii. CSR\_ADJUSTED\_SCORE\_CHILD\_SILVER
    - iv. CSR\_ADJUSTED\_SCORE\_CHILD\_BRONZE
    - v. CSR\_ADJUSTED\_SCORE\_CHILD\_CATASTROPHIC
  - b. Identify the metal level in which the child is enrolled (Table 1 row 27).
  - c. Create a child model score CSR\_ADJUSTED\_SCORE\_CHILD for the metal level in which the child is enrolled using directions in rows 100-102 in Table 7.
- 3. For each infant, create an infant model CSR-adjusted score for the metal level in which that infant is enrolled (corresponding to CSR-adjusted scores in Part IV).
  - a. Each infant will have one of the following CSR-adjusted risk scores that corresponds to his/her metal level. CSR-adjusted scores are created using directions in rows 176-195 in Table 8.
    - i. CSR\_ADJUSTED\_SCORE\_INFANT\_PLATINUM
    - ii. CSR\_ADJUSTED\_SCORE\_INFANT\_GOLD
    - iii. CSR\_ADJUSTED\_SCORE\_INFANT\_SILVER
    - iv. CSR\_ADJUSTED\_SCORE\_INFANT\_BRONZE
    - v. CSR\_ADJUSTED\_SCORE\_INFANT\_CATASTROPHIC
  - b. Identify the metal level in which the infant is enrolled (Table 1 row 27).
  - c. Create an infant model score CSR\_ADJUSTED\_SCORE\_INFANT for the metal level in which the infant is enrolled using directions in rows 200-202 in Table 8.

**End of Document**