Acute Care Capitation Risk Adjustment for Minnesota Special Needs Plans Serving People with Disabilities (SNBC/MnDHO): USERS GUIDE

Prepared by and for more information or assistance contact:

Greg Gifford, Ph.D. or
greg.gifford@health.state.mn.us
Minnesota Department of Health
Health Economics Program
Health Policy Division
85 East Seventh Place, Suite 300
P.O. Box 64975
St. Paul, Minnesota 55101
651-201-3553

Tim Ryan, M.A. or
Timothy.J.Ryan@state.mn.us
Minnesota Department of Human Services
Purchasing and Service Delivery Division
540 Cedar Street
St. Paul, Minnesota 55155
651-431-2503

Jerry Storck, Ph.D.
Jerry.T.Storck@state.mn.us
Minnesota Department of Human Services
Mental Health Division
540 Cedar Street
St. Paul, Minnesota 55155
651-431-2229

CONTENTS

INTRODUCTION

This guide provides instructions for, and a demonstration of the acute care capitation risk adjustment method designed for application to the **Minnesota Special Needs Basic Care (SNBC) project**. The SNBC capitation risk adjustment method consists of two individualized risk adjustment models - one for "MA Only" recipients and one for Medicare/Medicaid "Dual Eligible" recipients. The MA Only and Dual Eligible models are identical with respect to both structure, and selected risk factors. However, they are distinct with respect to the specific values of the risk factor weights. These models are based on regression-based analyses of calendar year (CY) 2005 medical assistance (MA) reimbursements for paid health services per member per month (pmpm) for all persons enrolled in fee-for-service (FFS) MA for one or more months in CY 2005 separately for the 45,642 MA Only and 48,937 Dual Eligible enrollees.[1] All instructions necessary to fully implement these models are supplied in this document.

Equation (1) defines the basic structure and characteristics of the models that are common to both the MA Only and Dual Eligible populations, except where noted.

$$\mathbf{Pred(\$)_i = Pred(SqrRt(\$))_i^2 + Pred(SqrRt(PE\$))_i^2} \quad (1)$$

This equation specifies that predicted CY 2005 reimbursements pmpm for each recipient month record (i) - $Pred(\$)_i$ – is equal to (a) the square of the predicted square root of reimbursements - $Pred(SqrRt(\$))_i^2$, where the predicted square root of reimbursements is predicted by a linear, additive function of a set of risk factors common to both populations; plus (b) the square of the predicted square root of the prediction error from (a) - $Pred(SqrRt(PE\$))_i^2$. $Pred(SqrRt(PE\$))_i^2$ is predicted by an additive function of $Pred(SqrRt(\$))_i^2$ and $Pred(SqrRt(\$))_i$.

The function by which $Pred(SqrRt(\$))_i$ is calculated for each recipient is a linear, additive function of a constant "weight" plus the sum of the products of the value of each risk factor (i.e., demographic and CDPS diagnosis categories) in the model (0,1) and its "weight." The weights associated with each risk factor reflect the unique or net effect of each factor on $Pred(SqrRt(\$))_i$. The result is then squared to produce $\mathbf{Pred(SqrRt(\$))_i^2}$. It is important to notice, however, that while the risk factors are common to both populations, the weights for each risk factor are

[1] Additional documentation regarding the development and testing of the "square root model" is available in "Capitation Risk Adjustment for High Risk Populations: A Non-Linear Model" (Gifford, G., Knutson, D., Manning, W., and Finch, M., 2004), a paper submitted to Health Care Financing Review, 2004.

unique to MA Only and Dual Eligible recipients.

The function by which **Pred(SqrRt(PE$))$_i^2$** is calculated is an additive function of the products of Pred(SqrRt($))$_i^2$ and Pred(SqrRt($))$_i$ and the "weight" for each of these risk factors. These weights reflect the unique affects of Pred(SqrRt($))$_i^2$ and Pred(SqrRt($))$_i$ on Pred(SqrRt(PE$))$_i^2$.

Since Pred(SqrRt($))$_i$ is (1) the basis of the model, and (2) a function of selected demographic and CDPS diagnosis category assignments for each recipient month,[2] data records that store data reflecting these assignments for each month a given recipient was enrolled in a program or health plan during a given assessment period comprise all the data necessary to implement these models. This statement is applicable to both the state, when calculating risk scores for recipients of a given program population (e.g., MA Only), and health plans, when calculating risk scores for recipients enrolled in a given health plan. The following instructions regarding the development of these data and the execution of these models applies to both SNBC populations - MA Only vs. Dual Eligible - except where noted in the text.

Users should **identify and separate relevant SNBC MA Only and Dual Eligible population data** prior to performing the steps described below. A recipient month data record should be created for each month in which the recipient was (1) enrolled in MA but not in Medicare (i.e., either Part A or B) during the month – **MA Only;** or (2) enrolled in both MA and in both Part A and B Medicare during the month - **Dual Eligible**.

The implementation of each of these models will require users complete **four basic steps**: **Step 1**: Develop a **recipient month record input data** set for each of the MA Only and Dual Eligible SNBC recipient populations separately. **Step 2**: Calculate **predicted CY 2005 reimbursements** per member per month (pmpm) for each recipient by executing the relevant acute care risk prediction model for the MA Only and Dual Eligible SNBC recipient populations separately. Predicted CY 2005 reimbursements reflect expected increases in the use of some mental health services from 2005 to 2008, which were projected through adjustments to CDPS psychiatric high and medium weights. These adjustments are reflected in Table 3. **Step 3**: Calculate **recipient month relative risk scores**, and **Step 4**: Calculate **health plan level relative risk score(s)** for the selected SNBC enrolled populations.

[2] Recipient month refers to a month in which a person was enrolled in, or a member of, a given program or health plan during an assessment period, regardless of whether the recipient received "paid" services during the month.

STEP 1: DEVELOP RECIPIENT MONTH RECORD INPUT DATA

Separate recipient month record "input" data sets for each of the MA Only and Dual Eligible populations separately must be created to implement these models. By a recipient month record data set we mean a data file comprised one data record for each month a recipient was eligible to receive Minnesota public program health services during a one year assessment period that includes selected demographic characteristics, and diagnosis data from their use of services for a designated one year risk assessment period. *Recipient month relative risk scores will be merged into these data sets in Step 4, so they should be retained for that purpose*.

The data elements required for each recipient month are 63 binary (0,1) "indicator" variables that reflect four (4) basic types of data as follows: (1) five age/gender indicators, one for each of 5 possible age/gender categories into which recipient months were classified during the assessment period; (2) one indicator reflecting whether or not recipients resided in either a nursing facility (NF) or ICF-MR during the month; (3) one indicator reflecting whether or not recipients received any "waivered" services during the month; and (4) *a single array of* fifty-six (56) medical *diagnosis indicator variables* assigned by the Chronic Illness and Disability Payment System (CDPS)[3] based on the classification of all available ICD9 diagnosis codes from paid health care claims or encounters *for each recipient* across the relevant one year assessment period. For each recipient, this array will be fixed across all recipient month records for a given recipient and risk assessment period.

Tables 1 and 2 show examples of these data elements and suggested record formats for the data files required to implement the risk adjustment method for five "SNBC Eligible" recipients who were among the population of 45,642 enrolled in Fee-For-Service (FFS) MA Only in calendar year (CY) 2005. Relevant data regarding these individuals was selected and reproduced here to illustrate data file requirements. Although detailed data file specifications are provided below separately for each category of data - demographic, CDPS condition indicators, given the relatively small number of data elements required, users should recognize that it is possible to store all data elements on a single data record for each recipient/recipient. Recipient IDs for these five "SNBC" eligible recipients are not related to actual MA recipient IDs.

[3] Kronick, R., Gilmer, T., Dreyfus, T., and Lee, L., Improving Health-Based Payment for Medicaid Beneficiaries: CDPS, HEALTH CARE FINANCING REVIEW, Spring 2000, Vol. 21, No. 3.

Create Demographic Indicator Variables

Table 1 shows the data record format for the demographic data required to execute the risk score prediction models for our 5 MA Only "SNBC Eligible" recipients who were enrolled in Fee-For-Service (FFS) MA Only in calendar year (CY) 2005. Although Table 1 shows only one data record for each "Recip ID," as indicated above, these data elements will need to be reproduced on each recipient month record. For each SNBC recipient month record, age and gender data from Medical Assistance eligibility records are needed to create their respective indicator variables. Age should be defined as of the end of the one-year risk assessment period. The risk models require a separate "indicator" variable for each age and gender category specified in the models. However, the CDPS software automatically assigns these age/gender indicators to recipient IDs when it performs the diagnosis classification. It only requires that users supply recipient age (integer), and whether or not recipient is male (1 = yes, 0 = no or female; See CDPS Instruction Manual, Version 2.5 August 2005).

The "Waivered Services Recipient" indicator should indicate whether or not recipients were recipients of any "waivered" services for each recipient month during the assessment period. Waivered services are operationally defined by specific values of the "Category" field that are assigned to recipients on an ongoing basis. "Category" is a DHS internal classification of recipient status based on both living arrangement and type of "waivered" program.

Table 1

Demographic Data Format for SNBC Acute Care Capitation Risk Adjustment Model

Recip ID	Male 15-24	Female 25-44	Male 25-44	Female 45-64	Male 45-64	NF or ICF-MR	Waiv. SvsRcp
1	0	1	0	0	0	0	0
2	1	0	0	0	0	0	0
3	1	0	0	0	0	0	0
4	0	1	0	0	0	0	0
5	0	0	0	0	1	0	0

The specific values of the field "Category" that reflect use of "waivered" services are the following ten 13 character string values:

- 'CAC - Conv   '
- 'CAC - Div   '
- 'TBI NB - Conv'
- 'TBI NB - Div '
- 'CADI - Conv '
- 'CADI - Div   '
- 'MR/RC - Conv '
- 'MR/RC - Div '
- 'TBI NF - Conv' and
- 'TBI NF - Div '.

The "NF or ICF-MR" indicator should reflect whether or not recipients resided in either a nursing facility (NF) or ICF-MR for each month of the risk assessment period. The specific values of the variable "Category" that should be used to develop the NF or ICF-MR indicators are 'NF ' and 'ICF '.

Note Regarding Demographic "Reference Groups": Notice that indicator variables for (1) females age 14-25 years, months in which recipients who (2) did not receive "waivered" services, or (3) did not reside in a NF or ICF-MR during a given month of the assessment period do not need to be included in the data file. However, recipient month data records regarding these subpopulations nevertheless should be included in the data files because they comprise the "reference groups." A "reference group" is the sub-population with the lowest mean unadjusted reimbursement amount (i.e., before controlling for other predictors) relative to the other groups defined by the same dimension (e.g., gender and age), to which the net effects –weights - of the relevant indicator variable actually specified in a given model "refer" (i.e., are compared).

The demographic group females age 14-25 years was defined as the age/gender reference group for both the MA Only and Dual Eligible models, because, for the MA Only CY 2005 population, it was the age/gender category with the lowest mean unadjusted CY 2005 reimbursements pmpm. Since indicators of recipients of "waivered" services, and recipients that resided for one or months in a NF or ICF-MR were included in both models, recipients who were not recipients of waivered services, or who did not reside in a NF or ICF-MR at any time during an assessment period are, in effect, defined as the reference groups for these indicators.

Create CDPS Diagnosis Indicator Variables

A single array of values for the CDPS medical diagnosis indicator variables must be assigned to each of the recipient month records for a given recipient and assessment period. This is accomplished by executing the CDPS Version 2.5 software code within a Statistical Analysis Software (SAS) application using the following input data file specifications. The CDPS program is a SAS program that classifies the ICD-9-CM diagnosis codes assigned to recipients by health providers from a given assessment period into from 0 to 1 or more of 56 distinct CDPS condition categories.

Because CDPS assignment is "hierarchical" within major disease category (e.g., cardiovascular), the maximum number of CDPS condition categories assigned to a given recipient is less than 56. Each CDPS disease category is grouped into one of 19 major diagnosis categories that reflect major body systems or disease types (e.g., cardiovascular, metabolic). Most disease categories within major category are "hierarchically" ordered by CDPS according to average cost, such that only the highest "cost" disease category within major disease category is counted. For example, if a recipient's ICD-9 codes could be classified into both cardiovascular low and medium of the 56 CDPS disease categories, only cardiovascular medium was counted. As an example of what might be expected, when CDPS 2.5 diagnosis categories were assigned to the CY 2005 FFS MA Only population, the number of major disease categories assigned ranged from "0" to "15."

The CDPS software and CDPS Instruction Manual can be obtained at http://cdps.ucsd.edu. The CDPS software requires the user develop two input SAS data files – the CDPS eligibility input data file and the CDPS diagnoses input data file. (See CDPS Instruction Manual for more information.)

CDPS Eligibility Input Data File. The CDPS Eligibility Input Data File should contain one data record per recipient, and each record should include five data elements: (1) a recipient identification number, (2) category of assistance, (3) recipient age, (4) a recipient gender indicator, and (5) months enrolled. The CDPS software requires that the five data elements in the Eligibility Data File be created according to the following specifications.

Element	Required Variable Name	Type(Specifications)
Recipient Identification Number	RECIPNO	Character (any length)
Category of Assistance	AIDCAT	Numeric (1 character in length, 1 = AFDC/TANF Adult 2 = AFDC/TANF Child 3 = Disabled Adult 4 = Disabled Child.
Recipient Age	AGE	Numeric (not grouped range)
Recipient Gender	MALE	Numeric (1 character in length, 1 = male, 0 = female
Months Enrolled	ELIG	Number of months enrolled

For the SNBC project, since all recipients are adults with disabilities, the AIDCAT element in the Eligibility Data File should be set to "3" for all recipients.

CDPS Diagnoses Input Data File. The second data file required by the CDPS software is a file that contains the ICD-9 diagnosis codes assigned to recipients during the assessment period. According to CDPS requirements, this data file can have zero (0) to multiple records per recipient, and should include only diagnoses from paid utilization of "acute care" services. For each recipient, the diagnosis data file should include records that include a recipient identification number that must be named RECIPNO, and should be specified exactly as in the Eligibility Data File. The diagnoses must be named DIAG1, DIAG2, DIAG3, etc., and each diagnosis code should be specified as a five character, left justified string with no decimals, and with leading zeros for diagnoses in the range of 001 - 099. If the 4th and 5th characters are blank, for instance, blanks should be written in the 4th and 5th positions of the diagnosis string.

In addition, while CDPS software developers indicate that the software permits up to 14 diagnoses per record,[4] they recommend users set the number of diagnosis codes per record to 10. Then, to accommodate all diagnosis codes available for each recipient, users can create multiple diagnosis records per recipient. As an example, when the CY 2005 FFS data used to develop the models was prepared, since the maximum number of valid ICD-9 codes per recipient month record was 80, we created 8 diagnosis records per recipient month, and 10 diagnosis codes per diagnosis record. Since their were 427,503 recipient month records for the CY 2005 MA Only population, and 8 diagnosis records per recipient month, we created a total of 3,420,024 ICD9

[4] Source: CDPS/UCSD support staff , although CDPS documentation suggests 10 diagnoses per record.

diagnosis records for input into the CDPS software for this population.

Using these Eligibility and Diagnoses SAS data sets as input data files, CDPS outputs a SAS data set comprised of as many person or recipient level records as there are recipient IDs in the input Diagnoses Data File. The SAS data set produced is named xxxxxxxg.sas7bdat, where the "xxxxxxx" is a user specified file name prefix. When the total number of recipient IDs in the CDPS output SAS data set is less than the number of Recipient IDs in the input eligibility set, it is because there were no diagnosis records for some of the recipient IDs in the eligibility file.[5] Each record in the CDPS output SAS data set includes the Recipient ID number, 56 CDPS chronic condition or disease category indicator (0,1) variables, the age/gender indicators, a number of other variables, and a risk score for each recipient. The risk scores assigned by the CDPS software to recipient IDs should be ignored. We use the CDPS software only to assign the available ICD-9 codes to one or more of the 56 possible CDPS diagnosis indicators, and to assign values (0,1) to the age/gender indicators for each recipient.

Since the 56 CDPS-generated indicators and the age/gender variables are themselves (0,1) indicators, they can be used in the models without modification. Table 2 shows the values for the 56 CDPS-based indicator variables used by the SNBC models for our same 5 example recipients whose demographic data appear in Table 1. These indicators must be transferred to each of the original recipient month records available for each recipient ID to calculate predicted CY 2005 reimbursements$_{pmpm}$.

STEP 2: CALCULATE PREDICTED CY 2005 REIMBURSEMENTS PMPM

The next step is to use the demographic and CDPS disease indicator data developed in Step 1 and the risk prediction models to calculate predicted CY 2005 acute care reimbursements per member per month (pmpm). As indicated above, while the MA Only and Dual Eligible models are identical with respect to both structure and selected risk factors, they are distinguished by the values of the weights used in the models. In addition, the weights for the CDPS psychiatric high and medium categories were set to values that achieved projected increases in mental health utilization, separately for MA Only and Dual Eligible recipients – the "Mental Health Enhancement" (MHE). When the instructions refer to predicting CY 2005

[5] e.g., "new enrollees" in the eligibility file whose diagnosis data was therefore not included in the assessment file.

Table 2

CDPS-Based Indicator Variable Data Format for Calculating
Predicted Acute Care Reimbursements
Special Needs Basic Care

RecipID	CARVH	CARM	CARL	CAREL	PSYH	PSYM	PSYL	SKCM	SKCL	SKCVL
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	1	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	1	0	0	1	0	1	0
RecipID	SKCEL	CNSH	CNSM	CNSL	PULVH	PULH	PULM	PULL	GIH	GIM
1	0	0	0	0	0	0	0	1	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0
5	0	0	0	1	0	0	0	0	0	0
RecipID	GIL	DIA1H	DIA1M	DIA2M	DIA2L	SKNH	SKNL	SKNVL	RENVH	RENM
1	0	0	0	0	0	0	0	1	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
RecipID	RENL	SUBL	SUBVL	CANH	CANM	CANL	DDM	DDL	GENEL	METH
1	1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
RecipID	METM	METVL	PRGCM	PRGINC	EYEL	EYEVL	CERL	HIVM	AIDSH	INFH
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
RecipID	INFM	INFL	HEMEH	HEMVH	HEMM	HEML				
1	0	0	0	0	0	0				
2	0	0	0	0	0	0				
3	0	0	0	0	0	0				
4	0	0	0	0	0	0				
5	0	0	0	0	0	0				

Reimbursements$_{pmpm}$ we mean CY 2005 Reimbursements$_{pmpm}$ for MA Only population or Dual Eligible recipients, after MHE.

As suggested above, this step involves developing two prediction equations that derive from Equation (1) that calculate predicted CY 2005 Reimbursements$_{pmpm}$ – hereafter referred to as Pred($)$_i$ in the equations - for the MA Only and Dual Eligible recipients separately based on the weights from Table 3, and recipient values for the demographic and CDPS indicator variables. Table 3 shows the risk factor weights for each of the MA Only and Dual Eligible populations after MHE. Therefore, please note that Table 3 weights should be used when calculating predicted CY 2005 Reimbursements$_{pmpm}$. These weights are based on multiple regression analyses of data regarding the relevant CY 2005 FFS "SNBC Eligible" recipient populations (i.e., 45,624 MA Only recipients; and 48,937 Dual Eligible recipients).

The following instructions are applicable to the models for both populations. As result, the SNBC sub-populations are not indicated in the specifications. STEP 2 has two components, identified by the terms to the right of the = sign in Equation (1) below. Each of these two components comprise the two distinct sub-steps of STEP 2 – labeled 2.a and 2.b.

$$\mathbf{Pred(\$)_i = Pred(SqrRt(\$))_i^2 + Pred(SqrRt(PE\$))_i^2} \quad (1)$$

where: Pred($)$_i$ = predicted CY 2005 Reimbursements in $s pmpm for recipient i;
Pred(SqrRt($))$_i^2$ = the square of the predicted square root of CY 2005 reimbursements pmpm for recipient in $s, before adjusting for prediction error; and
Pred(SqRt(PE$))$_i^2$ = predicted, prediction error also in $s pmpm for recipient$_i$.

Before proceeding to Step 2.a, users must transfer the demographic and CDPS indicator data onto the recipient month record data files, as defined in Step 1.

Step 2.a: Calculate Unadjusted Predicted CY 2005 Reimbursements$_{pmpm}$ - Pred(SqrRt($))$_i^2$

As indicated by equation (1), to calculate predicted CY 2005 reimbursements$_{pmpm}$ for each recipient month - Pred($)$_i$ - we first calculate the predicted square root of CY 2005 reimbursements for each recipient month - Pred(SqrRt($))$_i$. For each recipient month record, Pred(SqrRt($))$_i$ is calculated using equation (2a) below. Equation (2a) specifies that for each recipient month, Pred(SqrRt($))$_i$ = the sum of the constant weight "a", and the sum "∑" of the

product of the risk factor weights "wt_j" for each risk factor "j" and the value of each risk factor "j" (i.e., 0,1 indicator) for each recipient month record "i" across all 63 risk factors "j" - $(\text{Dem, CDPS Conds.})_{ij}$. As a result, weights for risk factors that are false or "0" for a given recipient month record drop out of the equation.

$$\text{Pred(SqrRt(\$))}_i = a + \sum wt_j(\text{Dem, CDPS Conds.})_{ij} \qquad (2.a)$$

As a result, because equation (2a) predicts the square root of predicted CY 2005 $\text{reimbursements}_{pmpm}$ for SNBC recipients, the risk factor weights reflect the unique effects of the presence of each risk factor on the square root of CY 2005 reimbursements pmpm scale. That is, they quantify the increase (or decrease) in the square root of CY 2005 reimbursements pmpm uniquely due to the presence of each risk factor. (Please note that you cannot square the weights to produce "cost" effects in terms of dollars, and then sum these, because of the mathematical nature of the prediction model.)

The result - $\text{Pred(SqrRt(\$))}_i$ - is then converted back into $s by simply squaring it to get $\text{Pred(SqrRt(\$))}_i^2$. This is the initial or unadjusted predicted CY 2005 $\text{reimbursements}_{pmpm}$, based on the ability of the demographic and CDPS condition indicators to predict reimbursements before adjusting for prediction error.

The last step, as indicated by equation (1), is to then add "predicted, prediction error" to $\text{Pred(SqrRt(\$))}_i^2$. Calculating "predicted, prediction error" comprises Step 2.b.

Step 2.b: Adjust $\text{Pred(SqrRt(\$))}_i^2$ by Predicted, Prediction Error - $\text{Pred(SqrRt(PE\$))}_i^2$

The "Predicted, Prediction Error" component - $\text{Pred(SqrRt(PE\$))}_i^2$ - is calculated as indicated by equation (2.b) below. The term $\text{Pred(SqrRt(PE\$))}_i^2$ is the predicted difference between recipients' actual and predicted reimbursements in $s. Equation (2.b) reflects that prediction error, which was originally on the square root scale, was also squared to convert it back into $s before the equation was developed. As a result, when it is greater than zero (0), it indicates the degree to which the difference **actual** reimbursements pmpm minus **predicted** reimbursements pmpm is predicted by $\text{Pred(SqrRt(\$))}_i$, and $\text{Pred(SqrRt(\$))}_i^2$.

To calculate $Pred(SqrRt(PE\$))_i^2$ for recipient month "i", we simply sum the 3 right hand side terms in the equation: (a) the constant weight "b", (b) the product of the weight for the predicted square root of reimbursements "wt_k" and the value of the predicted square root of reimbursements for recipient month i - $Pred(SqrRt(\$))_i$, and (c) the product of the weight for the initial predicted reimbursements pmpm in $ "wt_l" for recipient month i - and the value for the initial or unadjusted predicted CY 2005 $reimbursements_{pmpm}$ for recipient month i - $Pred(SqrRt(\$))_i^2$. Remember, $Pred(SqrRt(\$))_i$, and $Pred(SqrRt(\$))_i^2$ were calculated in Step 2.a..

$$Pred(SqrRt(PE\$))_i^2 = b + wt_k\,(Pred(SqrRt(\$))_i) + wt_l\,(Pred(SqrRt(\$))_i^2) \qquad (2.b)$$

The constant and weights for the prediction error equations for both MA Only and Dual Eligible populations are also supplied in Table 3. The only values users need to supply for each recipient month are the predicted square root of CY 2005 $reimbursements_{pmpm}$ - $(Pred(SqrRt(\$))_i$ - from step 2.a, and the square of that value - $(Pred(SqrRt(\$))_i^2$.

The resulting value is the "predicted, prediction error" - $\mathbf{Pred(SqrRt(PE\$))_i^2}$ - for each recipient. It represents, in effect, the amount by which recipients' total actual CY 2005 reimbursements pmpm exceed their predicted reimbursements pmpm from Step 2.a as predicted by $Pred(SqrRt(\$))_i$ and $Pred(SqrRt(\$))_i^2$ as specified by Equation 2.b. **When added to** the initial predicted reimbursements from Step 2.a - $\mathbf{Pred(SqrRt(\$))_i^2}$, the **result is** final, prediction error adjusted, predicted CY 2005 $reimbursements_{pmpm}$ - $\mathbf{Pred(\$)_i}$ – according to Equation (1).

STEP 3: CALCULATE RECIPIENT MONTH RELATIVE RISK SCORES

The next step is to convert $Pred(\$)_i$ into a recipient month relative risk score that is standardized to the average CY 2005 $reimbursements_{pmpm}$ of a given population. These risk scores can then be used to risk adjust the CY 2008 base capitation rate developed for each population. Relative risk scores for each recipient month of enrollment are calculated by dividing $\mathbf{Pred(\$)_i}$ by the **mean predicted CY 2005 mean MHE adjusted** $\mathbf{reimbursements_{pmpm}}$, hereafter referred to as - **Mean($)** - for the relevant FFS CY 2005 SNBC population, as indicated by Equation (3). **Mean($)** for the **MA Only** population = **$1,153.46**; and for the **Dual Eligible** population = **$305.70**. Since predictor variables/risk factors are constant within recipients,

relative risk scores will be constant across all recipient month records for a given recipient.

$$\text{Risk Score}_i = \text{Pred}(\$)_i / \text{Mean}(\$) \quad (3)$$

To illustrate this step, below we show and describe the calculations in detail for STEPs 2 & 3, for Recipient ID #1 from the 5 example recipients whose input data appears in Tables 1 - 2.

Recipient ID #1 was a female 25-44 years of age in CY 2005 (See Table 1), with three assigned CDPS diagnosis categories, and no other risk factors. These characteristics are shown in the following table as well as the relevant weights from Table 3.

Recipient #1 Predictors	Risk Factor Weights
Constant	8.120
Demographics	
Age/Gender: A_25_44_F	0.656
CDPS Categories	
PULL05	4.066
SKNVL05	2.316
RENL05	3.803

Step 2: Calculate Predicted CY 2005 $\text{Reimbursements}_{pmpm}$

Step 2a: Calculate Unadjusted Predicted CY 2005 $\text{Reimbursement}_{pmpm}$ - $\text{Pred}(\text{SqrRt}(\$))_i^2$

Since there are 4 predictors from equation (2) that are true or = 1 for this recipient, we sum the constant weight and the weights for the 4 risk factors that are set to 1 to get $\text{Pred}(\text{SqrRt}(\$)_i$, and then square it to get $\text{Pred}(\text{SqrRt}(\$))_i^2$. For all Recipient ID #1's monthly records, $\text{Pred}(\text{SqrRt}(\$))_i$ =18.961, and so $\text{Pred}(\text{SqrRt}(\$))_i^2$ = $359.52.

Step 2b: Adjust $\text{Pred}(\text{SqrRt}(\$))_i^2$ by Predicted, Prediction Error

For a given recipient, $\text{Pred}(\text{SqrRt}(\text{PE}\$))_i^2$ is the sum of: (a) the constant weight "b", a

constant for all recipients; (b) the product of the weight - "wt_k" - for the predicted square root of reimbursements - $Pred(SqrRt(\$))_i$ - and recipient i's predicted value for $Pred(SqrRt(\$))_i$, and (c) the product of the weight - "wt_l" - for the predicted value for $Pred(SqrRt(\$))_i^2$ and recipient i's predicted value for $Pred(SqrRt(\$))_i^2$. (We reproduced Equation (2.b) below to help clarify the substitutions.)

$$Pred(SqrRt(PE\$))_i^2 = b + wt_k\,(Pred(SqrRt(\$))_i + wt_l\,(Pred(SqrRt(\$))_i^2 \qquad (2.b)$$

For Recipient ID #1, the calculation would take place as follows:

$$Pred(SqrRt(PE\$))_1^2 = 143.441 - 7.188\,(18.961) + 0.252\,(359.52) = 97.748.$$

We now have calculated values for all the elements of Equation (1) to complete the calculation of $Pred(\$)_i$ for this recipient. Equation (1) is reproduced below followed by its application to calculating $Pred(\$)_i$ for this recipient.

$$Pred(\$)_i = Pred(SqrRt(\$))_i^2 + Pred(SqrRt(PE\$))_i^2 \qquad (1)$$

By substituting this recipient's values for the right hand terms in Equation (1) we can now calculate $Pred(\$)_1$ as shown below.

$$Pred(\$)_1 = \$359.52 + 97.748 = \$457.27.$$

From Equation (3), then, the relative risk score for Recipient ID #1 is calculated as follows:

$$\text{Risk Score}_1 = \$457.27 / \$1{,}153.46 = 0.3964.$$

STEP 4: CALCULATE HEALTH PLAN RELATIVE RISK SCORES

The fourth and final implementation step is to *calculate health plan risk scores - by which acute care base rates* can be adjusted for participating health plans (hp). There are two methods by which health plan risk scores – $HPRS_{hp}$ - can be calculated. The first method is to average the risk scores across all recipient month records for a given health plan, as shown in Equation (4a) below.

$$HPRS_{hp} = \frac{\text{Sum of Risk Scores}_{i\,hp}\ (\text{across all recipient month records}_{\ i\ hp})}{\text{Total recipient months records}_{\ hp}} \qquad (4a)$$

The second method is calculated by taking the ratio of the average of the "Predicted CY 2005 MHE adjusted Reimbursements$_{pmpm\ hp}$" (i.e., across all recipient month records) for a given health plan (hp) over the **Mean($)**.

$$HPRS_{hp} = \frac{\text{Mean of Pred(\$) pmpm}_{\ hp}}{\text{Mean(\$)}} \qquad (4b)$$

We used the data regarding our five example cases as a very small, simulated health plan to illustrate the calculation of health plan risk scores. We calculated the predicted CY 2005 MHE adjusted reimbursements pmpm and the recipient level risk scores for the remaining four example recipients. Those calculations and the results are shown in Table 4. Table 4 shows total CY 2005 actual reimbursements, the number of CY 2005 recipient month records for each Recipient ID, actual CY 2005 reimbursements pmpm, predicted CY 2005 MHE adjusted reimbursements, recipient relative risk scores pmpm, the sum of recipient risk scores across all recipient month records, and the "Health Plan Risk Score" calculated using both methods for this hypothetical health plan with five recipients/recipients.

Using equation 4a, the risk score for each recipient month is the ratio of each individual recipient's predicted CY 2005 reimbursements pmpm to the mean MHE CY 2005 reimbursements pmpm for the CY 2005 "SNBC Eligible" MA Only population - $1,153.46. Since the sum of the risk scores for this population across all recipient months is 22.1395, the mean risk weight for the CY 2005 "SNBC Eligible" MA Only population across all 52 months of eligibility is 22.1395/52 = 0.4258. As a result, the mean risk weight for any population to which this model is applied, across all months enrolled, will reflect the acute care risk of the relevant population relative to, or standardized, to that of the CY 2005 FFS "SNBC Eligible" population (MA Only vs. Dual Eligible), and comprises the health plan risk score.

The Health Plan Risk Score in Table 4 - 0.4258 - is therefore a simulated $HPRS_{hp}$ for our five-person sample of recipients. Conceptually, using equation 4a, it is the weighted[6] average of the five recipients' risk scores. This is mathematically equivalent to assigning each recipient's risk score to each recipient month record, and then calculating a simple average across all recipient month records. Using equation 4b, it reflects the relative degree to which the average predicted CY 2005 MHE adjusted reimbursements pmpm for the health plan recipients is above or below the average predicted CY 2005 MHE reimbursements pmpm for the relevant SNBC population - for MA Only: $1,153.46; Dual Eligible: $305.70. When individuals enroll in SNBC, these health plan risk scores can be used to adjust SNBC base capitation rates for the MA Only and Dual Eligible populations.

Since health plan risk scores - $HPRS_{hp}$s - reflect the acute care risk of the enrolled populations relative to the relevant CY 2005 FFS "SNBC Eligible" populations (i.e., MA Only vs. Dual Eligible), $HPRS_{hp}$s can be standardized (normalized) to an actual prepaid SNBC population by taking the ratio of $HPRS_{hp}$s to the overall SNBC-wide HPRS across all participating health plans for a given SNBC sub-population - MA Only or Dual Eligible.

[6] "weighted" means each recipient record is weighted by the total recipient months for each recipient.