

The following is an overview of software for the CMS-RXHCC risk-adjustment model. The software includes a SAS program - **R0518P1P** that calls several SAS Macros to create 8 RXHCC score variables using coefficients from PDP+MAPD regression models. There are 8 regression models as follows:

- 1) Community, Non-Low Income, Aged, Continuing Enrollee
- 2) Community, Non-Low Income, Non-Aged,¹ Continuing Enrollee
- 3) Community, Low Income, Aged, Continuing Enrollee
- 4) Community, Low Income, Non-Aged, Continuing Enrollee
- 5) Institutional Continuing Enrollee
- 6) Community, Non-Low Income, New Enrollee
- 7) Community, Low Income, New Enrollee
- 8) Institutional New Enrollee

Software description

The software consists of a main program **R0518P1P** that supplies user parameters to the main SAS Macro program **R0518P1M**. This macro program reads in two input files and assigns RXHCCs for each person. First, the program crosswalks diagnoses to Condition Categories (RXCCs) using SAS formats which were previously stored in the FORMAT library. Then the program creates Drug Hierarchical Condition Categories (RXHCCs) by imposing hierarchies on the RXCCs. For persons without claims, zeros are assigned to all RXHCCs. A person may be assigned none, one, or multiple RXHCCs.

After RXHCCs are created, the program computes predicted scores from 8 regression models.

The main macro R0518P1M uses 5 external SAS Macro programs:

- %AGESEXV4 - creates age/sex, originally disabled, and non-aged interaction variables
- %R05I0ED2 - performs edits to ICD10 diagnosis codes
- %R05X76L1 - assigns labels to RXHCCs
- %R05X76H1 - sets RXHCC=0 according to hierarchies
- %SCOREVAR - calculates a score variable

The main program, main macro and 5 external macros have a .txt extension to make the files easier to view. Please rename them to have a .sas extension before running the software.

Steps performed by the software:

- step1: include external macros
- step2: define internal macro variables
- step3: merge person and diagnosis files outputting one record per person for each input person level record
 - step3.1: declaration section
 - step3.2: bring in regression coefficients

¹ The term “non-aged” is used for those younger than 65 because this group includes beneficiaries eligible for Medicare because of end-stage renal disease as well as those eligible because of disability.

step3.3: merge person and diagnosis files
step3.4: for the first record for a person, set RXCCs to 0 and create person's age
step3.5: if there are any diagnoses for a person, then do the following:

- perform diagnosis edits using macro R05I0ED2
- create RXCC using format provided in format library
- create additional RXCC using additional formats provided in format library

step3.6: for the last record for a person do the following:

- create demographic variables needed for score calculation (macro **AGESEXV4**)
- create RXHCCs using hierarchies (macro **R05X76H1**)
- create RXHCC by non-aged interaction variables
- set RXHCCs and interaction variables to zero if there are no diagnoses for a person
- create scores for 5 continuing enrollee models (macro **SCOREVAR**)
- create scores for 3 new enrollee models (macro **SCOREVAR**)

step4: data checks and proc contents

PART 1. Files supplied by the software.

The following SAS programs and files are included in this software:

R0518P1P - main program that has all the parameters supplied by a user (see below for parameter and variable list). It calls main macro R0518P1M.
R0518P1M - main macro that creates RXHCC and SCORE variables by calling other external macros.
AGESEXV4 - creates age/sex, originally disabled, and non-aged interaction variables.
R05I0ED2 - performs edits to ICD10 code if wanted. Medicare Code Editor (MCE) is source of edits.
R05X76L1 - assigns labels to RXHCCs.
R05X76H1 - sets RXHCC=0 according to hierarchies.
SCOREVAR - calculates a score variable.
F0518P1Q.TXT - a txt version of the format that has a cross-walk from ICD10 codes to RXCC categories (use for reference only). This format contains ICD10 codes valid in FY2017/FY2018.
F0518P1Q.TRN - format library containing all the formats necessary for the software.
R0515P2Q.TRN - relative coefficients for 8 regression models created on CY2014/2015 data using denominator \$1,047.96 (CMS 12-12-2016).

The last 2 files are SAS transport files and have the extension .trn. These transport files are special SAS files that may be used on any platform running SAS after uploading and converting using PROC CIMPORT. The user should use the following program to convert them.

Code for converting coefficients transport file to SAS file:

```
filename inc "C:\user defined location of the transport
file\R0515P2Q.TRN";
libname incoef "C:\user defined location of the sas coefficients
file";
```

```
proc cimport data=incoef.rxcoeff infile=inc;
run;
```

Code for converting formats transport file to SAS file:

```
filename inf "C:\user defined location of the transport file\
F0518P1Q.TRN";
libname library "C:\user defined location of the sas formats file";
proc cimport library=library infile=inf;
run;
```

If you are operating in an MVS - z/OS environment, the transport files should be uploaded using the following parameters:
RECFM(F or FB) LRECL(80) BLKSIZE(8000)

PART 2. Files supplied by a user.

Two SAS input files needed for the software must be presorted in ascending order by the person ID variable

- 1) **PERSON** file--a person-level file of demographic and enrollment information
- 2) **DIAG** file--a diagnosis-level input file of diagnoses

Data requirements for the SAS input files: The variable names listed are required by the programs as written:

1) **PERSON** file

- **HICNO** (or other person identification variable. It must be set in the macro variable IDVAR)
-character or numeric type and unique to an individual
- **SEX**
-one character, 1=male; 2=female
- **DOB**
-SAS date format, date of birth
- **OREC**
-one character, original reason for entitlement with the following values:
 - 0 - OLD AGE (OASI)
 - 1 - DISABILITY (DIB)
 - 2 - ESRD
 - 3 - BOTH DIB AND ESRD
- **ESRD**
-numeric, end stage renal disease indicator with the following values for the payment year:
 - 0 - no ESRD
 - 1 - if person is in any of the following statuses:
ESRD dialysis, transplant, post graft.

ESRD variable is needed for New Enrollee models. If missing, the New Enrollee scores for the beneficiary will be missing. Set to 0 if not known to get the non-ESRD score, the most common situation.

- 2) **DIAG** file--a diagnosis file with at least one record per person-specific diagnosis for persons with diagnoses.
- **HICNO** (or other person identification variable that must be the same as in PERSON file)
 - person identifier of character or numeric type and unique to an individual
 - **DIAG**
 - Diagnosis code, 7 character field, no periods, left justified. The user may include all diagnoses or limit the codes to those used by the model. Codes should be to the greatest level of available specificity. Diagnoses should be included **only** from acceptable sources, depending on whether you are using RAPS submission or encounter data.

PART 3. Parameters supplied by a user:

NOTE: All user-supplied parameters should be reentered by the user. The default settings are examples only, and should not be used.

The user must supply the following in the **R0518P1P** program:

- **INP** - SAS input person dataset name
- **IND** - SAS input diagnosis dataset name
- **OUTDATA** - SAS output dataset name
- **IDVAR** - variable name for Beneficiary ID (HICNO for Medicare data)
- **KEEPVAR** - variables kept in the output dataset. There is a list of KEEP variables in the program, but the user can alter the list.
- **SEDITS** - a switch that controls whether to perform MCE edits on ICD10 codes:
 - 1-YES, 0-NO
- **DATE_ASOF**- reference date to calculate age. Set to February 1 of the payment year for consistency with CMS.

PART 4. Variables output by the software.

The software outputs a person level file. Any variables that the user wants to keep in it should be specified in the main program **R0518P1P** in the **KEEPVAR** parameter of macro **R0518P1M** call. The following variables can be specified:

1) Any person level variables from the original person level file

2) Demographic variables created by the software:

AGEF ORIGDS NONAGED

**F0_34 F35_44 F45_54 F55_59 F60_64 F65_69
 F70_74 F75_79 F80_84 F85_89 F90_94 F95_GT
 M0_34 M35_44 M45_54 M55_59 M60_64 M65_69
 M70_74 M75_79 M80_84 M85_89 M90_94 M95_GT**

(These are age/sex variables for continuing enrollees defined in the main program **R0518P1P** by the macro variable **&AGESEXVARS.**)

- 3) **RXHCCs** defined in the main program **R0518P1P** by the macro variable **&RXHCCV5_list76**
- 4) **RXCCs** (condition categories assigned before hierarchies are applied) defined in the main program **R0518P1P** by the macro variable **&RXCCV5_list76**
- 5) Score variables:
 - SCORE_CE_NoLowAged**
 - SCORE_CE_NoLowNoAged**
 - SCORE_CE_LowAged**
 - SCORE_CE_LowNoAged**
 - SCORE_CE_LTI**
 - SCORE_NE_NonLowCommunity**
 - SCORE_NE_LowCommunity**
 - SCORE_NE_LTI**

The user should determine which of the scores is appropriate for the beneficiary depending upon the status of that beneficiary.