

# **Technical Notes on the Standardized Mortality Ratio**

*For the Dialysis Facility Reports*

September 2010

## Technical Notes on the Standardized Mortality Ratio For the Dialysis Facility Reports

### Table of Contents

Table of Contents .....	2
Introduction .....	3
Assignment of Patients to Facilities for the SMR Calculation .....	4
General Inclusion Criteria for Dialysis Patients .....	4
Identifying Facility Treatment Histories for Each Patient .....	4
Days at Risk for Each Patient .....	5
Model for Calculating Expected Mortality .....	5
Missing Data .....	6
Calculation of Expected Deaths at a Facility .....	7
Example .....	8
Caveats .....	10
References .....	11

## Introduction

The Standardized Mortality Ratio (SMR) in Table 1 of the Dialysis Facility Reports (DFR) is designed to summarize the mortality at a facility relative to the mortality that would be expected, based on the characteristics of the patients at that facility. The SMR equals the ratio of the actual number of deaths divided by the expected number of deaths. The SMR estimates the relative death rate ratio for a facility, as compared to the national death rate. Qualitatively, the degree to which the facility's SMR varies from 1.00 is the degree to which it exceeds ( $>1.00$ ) or is under ( $<1.00$ ) the national death rates for patients with the same characteristics as those in the facility.

An important change to the report this year is that the SMR for a particular calendar year is now compared to the US mortality rates for that same year rather than to the entire 4-year period. The advantage to this is that the reference year for a particular estimate will be the same in each DFR and therefore the SMR value will change less between DFRs. In the past, because these statistics were compared to a different reference population in each DFR, the values changed more over time, even for the same year across reports. The use of a different reference year for each year's estimate will allow you to identify trends over time at your facility beyond the overall US trend over time. In other words, if the SMR for your facility decreases over the time period, this means that mortality at your facility has decreased more over that time period than the overall US average mortality decreased. If mortality at your facility decreased over the four year period at the same rate that overall US mortality decreased over this time period, the SMR for your facility would be the same for each year.

The SMR is adjusted for age, race, ethnicity, sex, diabetes, duration of ESRD, nursing home status, comorbidities at incidence, body mass index (BMI) at incidence, calendar year, and race-specific state population death rates. The SMR indicates whether patients treated in the facility had higher or lower mortality than expected when adjusted for age, race, ethnicity, sex, diabetes, years of ESRD, comorbidities, BMI, year, and population death rates. Quantitatively, if the facility's death rates equal the national death rates (in deaths per patient year or per year at risk) times a multiplicative constant, then the SMR estimates that multiplicative constant. If the multiplicative constant varies for different subgroups of patients, then the SMR estimates a weighted average of those constants according to the facility's patient mix. For example, an  $SMR=1.10$  would indicate that the facility's death rates typically exceed national death rates by 10% (e.g., 22 deaths observed where 20 were expected, according to the facility's patient mix). Similarly, an  $SMR=0.95$  would indicate that the facility's death rates are typically 5% below the national death rates (e.g., 19 observed versus 20 expected deaths). An  $SMR=1.00$  would indicate that the facility's death rates equal the national death rates, on average.

Similarly, the regional SMR values in the DFR are calculated as the ratio of the total number of observed deaths among patients from each region to the number of expected deaths among patients from each region.

## **Assignment of Patients to Facilities for the SMR Calculation**

This section describes the methods we used to assign patients to a facility in order to calculate the SMR. Because some patients receive dialysis treatment at more than one facility in a given year, we use standard methods based on assigning person-years to a facility, rather than on assigning a patient's entire follow-up to a facility. We developed conventions which define the group of patients assigned to a facility at any time during the particular year. This method is described below.

### ***General Inclusion Criteria for Dialysis Patients***

A patient's follow-up in the database can be incomplete during the first 90 days of ESRD therapy. For the purposes of this report, we only entered a patient's follow-up into the tabulations after that patient had received chronic renal replacement therapy for at least 90 days. Mortality and survival during the first 90 days do not enter into the calculations. This minimum 90-day period assures that most patients are eligible for Medicare insurance — either as their primary or secondary insurer. It also excludes from analysis patients who died during the first 90 days of ESRD, since such patients may have incomplete data.

In order to exclude patients who only received temporary dialysis therapy, we assigned patients to a facility only after they had been on dialysis there for at least 60 days. This 60 day period is used both for patients who started ESRD for the first time and for those who returned to dialysis after a transplant. That is, deaths and survival during the first 60 days of treatment at a facility do not affect the SMR of that facility.

### ***Identifying Facility Treatment Histories for Each Patient***

For each patient, we identified the dialysis provider at each point in time using data from a combination of Medicare-paid dialysis claims, the Medical Evidence Form (Form CMS-2728), and paid dialysis claims. Starting with day 91 of ESRD, we determined facility treatment histories for each patient, and then listed each patient with a facility only once the patient had been treated there for 60 days. When a patient transferred from a facility, the patient remained assigned to it in the database for 60 days. This continued tabulation of the time at risk for 60 days after transfer from a facility attributes to a facility the sequelae of treatment there for 60 days, even when a patient was transferred to another facility (such as a hospital-based facility) after the patient's condition worsened.

In particular, we placed patients in their initial facility on day 91 of ESRD once that facility had treated them for at least 60 days. If on day 91 a facility had treated a patient for fewer than 60 days, we waited until the patient reached day 60 of treatment at that facility before placing him or her there. State and Network summaries do not include patients who were not assigned to a facility; these patients are, however, included in the U.S. summaries.

Using SIMS data and paid dialysis claims to determine whether a patient has transferred to another facility, we attributed patient outcomes to the patient's original facility for 60

days after transfer out. On day 61 after transfer from a facility, we placed the patient in the new facility once s/he had been treated at the new facility for 60 days. When a patient was not treated in a single facility for a span of 60 days (for instance, if there were two switches within 60 days of each other), we did not attribute that patient to any facility.

Patients were removed from a facility's analysis upon receiving a transplant. Patients who withdrew from dialysis or recovered renal function remained assigned to their treatment facility for 60 days after withdrawal or recovery. Additionally, patients for whom the only evidence of dialysis treatment is the existence of Medicare claims were considered lost to follow-up and removed from a facility's analysis one year following the last claim, if there was no earlier evidence of transfer, recovery, or death. In other words, if a period of one year passed with neither paid Medicare dialysis claims nor SIMS information to indicate that a patient was receiving dialysis treatment, we considered the patient lost to follow-up, and did not continue to include that patient in the analysis. If evidence of dialysis re-appeared, the patient was entered into analysis after 60 days of continuous therapy at a single facility. Finally, all SIMS records noting continuing dialysis were extended until the appearance of any evidence of recovery, transfer, or death. Periods of lost to follow-up were not created in these cases since the instructions for SIMS only require checking patient data for continued accuracy, but do not have a requirement for updating if there are not any changes.

### ***Days at Risk for Each Patient***

After patient treatment histories are defined as described above, periods of follow-up time are created for each patient. A new time period begins each time the patient is determined to be at a different facility and at the start of each calendar year. The number of days at risk starts over at zero for each time period so that the number of days at risk for any patient-year-facility period is always a number between 0 and 365 (or 366 for leap years). Therefore, a patient who is in one facility for all four years is analyzed the same way as four separate patients in the facility for one year each. When patients are treated at the same facility for two or more separate time periods during a year, the days at risk at the facility is the sum of all time spent at the facility for the year. For example, consider a who patient spends two periods of 100 days assigned to a facility, but is assigned to a different facility for the 165 days between these two 100-day periods. This patient will have one period of 200 days at risk at the first facility, and a separate period of 165 days at risk at the second facility.

The number of days at risk ( $t_i$ ) in each of these patient-year-facility time periods is used to calculate the expected number of deaths for the patient during that period as described in the "Calculation of Expected Deaths at a Facility" section below. The SMR for a facility is the ratio of the total number of observed to the total number of expected deaths during all time periods at the facility.

### **Model for Calculating Expected Mortality**

The SMR uses expected mortality calculated from a Cox model (SAS Institute Inc., 2004; Andersen, 1993; Collett, 1994). The model is fit in two stages. The stage 1 model is a

Cox model stratified by facility and adjusted for patient age, race, ethnicity, sex, diabetes, duration of ESRD, nursing home status, patient comorbidities at incidence, calendar year, body mass index (BMI) at incidence. This model allows the baseline survival probabilities to vary between strata (facilities), and assumes that the regression coefficients are the same across all strata. The linear predictor for each patient based on the regression coefficients in the stage 1 model is then used as an offset in the stage 2 model, which also includes adjustment for race-specific state population death rate.

The patient characteristics included in the stage 1 model as covariates are age, race, ethnicity, sex, cause of ESRD (diabetes or other), duration of ESRD (<1 year, 1-2 years, 2-3 years, 3+ years as of the period start date), nursing home status, comorbidity index at incidence, calendar year, BMI at incidence, and interaction terms between race, sex and duration and cause of ESRD. Age as of the period start date is included as a piecewise continuous variable with different coefficients based on whether the patient is 0-13 years old, 14-60 years old, or 61+ years old, and whether the patient is black or not. Ethnicity is included with different coefficients for white and non-white patients. The comorbidity index is included as a linear variable. BMI and race-specific state population death rates are included as log-linear terms. Categorical indicator variables are included as covariates in the stage 1 model to flag records missing values for cause of ESRD, comorbidity index, calendar year, and BMI. These variables have a value of 1 if the patient is missing the corresponding piece of information and a value of 0 otherwise. The stage 2 model includes the population death rate for patients of that race in that state as a covariate. The example below shows how these coefficients are used to carry out the calculations.

The Microsoft Excel file available with this document indicates the value of the coefficient for each characteristic in the stage 1 and 2 models (beta) as well as the corresponding standard error and a p-value indicating if the coefficient is significantly different than 0. The file also includes the baseline survival curve for the stage 2 model. The comorbidity index is calculated as a weighted linear combination of comorbidities reported on the Medical Evidence Form (CMS-2728). These weights are also provided in the Excel file.

Age adjusted population death rates (per 100,000) by state and race are obtained from the U.S. Centers for Disease Control National Center for Health Statistics. The 2010 DFR used age-adjusted death rates for 2003-05 from Table 27 of the publication *Health, United States, 2008, With Chartbook on Trends in the Health of Americans* available at [http://www.cdc.gov/nchs/data/08.pdf](http://www.cdc.gov/nchs/data/hus/08.pdf).

### ***Missing Data***

Patients with missing data are not excluded from the model. Patients with missing diagnosis are included in the “other” diagnosis group strata. For the purposes of calculation, missing values for the comorbidity index and BMI are replaced with mean values for patients with similar age, race, sex, and cause of ESRD. When the cause of ESRD is missing, missing values are replaced with mean values for patients with similar age and sex. These mean values are included in the attached tables. Patients with missing race are included in the “other” race group strata and classified as non-White in the

model. Patients with missing ethnicity are classified as “unknown” ethnicity. No patients were missing age, sex, or date of first ESRD treatment.

As mentioned above, indicator variables identifying patients with missing values for cause of ESRD, comorbidity index, and BMI are also included as covariates in the model.

## Calculation of Expected Deaths at a Facility

The Cox model consists of two stages. Stage 1 yields estimates of the coefficients  $\beta_j$  for the 42 covariates in the model. Using these coefficients, a predicted value is calculated for each patient. Stage 2 of the model uses only one covariate, the log of the population death rate for the patient’s race within the state, and utilizes the patient’s predicted value from stage 1 as an offset. The predicted value from stage 1 and the baseline survival curve from stage 2 of the Cox model are then used to calculate the expected number of deaths for a specific patient.

Let  $p$  denote the number of patient characteristics in the model and  $x_{ij}$  be the specific value of the  $j^{\text{th}}$  characteristic for the  $i^{\text{th}}$  patient. In stage 1, for patient  $i$ , with characteristics

$$\mathbf{X}_i = x_{i1}, x_{i2}, \dots, x_{ip},$$

we calculate:

$$\text{stage 1 } \mathbf{X}_i' \mathbf{B} = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}$$

where  $\beta_j$  is the  $j^{\text{th}}$  coefficient from the model. For a categorical characteristic, the  $x_{ij}$  value is 1 if the patient falls into the category and 0 otherwise.

In stage 2, let  $x_j$  be the log of the state population death rate for a specific state and race. We utilize **stage 1**  $\mathbf{X}_i' \mathbf{B}$  as an offset to calculate:

$$\text{Stage 2 } \mathbf{X}_i' \mathbf{B} = \beta_1 x_j + \text{stage 1 } \mathbf{X}_i' \mathbf{B}$$

Suppose that  $t_i$  is the end of follow-up time for patient  $i$  and that  $S_{0k_i}(t_i)$  is the baseline survival probability at time  $t_i$ . The survival probability for this patient at time  $t_i$  is:

$$S_{ik_i}(t_i) = [S_{0k_i}(t_i)]^{\exp(\text{stage 2 } \mathbf{X}_i' \mathbf{B})}$$

The expected number of deaths for this patient during follow-up is then

$$-\ln(S_{ik_i}(t_i)) = -e^{\text{stage 2 } \mathbf{X}_i' \mathbf{B}} \ln[S_{0k_i}(t_i)]$$

and summing these values for the  $N$  patients at the facility

$$\sum_{i=1}^N -\ln[S_{ik_i}(t_i)] = -\sum_{i=1}^N e^{\text{stage 2 } \mathbf{X}_i' \mathbf{B}} \ln[S_{0k_i}(t_i)]$$

results in the expected number of deaths during follow-up at the facility. Thus, patients with 100 days of follow-up, who are otherwise the same, have the same expected mortality even if the 100 day period started at different dates during the year. This approximation is made to simplify the calculations.

As stated above, the SMR is the ratio of this expected value to the total number of deaths observed at the facility during follow-up.

**Example**

As an example, we calculate the one-year SMR for a hypothetical facility in Florida that treated 5 patients in 2008. Table 1 describes the patients and their sequelae of treatment.

**Table 1.** Description of example patients at example facility for 2008 SMR calculation

Patient	Start of ESRD	Dates treated at this facility	Characteristics
1	4/15/2006	4/15/2006 to 12/31/2008	Male, White, 51, diabetic, alcohol dependence at incidence, BMI=32.2
2	3/1/2008	3/1/2008 to 12/31/2008	Female, Black, 70, BMI=23.0
3	11/1/2007	11/1/2007 to 11/10/2008 death	Female, Black, 78, BMI=22.1
4	8/15/2002	7/15/2008 to 7/31/2008 transfer	Female, Asian, 66, diabetic, Hispanic, BMI=18.7
5	11/1/2003	11/1/2003 to 4/1/2008 transfer and 8/1/2008 to 12/31/2008	Male, 83, missing race and cause of ESRD, COPD at incidence , BMI=28.5

First we determine which patients are assigned to the facility and for how many days each assignment lasts. Patient 1 started treatment at our example facility in 2006 and is assigned to the facility for the entire year of 2008. Patient 2 started renal replacement therapy on 3/1/2008 and is assigned to our facility after 90 days. Patient 3, similarly, started RRT on 11/1/2007 is assigned to our facility after 90 days until her death on 11/10/2008. Patient 4 was only in our facility 16 days, so was never assigned to our facility. Patient 5 started treatment at our facility in 2003 but was treated at another facility from 4/1/2008 through 8/1/2008, so has two treatment periods at the facility which are combined. Table 2 summarizes the assignment periods.

**Table 2.** Patient assignment periods

Patient	Dates Assigned	Days ( $t_i$ )	Notes
1	1/1/2008 to 12/31/2008	366	
2	5/30/2008 to 12/31/2008	216	Eligible starting with day 91 of ESRD
3	1/30/2008 to 11/10/2008	286	Eligible starting with day 91 of ESRD
5	1/1/2008 to 5/31/2008 and 10/1/2008 to 12/31/2008	152+ 92 = 244	Segment 1: Remains assigned to facility for 60 days after transfer out + Segment 2: Eligible starting with day 61 at facility after transfer in

For each patient period, we calculate the **stage 1**  $X_i'B$  using the comorbidity index weights table, the mean values for imputation of comorbidity index and BMI table, and the coefficients table in the Excel file. Table 3 shows these details for the example. Note

the calculations can be affected by rounding. We show only four decimal places for ease of display.

**Table 3.** Stage 1 Calculations for each patient period

Patient	Comorbidity index	$X_i'\beta$ *	Stage 1 $X_i'\beta$
1	CI=0.3123	(age)(- 0.099) + (age-14)(.129) + (CI)(0.796) + bmi(-0.309) + (vin23)(-.133) + (diab)(.286) + (year2008)(.031)+ (hispanic)(-.369) + (vin23*diab)(-.117)	-1.524
2	CI=0	(age)(- 0.099)+(age-14)(.129)+(age-60)(-.00003)+bmi(-0.309) + (black)(-.351) + (female)(-.061) + (vin01)(-.011) + (year2008)(.031) + (black*female)(.080) + (vin01*black)(.034) + (vin01*female)(.048) + (black*age)(.022) + (black*age-14)(-.029) + (black*age-60)(.008)	-0.884
3	CI=0	(age)(- 0.099) + (age-14)(.129) + (age-60)(-.00003) + bmi(-0.309) + (black)(-.351) + (vin12)(-.186) + (female)(-.061) + (year2008)(.031) + (black*female)(.080) + (black*age)(.022) + (black*age-14)(-.029) + (black*age-60)(.008)	-0.884
5	CI=0.2260	(age)(- 0.099) + (age-14)(.129) + (age-60)(-.00003) + (CI)(0.796) + bmi(-0.309)+(missc)(-.044) + (year2008)(.031) + (unknown)(-.197) + (unknown*nonwhite)(.263)	-0.084

\* CI = Comorbidity index, add weights from table in attachment

bmi = natural logarithm of body mass index

missc = missing cause of ESRD, 0 for no, 1 for yes

Next we use the **stage 1  $X_i'\beta$**  as an offset in step 2 of the model, which includes only the race-specific state population death rate as a covariate.

**Table 4.** Stage 2 Calculations for each patient period

Patient	$X_i'\beta$ *	Stage 2 $X_i'\beta$
1	(popdrtate)(.4165) + Stage 1 $X_i'\beta$	-1.429
2	(popdrtate)(.4165) + Stage 1 $X_i'\beta$	1.614
3	(popdrtate)(.4165) + Stage 1 $X_i'\beta$	1.732
5	(popdrtate)(.4165) + Stage 1 $X_i'\beta$	-0.109

\* popdrtate = log of the race-specific state population death rate

We also use the Excel file to find the baseline survival probability  $S_{0k_i}(t_i)$ , by finding the corresponding survival value given the number of days at risk in the patient period.

Table 5 shows these details for the example. Again, note the baseline survival probabilities are shown to four decimal places in this example.

**Table 5.** Baseline survival values

Patient	Days ( $t_i$ )	$S_{0k_i}(t_i)$
1	365	0.6649
2	216	0.7804
3	285	0.7262
5	244	0.7584

Finally, we calculate  $-e^{\text{stage } 2 \text{ } X_i' \beta} \ln[S_{0k_i}(t_i)]$ , the expected number of deaths for each of these patients.

**Table 6.** Calculate expected deaths for each patient

Patient	Stage 2 $X_i' \beta$	$-e^{\text{stage } 2 \text{ } X_i' \beta}$	$\ln[S_{0k_i}(t_i)]$	expected deaths
1	-1.552	-0.212	-0.408	0.086
2	-0.905	-0.405	-0.248	0.100
3	-0.905	-0.405	-0.320	0.129
5	-0.109	-0.897	-0.277	0.248

The total expected number of deaths in this facility for 2008 is the sum of the expected number of deaths for all the patient periods in that facility, or in this case 0.564. Because there was one death in the facility during 2008, the SMR is  $1/0.564$ , or 1.77.

### *Caveats*

Calculation of the SMR using this method may differ from the SMR published in the DFR for several reasons. For example, the DFR includes deaths within 60 days after transfer out of a facility, but this information may not be available to other researchers. Other differences in the calculation of days at risk will affect expected mortality and may be associated with events such as transfer, transplant, withdrawal from dialysis, hospitalization, and loss to follow-up. Differences in the coding of patient characteristics may also cause other researcher's calculations to differ from those published in the DFR.

## References

- Andersen PK, Borgun O, Gill RD, Keiding N. Statistical Models Based on Counting Processes. Springer-Verlag, New York. 1993. See pages 334 and 406-407.
- Breslow NE, Day NE. Statistical methods in cancer research (Volume II), IARC, Lyon, 1987.
- Colett D, Modeling Survival Data in Medical Research. Chapman and Hall, London, 1994. See page 153 equation 5.6 and page 151, equation 5.1.
- Cox DR. Regression Models and life tables (with discussion). J R Stat Soc 1972; 34: 197-220.
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc. 1972; 53:457-481.
- SAS Institute Inc. SAS/STAT 9.1 User's Guide. Cary, NC: SAS Institute Inc. 2004; 3213-3329.
- Zucker DM. Restricted mean life with covariates: modification and extension of a useful survival analysis method. J. Amer. Statist. Assoc. 1998; 93:702-709.
- National Center for Health Statistics. Health, *United States, 2007 With Chartbook on Trends in the Health of Americans*. Centers for Disease Control and Prevention, Health and Human Services Dept., 176-177.