Dialysis Facility Compare Star Rating System

Introduction

CMS has contracted with the University of Michigan KECC to create and implement a Star Quality Rating System for dialysis facilities. The Star quality ratings will be displayed on the CMS Dialysis Facility Compare (DFC) website as part of a plan to make quality information more easily available to health care consumers and other key audiences (e.g., providers and policymakers) for all Compare sites. The goal is to increase consumer knowledge, understanding, and satisfaction with the Compare sites.

The Star Rating System includes nine DFC quality measures assessing patient health outcomes and processes of care for ESRD dialysis patients. The three QMs that measure Kt/V are combined resulting in seven measures used to rate the dialysis facilities. Facilities with 5 stars are considered to deliver care that is much above average quality and those with 1 star are considered to deliver care that is rated much below average quality when compared to other dialysis facilities in the nation.

Overview of Measures

A set of eleven Quality Measures (QMs) that represent quality of care at the facility level is currently in use on DFC. The nine measures that were used in the algorithm to develop the Star Rating for facilities and are listed below:

1. Standardized Transfusion Ratio (STrR)
2. Standardized Mortality Ratio (SMR)
3. Standardized Hospitalization Ratio (SHR)
4. Percentage of adult hemodialysis (HD) patients who had enough wastes removed from their blood during dialysis: Kt/V greater than or equal to 1.2
5. Percentage of pediatric hemodialysis (HD) patients who had enough wastes removed from their blood during dialysis: Kt/V greater than or equal to 1.2
6. Percentage of adult peritoneal dialysis (PD) patients who had enough wastes removed from their blood during dialysis: Kt/V greater than or equal to 1.7
7. Percentage of adult dialysis patients who had an average calcium over the past three months greater than 10.2 mg/d (hypercalcemia)
8. Percentage of adult patients who received treatment through arteriovenous fistula (AVF)
9. Percentage of adult patients who had a catheter (tube) left in a vein longer than 90 days, for their regular hemodialysis treatment (catheter > 90)

For the purposes of analysis, the three Kt/V measures (adult PD, adult HD, and pediatric HD) were combined into a single Kt/V measure. The resulting pooled measure (All Kt/V) represents the percentage of the combined population of adult HD, adult PD, and pediatric HD patients who had enough wastes removed from their blood (Kt/V greater than or equal to specified threshold). After these measures were combined, there were seven final measures used to rate the dialysis facilities.

Analytic Approach

A straight forward way of creating an overall rating would be to use the un-weighted average of the measures. However, if some closely associated QMs measure a similar aspect of a facility and fewer QMs measure a different quality, equal weighting artificially treats the former quality as more important. This limitation is addressed by using factor analysis loadings to group the QMs into distinct factors or domains or groups, allowing us to weight the final score based on the groupings instead of individual measures.

The Star System to be reported in October 2014 on DFC was developed based on data released on the DFC website in January 2014. In this dataset, the SHR and STrR were calculated for calendar year 2012, the SMR for 2009-2012, hypercalcemia for July 2012-March 2013, and the remaining measures for April 2012-March 2013.

Standardization of Measures

The DFC QMs are clearly different in distributions as well as scales. In order to make measures comparable across facilities and to reduce the impact of outliers, we standardize the measures by using their ranks (instead of the original values) and align all the measures in the same direction. Specifically, for each QM, the facility performances are separated into 100 groups or "percentile ranks" ranging from 0.5 to 99.5 increasing by 1 where higher rank indicates a better score on a measure. To further identify facilities that performed exceptionally well or poorly, these percentile ranks (pRanks) were "normalized" or mapped from the uniform percentile rank distribution to a normal distribution (nRanks). This method allows all measures to be scored in the same manner, preventing different weighting on measures due to different distributions and scales. This method also prevents outliers from having scores that differ a lot from the other facilities while recognizing that very high or low values should be distinguished.

Quality Measure Domains

Principal factor analysis identified three domains or groups to be used for grouping the QMs. Based on the resulting factor loadings, the three empirically derived groups (domains) were found to match to related outcomes at the facility level. The three outcome measures for transfusions, mortality and hospitalization (STrR, SMR and SHR) formed the first grouping which was named the "Standardized Outcomes" (SHR, SMR, STrR), The arteriovenous fistula and catheter measures formed the second grouping which was named "Other Outcomes 1 (AV fistula, tunneled catheter)". The All Kt/V and hypercalcemia QMs formed the third grouping which was named "Other Outcomes 2 (Kt/V, hypercalcemia)". Together, these empirically derived groupings contain measures that are most correlated with one another. This is further evidence that grouped measures provide information on similar qualities about a facility.

Missing Data

Because the vascular access measures in "Other Outcomes 1 (AV fistula, tunneled catheter)" do not apply to PD patients, facilities that serve PD patients only (N=92) do not have values for measures in this domain or group. These PD-only facilities are not excluded from the ratings, but are instead rated based on the other two domains.

Facilities are given ratings as long as they have at least one measure in each of the three domains (groups), or, for PD-only facilities, in each of the two domains not related to vascular access. Among the 6,033 facilities in the January 2014 dataset, 542 (9%) were unrated. Most facilities (81%) had all seven of the measures used to create a star rating.

Missing values for facilities that qualified for ratings (i.e. for facilities that had at least one measure in each domain) are assigned median pRanks and nRanks of 50. This method of imputation ensures that one measure is not too influential in the final rating.

Overall Star Rating for each Facility

To calculate the Star rating for a facility, each domain is first given a score between 0 and 100 by averaging the normalized scores for measures within that domain. A final score between 0 and 100 is then created by averaging these three domain scores (or two domain scores for PD only facilities). Finally to ensure acknowledgement of facilities that perform exceptionally well or poorly, the facilities are assigned stars as follows based on the average of the domain scores:

- Facilities with top 10% final scores were given a star rating of 5.
- Facilities with the next 20% highest final scores were given a star rating of 4.
- Facilities within the middle 40% of final scores were given a star rating of 3.
- Facilities with the next 20% lowest final scores were given a star rating of 2.
- Facilities with bottom 10% final scores were given a star rating of 1.

In the January 2014 release dataset used for development, there was a noticeable systematic improvement of all average measure values with higher star rating.

Conclusions

This report presents an overview of the DFC Star Rating of facilities based on the groupings of correlated quality measures that are currently reported on the Medicare DFC website. In future years, if DFC measures change or new ones are added, the general algorithm described here will be used to update the measure domains used to produce the rating. In the Star System algorithm developed with January 2014 data, average measure values are consistently better with higher Overall Star Ratings.

Analysis of the ratings over time was limited because data for some measures have been available only recently. However, the data available provided evidence that the ratings are likely to be stable over time.

An advantage to the Overall Star Rating is the grouping of QMs based on systematic empirical methods, specifically, factor analysis. This method limits the possibility of overweighting QMs that measure similar qualities of facility care. The Star Rating is updated annually, to align with the annual updates of the standardized measures..