

Technical Notes on the Standardized Mortality Ratio (SMR)

For the Dialysis Facility Reports

July 2020

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

Table of Contents

Introduction..... 3

Assignment of Patients to Facilities for the SMR Calculation 3

 General Inclusion Criteria for Dialysis Patients 4

 Identifying Patients Treated at Each Facility..... 4

 Days at Risk for Each Patient-Record 5

Model for Calculating Expected Mortality 6

 Missing Data 7

 P-value, Confidence Intervals, and flagging rules 8

Final Remarks 10

References..... 11

Introduction

The Standardized Mortality Ratio (SMR) in Table 3 of the Dialysis Facility Reports (DFR) is designed to compare the mortality at a facility to the mortality that would be expected given national death rates for patients with similar characteristics. Numerically, the SMR is calculated as the ratio of two numbers: the numerator (“observed”) is the actual number of deaths for the patients in a facility over a specified time period, and the denominator (“expected”) is the number of deaths that would have been expected for the same patients if they were in a facility conforming to the national norm.

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. 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.

The SMR is adjusted for age, race, ethnicity, sex, diabetes as cause of ESRD, 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 above factors.

The mortality rate for a facility in a particular calendar year is compared to the US mortality rates for dialysis patients in that same year rather than to the average mortality rates over the 4-year period. The advantage of 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. 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.

In the DFR, we also report SMR for a given region (i.e., state, network). A region’s SMR is calculated as the ratio of the total number of observed deaths among patients from that region, to the expected number of deaths for that region’s patients adjusted for the patient characteristics described below. The regional SMRs are provided for comparison purposes, so that each facility’s SMR can be compared to the SMR for the region in which it is located.

Assignment of Patients to Facilities for the SMR Calculation

This section describes the methods we use to assign patients to a facility in order to calculate the SMR. As patients can receive dialysis treatment at more than one facility in

a given year, we assign each patient day to a facility (or no facility, in some cases) based on a set of conventions that are described below.

General Inclusion Criteria for Dialysis Patients

As 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 enter a patient's follow-up into the tabulations after that patient had received chronic renal replacement therapy for more than 90 days. In other words, mortality and survival during the first 90 days do not enter into the calculations. 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 starting renal replacement therapy for the first time and for those who returned to dialysis after a transplant. That is, deaths and survival during the first 60 days do not impact the SMR of that facility.

Identifying Patients Treated at Each Facility

For each patient, we identified the dialysis provider at each point in time using a combination of Medicare dialysis claims, the Medical Evidence Form (Form CMS-2728), and data from CROWNWeb. 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, even when a patient was transferred to another facility (such as a hospital-based facility) after his or her 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 CROWNWeb data and 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 the patient 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 facilities upon receiving transplants. 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 Medicare dialysis claims nor CROWNWeb 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 reappeared, the patient was entered into analysis after 60 days of continuous therapy at a single facility. Finally, all CROWNWeb 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 CROWNWeb only require checking patient data for continued accuracy, but do not have a requirement for updating if there are not any changes.

As 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 enter a patient's follow-up into the tabulations after that patient had received chronic renal replacement therapy for more than 90 days. In other words, mortality and survival during the first 90 days do not enter into the calculations. 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 received only temporary dialysis therapy at a facility, 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.

Days at Risk for Each Patient-Record

After patient treatment histories are defined as described above, periods of follow-up time (or patient-records) are created for each patient. A patient-record 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 patient record so that the number of days at risk for any patient-record 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 gives rise to four patient-records and is analyzed the same way as would be four separate patients in that 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 so that a given patient can generate only one patient-record per year at a given facility. 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 give rise to one patient-record of 200 days at risk at the first facility, and a separate patient-record of 165 days at risk at the second facility.

The number of days at risk (t_i) in each of these patient-records is used to calculate the expected number of deaths for that patient-record as described in the "Model for Calculating Expected Mortality" 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 patient-records at the facility.

Model for Calculating Expected Mortality

The SMR is based on expected mortality calculated from a Cox model (Cox, 1972; SAS Institute Inc., 2004; Kalbfleisch and Prentice, 2002; Collett, 1994). The model used is fit in two stages. Model coefficients and baseline survival curves are available for download here: <https://dialysisdata.org/sites/default/files/content/Methodology/SMRModelInfo.xls>.

In stage I, the patient characteristics included 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 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. BMI is included as a categorical term indicating underweight, normal, overweight, or obese according to the World Health Organization criteria (WHO, 2020). Missing BMI is imputed as obese. Categorical indicator variables are included as covariates in the stage 1 model to flag records missing values for cause of ESRD, Form CMS-2728, 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. A categorical indicator variable also flags records with at least one comorbidity 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. Stratification by facility at this stage avoids biases in estimating regression coefficients that can occur if the covariate distributions vary substantially across centers. The results of this analysis are estimates of the regression coefficients in the Cox model and these provide an estimate of the relative risk for each patient.

Let p denote the number of patient characteristics in the model and x_{ij} be the specific value of the j^{th} characteristic for the i^{th} patient-record. In stage 1, for patient-record i , we denote the measured characteristics or covariates as

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

and use this to define the regression portion of a Cox model in which facilities define the strata. Note that for a categorical characteristic, the x_{ij} value is 1 if the patient falls into the category and 0 otherwise. The output of stage I is a set of regression coefficients, $\beta_1, \beta_2, \dots, \beta_p$ and the corresponding predicted value for the i^{th} patient-record is given by

$$\mathbf{X}_i\boldsymbol{\beta} = \beta_1x_{i1} + \beta_2x_{i2} + \dots + \beta_px_{ip}. \tag{1}$$

At stage II, the relative risk estimates from the first stage are used as an offset. The stage II model includes the age-adjusted population death rates for patients of that race in that state as a covariate. In the stage 2 model, there is no stratification and there is a single baseline survival curve, which is estimated along with the estimates of the stage 2

regression parameters. The estimate of the baseline survival curve also arises from the fitting of the Cox model and is analogous the Kaplan-Meier (1958) estimate, except that it is adjusted for variation among patients.

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 FY 2020 DFR used age-adjusted death rates for 2014-16 from Table 16 of the publication *Health, United States, 2017*, available at <http://www.cdc.gov/nchs/hus/contents2017.htm>.

After stage II, the linear prediction is

$$\mathbf{A}_i = \beta_0 X_{i0} + \mathbf{X}_i \boldsymbol{\beta} = \beta_0 X_{i0} + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} \quad (2)$$

Starting in Fiscal Year 2021, to account for patients starting the follow-up period at different dates during the year, left truncation has been implemented using the counting process syntax in the SAS procedure PROC PHREG (Allison, 2010). Suppose that t_{iL} and t_{iR} are the starting and end of follow-up time for patient-record i , respectively, so that $S_0(t_{iR})/S_0(t_{iL})$ is the baseline conditional survival probability at time t_{iR} conditional on that the patient start the follow-up period at time t_{iL} . The conditional survival probability for this patient-record i at time t_{iR} is:

$$S_i(t_{iR})/S_i(t_{iL}) = [S_0(t_{iR})/S_0(t_{iL})]^{\exp(A_i)}. \quad (3)$$

The expected number of deaths for this patient-record during follow-up time from t_{iL} to t_{iR} arises from considerations in the Cox model and can be written as

$$-\ln(S_i(t_{iR})/S_i(t_{iL})) = -e^{A_i} \ln [S_0(t_{iR})/ S_0(t_{iL})]. \quad (4)$$

The expected number of deaths at a given facility can now be computed simply by summing these expected values over the totality of patient-records in that facility. Specifically, the expected value is the sum over the N patient-records at the facility giving

$$\mathbf{E} = \sum_{i=1}^N -\ln[S_i(t_{iR})/S_0(t_{iL})] = -\sum_{i=1}^N e^{A_i} \ln[S_0(t_{iR})/S_0(t_{iL})]. \quad (5)$$

Let O be the total number of deaths observed at the facility during the total four year follow up period. As stated above, the SMR is the ratio of the total number of deaths observed to the expected number so that

$$\mathbf{SMR} = O/E. \quad (6)$$

Missing Data

Patients with missing data are not excluded from the model. Patients with missing diagnosis are included in the “other” diagnosis group. For the purposes of calculation, BMI missing values are included in the obese group. 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. The model also includes an indicator variable identifying patients with missing values for cause of ESRD.

P-value, Confidence Intervals, and flagging rules

The p-value measures the statistical significance (or evidence) for testing the two-sided hypothesis that the true ratio of death rates for the facility versus the nation is different (higher or lower) from 1.00. The p-value is the probability that, under the null hypothesis that the facility-specific mortality is the same as the population norm, the SMR would, just by chance, deviate from 1.00 as much as does the observed SMR, and is sometimes naively interpreted as the probability that the true SMR equals 1.00. A smaller p-value tends to occur when the ratio differs more greatly from 1.00 and when one uses more patient data to calculate the SMR value. A p-value of less than 0.05 is usually taken as evidence that the ratio of death rates truly does differ from 1.00. For instance, a p-value of less than 0.05 would indicate that the difference between the facility’s death rates and the nation’s is unlikely to have arisen from random fluctuations alone. The smaller the p-value, the more *statistically significant* the difference between national and individual facility death rates is. A small p-value helps rule out the possibility that an SMR’s variance from 1.00 could have arisen by chance. However, a small p-value does not indicate the degree of importance of the difference between the facility’s death rates and the nation’s.

The SMR’s actual quantitative value reflects the clinical importance of the difference between the facility’s and the nation’s death rates. An SMR that differs greatly from 1.00 is more important than an SMR in the range of 0.95 to 1.05.

For reporting purposes, we identify outlier facilities from amongst those with at least 3 expected deaths during the time period. Starting in Fiscal Year 2021, to address the problem of simultaneously monitoring a large number of facilities and to take account of the intrinsic unexplained variation among facilities, we used the approach described in Kalbfleisch and Wolfe (2013). This method is based on the empirical null as described in Efron (2004, 2007). The p-value for each facility is converted to a Z-score, stratified into four groups based on patient-years within each facility. The empirical null corresponds to a normal curve that is fitted to the center of each Z-score histogram using a robust M-estimation method. The mean and the standard deviation of empirical null distribution are then used for a reference distribution to compute the empirical null-based P-values, which are used to identify outlier facilities. This method aims to separate underlying intrinsic variation in facility outcomes from variation that might be attributed to poor (or excellent) care. If the empirical null-based p-value is less than 0.05 and SMR is greater than 1, the facility is said to have outcomes that are “worse than expected”. On the other hand, if the p-value is less than 0.05 and SMR is smaller than 1, the facility is said to be “better than expected”. If the p value is greater than 0.05, the facility is said to have outcomes that are “as expected”. Without empirical null

methods, a large number of facilities will be flagged, including some larger facilities with a relatively small difference between the rates of mortality. In contrast, the methods based on the empirical null adjusts for overdispersion. Using this method, facilities are flagged if they have outcomes that are extreme when compared to the variation in outcomes for other facilities of a similar size.

Mid-P-value Calculations

$$p\text{-value_min} = \text{poisson}(\text{expected deaths}, \text{observed deaths}) + \text{poisson}(\text{expected deaths}, \text{observed deaths} + 1)$$

$$p\text{-value_max} = (1 - \text{poisson}(\text{expected deaths}, \text{observed deaths} - 1)) + (1 - \text{poisson}(\text{expected deaths}, \text{observed deaths}))$$

$$p\text{-value} = \max(0.000001, \min(p\text{-value_min}, p\text{-value_max})/2)$$

Empirical Null-Based P-value Calculations

The mid-p-value for each facility is converted to a Z-score. The Z-scores across facilities are stratified into four groups based on patient-years within each facility. The empirical null corresponds to a normal curve that is fitted to the center of each Z-score histograms using a robust M-estimation method to obtain the mean and the standard deviation of the Z-scores within each group. The mean and standard deviation of empirical null distribution are then used for a reference distribution to identify outlier facilities. Specifically, let z be the original Z-score for the facility under evaluation computed based on the mid-p-value. Let μ and σ be the mean and standard deviation of the Z-scores for the corresponding group that the facility under evaluation belongs to. The empirical null-based p-value is computed as

$$P = 2 * \min(\Pr(Z > (z - \mu) / \sigma), 1 - \Pr(Z > (z - \mu) / \sigma)), \text{ where } Z \text{ is a random variable with a standard normal distribution.}$$

Empirical Null-Based Confidence Limits Calculations

Let $P(O, E)$ be the empirical null-based p-value, which is a function of the observed deaths O and expected deaths E in each facility.

If observed deaths O is less than or equal to the expected deaths E ,

If the empirical null-based p-value $P(O, E) > 0.05$

Lower CL = a value of E^* , where $0 \leq E^* \leq O$, such that the $P(O, E^*) = 0.05$

Upper CL = a value of E^* , where $E^* \geq E$, such that the $P(O, E^*) = 0.05$

If the empirical null-based p-value $P(O, E) \leq 0.05$

Lower CL = a value of E^* , where $0 \leq E^* \leq O$, such that the $P(O, E^*) = 0.05$

Upper CL = a value of E^* , where $O \leq E^* \leq E$, such that the $P(O, E^*) = 0.05$

If observed deaths O is greater than the expected deaths E ,

If the empirical null-based p-value $P(O, E) > 0.05$

Lower CL=a value of E^* , where $0 \leq E^* \leq E$, such that the $P(O, E^*) = 0.05$
Upper CL= a value of E^* , where $E^* \geq O$, such that the $P(O, E^*) = 0.05$

If the empirical null-based p-value $P(O, E) \leq 0.05$

Lower CL=a value of E^* , where $E \leq E^* \leq O$, such that the $P(O, E^*) = 0.05$

Upper CL= a value of E^* , where $E^* \geq O$, such that the $P(O, E^*) = 0.05$

Final Remarks

This document details the computation of SMR, designed to measure the performance of facilities in lower mortality. Our proposed methods are based on several statistical publications developed by the investigators of UM-KECC and have been tested in various settings. In general, the validity of any standardized measures largely depends on the validity of risk-adjustment models. As proper choices of risk adjusters are vital in the process, our practical principles lie in scientific relevance and caution of sequel. That is, we choose risk adjusters that are scientifically relevant to the outcome, while avoiding choosing those which may be affected by the quality of care.

References

- Allison, PD, *Survival Analysis Using SAS: A Practical Guide*, Second Edition, SAS Publishing, 2010.
- 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.
- Efron, B. (2004). Large-scale simultaneous hypothesis testing: the choice of a null hypotheses. *Journal of American Statistics Association*, 99(465), 96-104.
- Efron, B. (2013). *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*. Cambridge University Press, Cambridge, UK.
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.* 1972; 53:457-481.
- Kalbfleisch JD and Prentice RL. *The Statistical Analysis of Failure Time Data*. New York: Wiley 2002. See Sections 4.1 to 4.5.
- Kalbfleisch, J. D. and Wolfe, R. A. (2013). On monitoring outcomes of medical providers. *Statistics in Biosciences*, 5(2), 286-302.
- 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, 2011*. Centers for Disease Control and Prevention, Health and Human Services Dept., 110-111.
- World Health Organization (WHO). (2020). *Body Mass Index- BMI*. <http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi> . Accessed May 13, 2020.