Measure Justification Form and Instructions

Project Title: Hemodialysis Vascular Access: Standardized Fistula Rate

Date:

Information included is current on February 14th 2022

Project Overview:

The Centers for Medicare & Medicaid Services (CMS) has contracted with the University of Michigan Kidney Epidemiology and Cost Center to develop facility-level measures in the area of modality education for dialysis patients. The contract name is Kidney Disease Quality Measure Development, Maintenance, and Support. The contract number is 75FCMC18D0041, task order number 75FCMC18F0001. As part of its measure development process, the University of Michigan Kidney Epidemiology and Cost Center convenes groups of stakeholders who contribute direction and thoughtful input to the measure developer during measure development and maintenance

Measure Name/Title (NQF Measure Submission Form sp.01)

Standardized Fistula Rate for Incident Patients

1. Type of Measure

☐process
☐process: appropriate use
☐outcome
☐cost/resource use
☐experience with care
☐efficiency
☒outcome: PRO/PRO-PM
☐structure
☒outcome: intermediate outcome
☐composite

2. Importance (NQF Importance to Measure and Report)

2.1 Evidence to Support the Measure Focus (for reference only) NQF Measure evaluation criterion 1a.

2.1.1 This is a Measure of

☐process:
☐process: appropriate use:
☒outcome:
☐outcome: PRO:
☐cost/resource use:

☐experience with care:
☐efficiency:
☐structure:
☐intermediate outcome:
☐composite:

2.1.2 Logic Model (NQF Measure Submission Form, Importance to Measure and Report: Evidence 1a.01)

Several observational studies have demonstrated an association between type of vascular access used for hemodialysis and patient mortality. Arteriovenous fistulae (AVF) are associated with the lowest mortality risk while long term catheters have the highest mortality. Arteriovenous grafts (AVG) have been found to have a risk of death that is higher than AVF but lower than catheters.

The measure focus is the process of assessing AV Fistula use at chronic dialysis facilities

This process leads to improvement in blood stream infections / mortality as follows:

Measure AV Fistula Rate→ Assess value →Identify patients who do not have an AV Fistula→Evaluation for an AV fistula by a qualified dialysis vascular access provider →Increase Fistula Rate →Lower patient blood stream infection / mortality.

2.1.3 Value and Meaningfulness (NQF Measure Submission Form, Importance to Measure and Report: Evidence [Outcomes] 1a.02)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (e.g., Cochrane Collaboration, AHRQ Evidence Practice Center)

☐ Other (specify)

2.1.4 Empirical Data (for outcome measures) – as applicable (NQF Measure Submission Form, Importance to Measure and Report: Evidence [Outcomes] 1a.03)

Lok CE, Huber TS, Lee T, et al; KDOQI Vascular Access Guideline Work Group. KDOQI clinical practice guideline for vascular access: 2019 update. Am J Kidney Dis. 2020;75(4)(suppl 2):S1-S164.

https://www.ajkd.org/article/S0272-6386(19)31137-0/fulltext

2.1.5 Systematic Review of the Evidence (for intermediate outcome, process, or structure quality measures, include those that are instrument-based) – as applicable (Measure Submission Form, Importance to Measure and Report: Evidence [Process] 1a.02)

2.1.6 Other Source of Evidence – as applicable (NQF Measure Submission Form, Importance to Measure and Report: Evidence [Process] 1a.13)

N/A

2.1.6.1 Briefly Synthesize the Evidence (NQF Measure Submission Form, Importance to Measure and Report: Evidence [Process] 1a.14)

N/A

2.1.6.2 Process Used to Identify the Evidence (NQF Measure Submission Form, Importance to Measure and Report: Evidence [Process] 1a.15)

N/A

2.1.6.3 Citation(s) for the Evidence (NQF Measure Submission Form, Importance to Measure and Report: Evidence [Process] 1a.16)

N/A

2.2 Performance Gap – Opportunity for Improvement (NQF Measure evaluation criterion 1b)

2.2.1 Rationale (NQF Measure Submission Form, Importance to Measure and Report: Gap in Care/Disparities 1b.01)

The newly revised NKF K/DOQI guidelines indicate that an arteriovenous access is preferable to a long-term catheter in most incident and prevalent hemodialysis patients. Furthermore, if patient circumstances are favorable, an AV fistula is preferable to an AV graft in incident dialysis patients due to fewer long-term vascular access events such as thrombosis and need for interventions. Although prior KDOQI guidelines indicated that AV fistula were associated with improved outcomes (improved patency, lower cost, lower complications) relative to AV grafts or long-term catheters, concern has been raised about significant bias in the older literature on which these conclusions had been made. Still, recent studies consistently find that AV fistula have the lowest rates of infection, and are associated with the highest survival of the three access types. Over the period from 2005-2016, the percentage of patients initiating hemodialysis with an AV fistula increased from 12% to 17%. By the middle of 2017 62.8% of prevalent hemodialysis patients were dialyzing with an AV fistula (USRDS 2018).

As the accompanying literature review indicates (see Evidence form), there are a growing number of studies reporting that creating AVF in some patients is less likely to be successful in the presence of certain comorbidities. In addition, certain patient groups may have less incremental benefit from an AV fistula relative to an AV graft. By adjusting the fistula rate for patient characteristics and comorbidities associated with low AV fistula success rates, this measure accounts for patients where a graft or even a catheter may be a more appropriate option. This approach is supported by the recent KDOQI updated guidelines which stress a patient centered approach to vascular access.

This measure is intended to be jointly reported with Hemodialysis Vascular Access: Long-term Catheter Rate. These two vascular access quality measures, when used together, consider Arterial Venous Fistula (AVF) use as a positive outcome and prolonged use of a tunneled catheter as a negative outcome. With the growing recognition that some patients have exhausted options for an AVF or have comorbidities

that may limit the success of AVF creation, joint reporting of the measures accounts for all three vascular access options. The fistula measure adjusts for patient factors where fistula placement may be either more difficult or not appropriate and acknowledges that in certain circumstances an AV graft may be the best access option. This paired incentive structure that relies on both measures (SFR, long-term catheter rate) reflects consensus best practice, and supports maintenance of the gains in vascular access success achieved via the Fistula First/Catheter Last Project over the last decade.

United States Renal Data System. 2018 USRDS annual data report: Epidemiology of kidney disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2018

Lok CE, Huber TS, Lee T, et al; KDOQI Vascular Access Guideline Work Group. KDOQI clinical practice guideline for vascular access: 2019 update. Am J Kidney Dis. 2020;75(4)(suppl 2):S1-S164.

2.2.2 Performance Scores (NQF Measure Submission Form, Importance to Measure and Report: Gap in Care/Disparities 1b.02)

TBD

2.2.3 Summary of Data Indicating Opportunity (NQF Measure Submission Form, Importance to Measure and Report: Gap in Care/Disparities 1b.03)

N/A

2.2.4 Disparities (NQF Submission Form, Importance to Measure and Report: Gap in Care/Disparities 1b.04)

TBD

2.2.5 Provide summary of data if no or limited data (NQF Submission Form, Importance to Measure and Report: Gap in Care/Disparities 1b.05)

N/A

3. Scientific Acceptability (NQF Scientific Acceptability)

3.1 Data Sample Description (NQF Measure evaluation criterion 2)

This description should be the same as 3.17 Sampling in the MIF.

3.1.1 What Types of Data Were Used for Testing? (NQF Measure Submission Form, Scientific Acceptability: Reliability - Testing 2a.01)

☐ Assessment Data

☑ Claims

☐ Electronic Health Data

☐ Electronic Health Records

☐ Instrument-Based Data

☐ Management Data

☐ Other (specify)

☐ Paper Medical Records

☑ Registry Data

3.1.2 Identify the Specific Dataset (NQF Measure Submission Form, Scientific Acceptability: Reliability - Testing 2a.02)

National CROWNWeb data from January 2018-December 2019 and Medicare claims data from January 2017 – December 2019

3.1.3 What Are the Dates of the Data Used in Testing? (NQF Measure Submission Form, Scientific Acceptability: Reliability - Testing 2a.03)

01-01-2017 to 12-31-2019

3.1.4 What Levels of Analysis Were Tested? (NQF Measure Submission Form, Scientific Acceptability: Reliability - Testing 2a.04)

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

☐ Accountable Care Organization

☐ Clinician: Clinician

☐ Clinician: Group/Practice

☐ Clinician: Individual

☑ Facility

☐ Health Plan

☐ Integrated Delivery System

☐ Other (specify)

☐ Population: Community, County or City

☐ Population: Population

☐ Population: Regional and State

Measure specified to measure performance of *(must be consistent with data sources entered in the MIF 3.22)* (NQF Measure Submission Form, Measure Specifications sp.07)

☐individual clinician
☐group/practice
☒hospital/facility/agency
☐health plan
☐other (specify) Click or tap here to enter text.

Measure tested at level of

☐individual clinician
☐group/practice
☒hospital/facility/agency
☐health plan
☐other (specify) Click or tap here to enter text.

3.1.5 How Many and Which Measured Entities Were Included in the Testing and Analysis? (NQF Measure Submission Form, Scientific Acceptability: Reliability - Testing 2a.05)

Patients on both home and in-center hemodialysis during the last HD treatment of the month from January 2018-December 2019 and starting chronic dialysis within the prior 12 months of the reporting month were included in the analyses. The number of facilities per month ranged from 6,355-6,659 and the total number of patients per month ranged from 76,012- 79,823.

Public reporting of this measure on DFC or in the ESRD QIP would be restricted to facilities with at least 11 eligible patients throughout the reporting period for the measure, meaning facilities with <11 patients would have their results suppressed. Patients at those facilities are still included in the modeling of the measure unless they otherwise meet the patient-level exclusion criteria. We have applied this restriction to all the reliability and validity testing reported here.

3.1.6 How Many and Which Patients Were Included in the Testing and Analysis? (NQF Measure Submission Form, Scientific Acceptability: Reliability - Testing 2a.06)

There were a total of 1,871,951 eligible patient-months. Among those patient-months over the whole reporting period, the average age was 64.4 years, 41.8% were female, 64.0% were white, 28.6% were black, 7.4 % reported race as “other”, 16.9% were Hispanic and 51.0% had type II diabetes as the primary cause of ESRD.

3.1.7 Sample Differences, if applicable (NQF Measure Submission Form, Scientific Acceptability: Reliability - Testing 2a.07)

N/A

3.1.8 What Were the Social Risk Factors That Were Available and Analyzed? (NQF Measure Submission Form, Scientific Acceptability: Reliability - Testing 2a.08)

Patient level:

- Employment status 6 months prior to ESRD
- Sex
- Race
- Ethnicity
- Medicare dual eligible
- ZIP code level – Area Deprivation Index (ADI) from Census data (2011-2015*). Based on patient zip-code.

Data on patient level SDS/SES factors were obtained from Medicare claims and administrative data.

*University of Wisconsin School of Medicine and Public Health. 2015 Area Deprivation Index v3.0. Downloaded from https://www.neighborhoodatlas.medicine.wisc.edu/

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.07 check patient or encounter-level data; in 2a.08 enter "see validity testing section of data elements"; and enter "N/A" for 2a.09 and 2a.10.

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.07 check patient or encounter-level data; in 2a.08 enter "see validity testing section of data elements"; and enter "N/A" for 2a.09 and 2a.10.

3.2 Reliability Testing **(for reference only)** (NQF Measure Submission Form, Scientific Acceptability: Reliability – Testing 2a)

3.2.1 Level of Reliability Testing (NQF Measure Submission Form, Scientific Acceptability: Reliability – Testing 2a.09

At what level of reliability was testing conducted? (check all that apply)

Choose one or both levels.

☐ Patient or Encounter-Level (e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements)

☑ Accountable Entity Level (e.g., signal-to-noise analysis)

3.2.2 Method of Reliability Testing (NQF Measure Submission Form, Scientific Acceptability: Reliability – Testing 2a.10)

We used January 2019 – December 2019 CROWNWeb data to calculate facility-level annual performance scores. The NQF-recommended approach for determining measure reliability is a one-way analysis of variance (ANOVA), in which the between-facility variation (σ_b^2) and the within-facility variation ($\sigma_{t,w}^2$) in the measure is determined. The inter-unit reliability (IUR) measures the proportion of the total variation of a measure (i.e., $\sigma_b^2 + \sigma_{t,w}^2$) that is attributable to the between-facility variation, the true signal reflecting the differences across facilities. We assessed reliability by calculating inter-unit reliability (IUR) for the annual performance scores. A small IUR (near 0) reveals that most of the variation of the measure between facilities is driven by random noise, indicating the measure would not be a good characterization of the differences among facilities, whereas a large IUR (near 1) indicates that most of the variation between facilities is due to the real difference between facilities.

Here we describe our approach to calculating IUR. Let $T_1,\ldots,T_N$ be the Incident Standardized Fistula Rate (SFR) for N facilities. Since the variation in $T_1,\ldots,T_N$ is mainly driven by the estimates of facility-specific intercepts ($\alpha_1,\ldots, \alpha_N$), we use their asymptotic distributions to estimate the within-facility variation in Incident SFR. Applying the delta method, we estimate the variance of T_i and denote the estimate as S_i^2. Calling on formulas from the one-way ANOVA, the within-facility variance in Incident SFR can be estimated by

$$s_{t,w}^2 = \frac{\sum_{i=1}^{N}[(n_i - 1)S_i^2]}{\sum_{i=1}^{N}(n_i - 1)}$$

and the total variation in Incident SFR can be estimated by

$$s_t^2 = \frac{1}{n'(N-1)}\sum_{i=1}^{N} n_i(T_i - \bar{T})^2,$$

where n_i is the number of subjects in the i^{th} facility, $\bar{T} = \sum n_i T_i \,/\, \sum n_i$, and $n' = \frac{1}{N-1}(\sum n_i - \sum n_i^2 \,/\, \sum n_i)$ is approximately the average facility size (number of patients per facility). Thus, the IUR = $\sigma_b^2 / (\sigma_b^2 + \sigma_{t,w}^2)$ can be estimated by $(s_t^2 - s_{t,w}^2)/s_t^2$.

To assess more directly the value of the measure in identifying providers with extreme outcomes, we also computed an additional metric, termed the profile IUR (PIUR). This was to address the challenge that the IUR could be small with many providers having outcomes around the national norm, even though the measure may still be able to identify facilities with extreme outcomes. The PIUR, based on the measure's ability to consistently flag extreme providers, was computed with a two-step approach: first, we evaluated the ability of a measure to consistently profile facilities with extreme outcomes; second, we mapped this reflagging ability to an IUR value computed by assuming no outlier facilities. This value was defined to be the PIUR. The difference between the PIUR and the IUR indicates the extent to which the measure identifies outliers.

The reliability calculation only included facilities with at least 11 patients during the two-year period.

3.2.3 Statistical Results from Reliability Testing (NQF Measure Submission Form, Scientific Acceptability: Reliability - Testing 2a.11)

The IUR is 0.705, which indicates that 70.5% of the variation in the Incident SFR can be attributed to between-facility differences in performance (signal) and 29.5% to the within-facility variation (noise).

The PIUR is 0.970 which is higher compared to the IUR, indicating the existence of outlier facilities that can be identified by the measure but were not captured by the IUR.

3.2.4 Interpretation (NQF Measure Submission Form, Scientific Acceptability: Reliability – Testing 2a.12)

The result of IUR and PIUR testing suggests a high degree of reliability.

3.3 Validity Testing **(for reference only)** (NQF Measure Submission Form, Scientific Acceptability: Validity - Testing 2b)

3.3.1 Level of Validity Testing (NQF Measure Submission Form, Scientific Acceptability: Validity – Testing 2b.01)

At what level(s) of validity was testing conducted? (check all that apply)

☐critical data elements (Note: Data element validity must address all critical data elements.)
☐performance measure score
☒empirical validity testing
☐systematic assessment of face validity of quality measure score as an indicator of quality or resource use (i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance)

Provide empirical validity testing at the time of maintenance review; if not possible, provide justification.

3.3.2 Method of Validity Testing (NQF Measure Submission Form, Scientific Acceptability: Validity – Testing 2b.02)

Validity was assessed using Poisson regression models to measure the association between facility level quintiles of performance scores and the 2018-2019 Standardized Mortality Ratio (SMR, NQF 0369), 2018-2019 Standardized Hospitalization Ratio (SHR, NQF 1463), and 2018 First Year Standardized Mortality Ratio (SMR) respectively. Facility-level performance scores were divided into quintiles (Q1 to Q5), and the relative risk (RR) for SMR (and SHR and first year SMR, separately) was calculated for each quintile, using Q5 as the reference group. A RR>1.0 would indicate a higher relative risk of mortality or hospitalization, compared to the lowest performance score quintiles.

For the all-cause hospitalization rate and vascular access infection related hospitalization rate, we used linear regression to test the association between the SFR quintiles and the 2018-2019 all-cause hospitalization rate, and 2018-2019 vascular access related infection hospitalization rate, respectively. For all-cause hospitalization and vascular access related infection hospitalization, the respective rate was calculated for each quintile and a trend test of the rates was performed.

- SMR: We expect a negative association with SMR since successfully creating an AVF is generally seen as representing a robust process to coordinate care outside of the dialysis facility. AVFs are also associated with lower risk of infection which may reduce the risk of a life threatening infection or other poor outcomes that place patients at higher risk of mortality. Higher standardized fistula rates will be negatively associated with SMR.

- SHR: We expect a negative association with SHR since successfully creating an AVF is generally seen as representing a robust process to coordinate care outside of the dialysis facility, and potentially reduces the risk for patients at such facilities going to hospital due to infections or other acute clinical events. Higher standardized fistula rates will be negatively associated with SHR.

- First Year SMR: We expect a negative association with the first year SMR as many incident patients begin with a catheter, and therefore face higher risk for infection compared to patients with an AVF. AVFs are associated with lower risk of infection which may reduce the risk of a life threatening infection or other poor outcomes that place patients at higher risk of mortality particularly in their first year of dialysis. Higher standardized fistula rates will be negatively associated with the first year SMR.

- All-cause hospitalization rate: We expect a negative association between all-cause hospitalization rates and higher AVF rates given the known risk of infection and other complications related to long-term catheter dependence, particularly in incident patients.

- Vascular access related infection hospitalization rate: We expect a negative association between access related infection hospitalizations and AVF rates because of the higher rates of catheter in patients in the first year of dialysis, which creates a higher risk of a catheter related infection.

3.3.3 Statistical Results from Validity Testing (NQF Measure Submission Form, Scientific Acceptability: Validity – Testing 2b.03)

Cut-points for the quintiles of the performance scores were defined as follows:

Q1: 0% - <30.8%

Q2: 30.8% - <38.3%

Q3: 38.3% - <44.6%

Q4: 44.6 - <52.1%

Q5: 52.1% - <99.0% as Reference

Results from the Poisson model indicated that the percent of patient-months with a fistula was significantly associated with the risks of mortality and hospitalization.

For the 2018-2019 SMR, the relative risk of mortality increased as the performance measure quintile decreased from the reference group (Q5)with the highest risk in quintile 1. For quintile 4, RR=1.02 (95%

CI: 1.00, 1.04; p<0.001), quintile 3, RR=1.06 (95% CI: 1.04, 1.08; p<0.001), quintile 2, RR=1.08 (95% CI: 1.06, 1.10; p<0.001), and quintile 1, RR=1.13 (95% CI: 1.11, 1.15; p<0.001).

Similarly for 2018-2019 SHR, the relative risk of hospitalization increased as the performance measure quintile decreased from the reference group (Q5) with the highest risk in quintile 1. For quintile 4, RR=1.06 (95% CI: 1.05, 1.06; p<0.001), quintile 3, RR=1.07 (95% CI: 1.06, 1.07; p<0.001), quintile 2, RR=1.11 (95% CI: 1.10, 1.12; p<0.001), and quintile 1, RR=1.15 (95% CI: 1.14, 1.15; p<0.001).

For the 2018 first year SMR, the relative risk of mortality increased as the performance measure quintile decreased from the reference group (Q5)with the highest risk in quintile 1. For quintile 4, RR=1.08 (95% CI: 1.03, 1.14; p=0.002), quintile 3, RR=1.11 (95% CI: 1.05, 1.16; p<0.001), quintile 2, RR=1.17 (95% CI: 1.12, 1.23; p<0.001), and quintile 1, RR=1.53 (95% CI: 1.46, 1.60; p<0.001).

For the 2018-2019 all-cause hospitalization, the hospitalization rate decreased as the performance measure quintile increased. Hospitalization rates for quintiles 1 to 5 are 1.06, 0.99, 0.95, 0.93, and 0.87 patient-years respectively (trend test p<0.001).

For the 2018-2019 vascular access related infection hospitalization, the hospitalization rate decreased as the performance measure quintile increased. Hospitalization rates for quintiles 1 to 5 are 0.22, 0.18, 0.17, 0.16, and 0.15 respectively (trend test p<0.001).

3.3.4 Interpretation (NQF Measure Submission Form, Scientific Acceptability: Validity – Testing 2b.04)

The results of the Poisson regression and trend test suggest that lower fistula use is associated with higher risk of mortality and hospitalization (measured by the respective standardized mortality, standardized hospitalization, and first year standardized mortality ratios), as well as all-cause and vascular access infection related hospitalization (measured by the hospitalization rates), as compared to facilities with higher standardized fistula rates.

3.4 Exclusions Analysis **(for reference only)** (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b)

3.4.1 Method of Testing Exclusions (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.16)

The following exclusions are applied to the denominator:

Patients with a catheter that have limited life expectancy. Limited life expectancy is defined as:

- Patients under hospice care in the current reporting month
- Patients with metastatic cancer in the past 12 months
- Patients with end stage liver disease in the past 12 months

- Patients with coma or anoxic brain injury in the past 12 months

The facility-level standardized fistula rate with and without the patient-month exclusions are calculated and compared.

3.4.2 Statistical Results from Testing Exclusions (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.17)

The following tables show the percent of patient-months at risk and the number of unique patients excluded as a result of the above mentioned exclusion strategy.

Table 1: Percent of patient-months at risk excluded, 2018-2019 data

Year	Before Exclusion	After Exclusion	Percent
2018-2019	1,915,119	1,864,647	2.64%

Table 2: Number and percent of unique patients excluded, 2018 data

Year	Before Exclusion	After Exclusion	Percent
2018-2019	276,606	270,702	2.13%

Table 3: Distribution of performance scores before and after the exclusion, 2018-2019 data

Standardized Fistula Rate	N	Mean	Std Dev	Minimum	Maximum	Median	Lower Quartile	Upper Quartile
Before exclusion	6664	40.369	12.459	0	94.425	40.472	31.966	48.619
After exclusion	6664	41.445	12.656	0	99.02	41.563	32.932	49.945

Figure 1: Scatterplot –Incident SFR with and without measure exclusions, 2018-2019 data

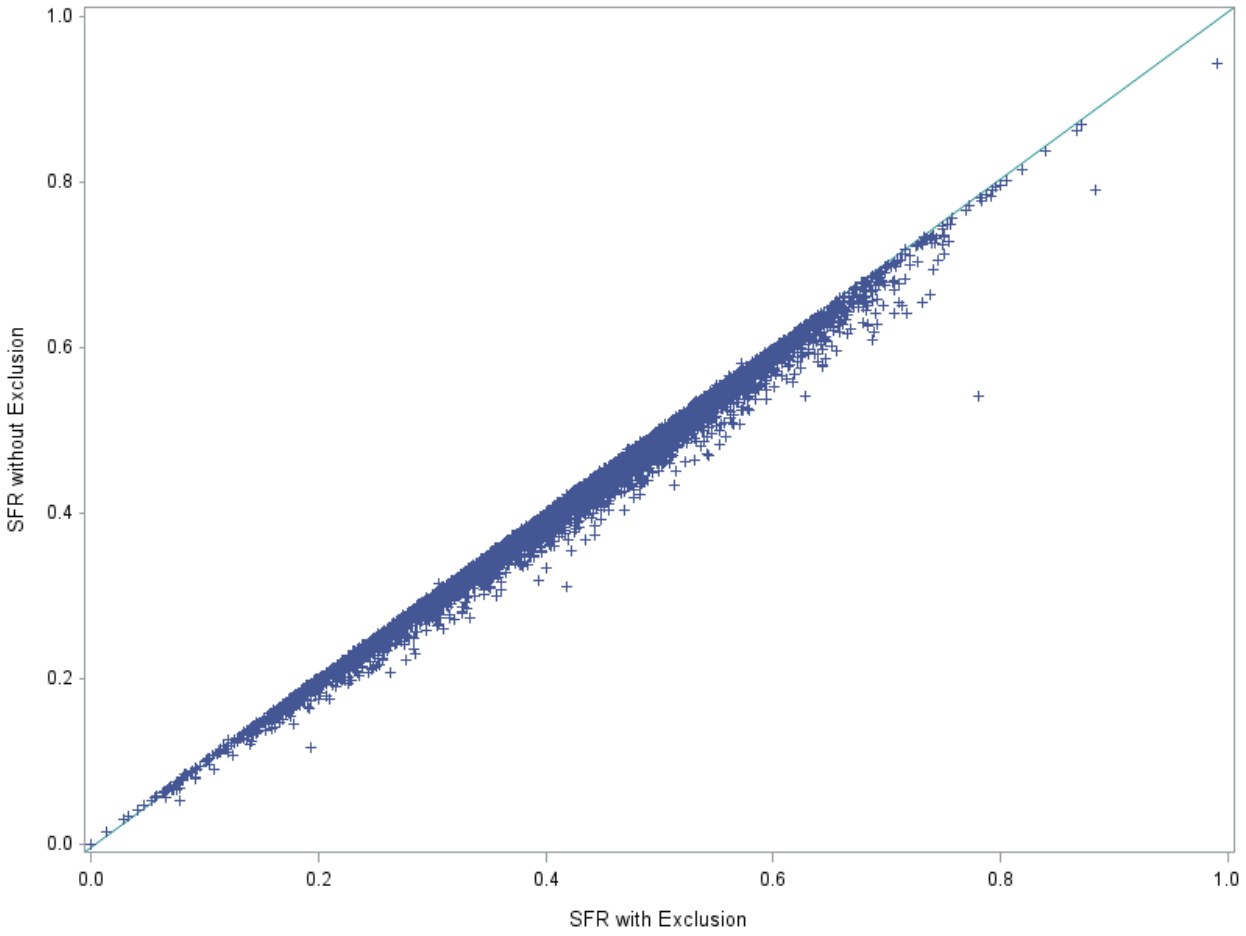
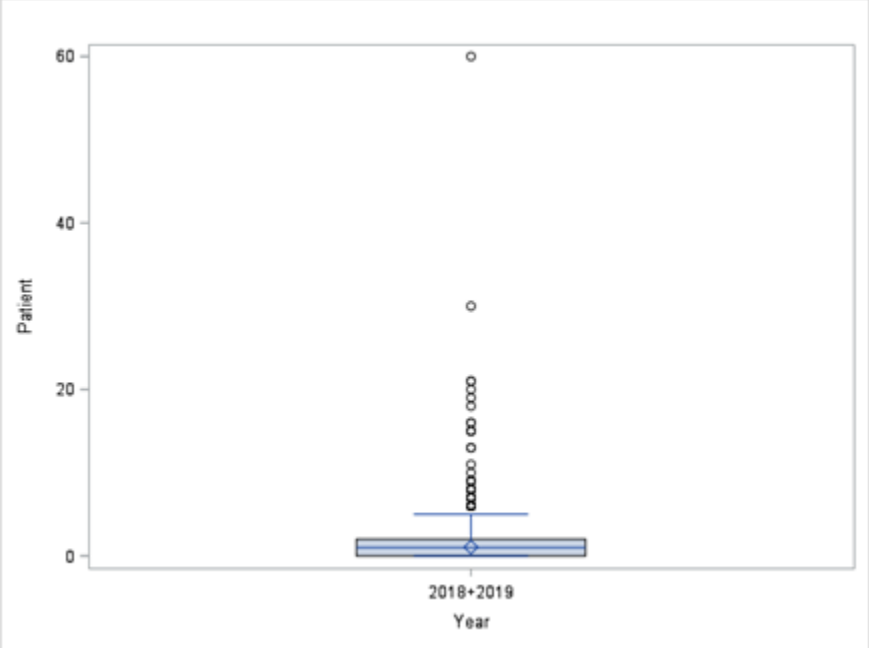


Figure 2. Distribution of Excluded Patients at facility level for 2018-2019



3.4.3 Interpretation (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.18)

Using 2018-2019 data, we show the exclusion criteria are necessary since the percentage of patients excluded at each facility is not evenly distributed across facilities (distribution shown in the boxplot above, figure 2). Due to the unequal distribution across facilities, the exclusion criteria take into account that some facilities treat a higher proportion of patients with limited life expectancy. Additionally, our results shown in both the scatter-plot (Figure 1) as well as the Pearson Correlation Coefficient of 0.996 (p-value <0.0001) between Incident SFRs with and without the exclusion suggest that the overall impact of the exclusion on the measure's validity is not substantial since the two are highly correlated.

3.5 Risk Adjustment or Stratification for Outcome or Resource Use Measures **(for reference only)** (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b)

3.5.1 Method of Controlling for Differences (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.19)

The method of controlling for differences in case mix is

☐no risk adjustment or stratification
☒statistical risk model with (specify number) risk factors
☐stratification by (specify number) risk categories
☐other (specify) Click or tap here to enter text.

3.5.2 Rationale for Why There Is No Need for Risk Adjustment (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.21)

N/A

3.5.3 Conceptual, Clinical, and Statistical Methods (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.20)

The Incident SFR measure is a directly standardized percentage, in that each facility's percentage of AVF use is adjusted to the national distribution of covariates (risk factors) (with 'national' here referring to all-facilities-combined). The Incident SFR for facility *i* is an estimate of what the facility's percentage of AVF would equal if the facility's patient mix was equal to that of the nation as a whole. The measure is adjusted for patient demographic and clinical characteristics based on a logistic regression model. This model includes the facility indicators and assumes that the regression coefficients of the risk factors are the same across all facilities. The common risk factor effects are assumed in order to improve computational stability in estimating facility-specific effects.

The patient characteristics included in the logistic regression model as covariates are:

- Age
    - 18-<25
    - 25-<59
    - 60-<75
    - 75+
- BMI at incidence
    - <18.5
    - 18.5-25
    - 25-30
    - >=30
- Nursing home status in the prior 12 months
    - No nursing home care: 0 days
    - Short term nursing home: 1-89 days
    - Long term nursing home: ≥90 days
- Nephrologist's care prior to ESRD (CMS-2728 form)
- Diabetes as primary cause of ESRD

- Comorbidities at ESRD incidence (CMS-2728 form)
    - Diabetes
    - Congestive heart failure
    - Other heart diseases
    - Peripheral vascular disease
    - Cerebrovascular disease
    - Chronic obstructive pulmonary disease
    - Drug dependence
    - Inability to ambulate/transfer at ESRD incidence
- Incident comorbidities: Incident comorbidity information was obtained from the CMS-2728. The covariates for comorbidities included in the final model take a value of 1 if there was any evidence of the condition from the CMS-2728, otherwise 0. We also use two binary indicators:
    - Indicator for missing a CMS-2728 form: this would signal an inability to capture incident comorbidities.
    - Indicator for having at least one of the ESRD incident comorbid conditions listed above: this would signal the possibility that a patient has a CMS-2728 form, but the comorbidity section could have been left blank.
- Medicare Indicator:
    - Medicare coverage for at least 6 months or more in the prior 12 months OR, Medicare Advantage coverage for one or more months in the prior 12 months. We used 1 month of MA because ESRD patients were restricted from enrolling in MA plans (prior to 2021). They already had to be Medicare beneficiaries in an MA plan prior to becoming ESRD therefore they are considered as having a sufficient Medicare coverage history during this period.

3.5.4 Conceptual Model of Impact of Social Risks (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.22)

☒published literature
☒internal data analysis
☐other (specify) Click or tap here to enter text.

3.5.5 Statistical Results (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.24)

In the table below, we list results from the adjusted model described above. For a given covariate, the regression coefficient represents the logit of the rate. We also report the odds ratio for each covariate. With a few exceptions (Diabetes not as primary cause of ESRD, and at least one of the comorbidities), all main effects are statistically significant at the ≤ 0.05 level.

Table 4. Model Coefficients and Odds Ratios, using 2018-2019 data

Covariate	Coefficient	Odds Ratio	P-value
Age	**	**	**
18-<25	-0.326	0.722	0.002
25-<59	0.044	1.045	0.036
60-<75	reference	**	**
75+	-0.140	0.869	<0.001
BMI	**	**	**
Underweight (< 18.5)	-0.306	0.737	<0.001
Normal (18.5 - 24.9)	reference	**	**
Overweight (25-29.9)	0.167	1.181	<0.001
Obese (30+)	0.206	1.228	<0.001
Nursing home during the prior 12 months	**	**	**
No nursing home care (0 day)	reference	**	**
Short-term nursing home care (<90 days)	-0.621	0.537	<0.001
Long-term nursing home care (90+ days)	-0.600	0.549	<0.001
Nephrologist's Care prior to ESRD*	0.560	1.750	<0.001
Primary Cause of ESRD	**	**	**

Covariate	Coefficient	Odds Ratio	P-value
Diabetes	0.112	1.119	<0.001
Other	reference	**	**
Comorbidities*	**	**	**
Diabetes (NOT as primary cause of ESRD)	-0.061	0.941	0.060
Congestive Heart Failure	-0.235	0.790	<0.001
Other Heart Diseases	-0.088	0.916	<0.001
Peripheral Vascular Disease	-0.070	0.933	0.028
Cerebrovascular Disease	-0.085	0.918	0.007
Chronic Obstructive Pulmonary Disease	-0.155	0.856	<0.001
Drug Dependence	-0.369	0.691	<0.001
Inability to ambulate/transfer	-0.491	0.612	<0.001
At least 6 months of Medicare covered months OR at least 1 month of MA covered months in prior 12 months	0.547	1.728	<0.001
Missing a CMS-2728 form	-1.487	0.226	<0.001
At least one of the comorbidities listed above	-0.063	0.939	0.068

*** "No" was used as the reference.**

**** This cell is intentionally left blank.**

3.5.6 Analyses and Interpretation in Selection of Social Risk Factors (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.25)

The table below shows the regression coefficient estimates and odds ratio for patient and area level SDS/SES variables based on a logistic regression model for AV fistula use that included all these variables along with all the other clinical covariates used for adjustment in the Incident SFR. Here we only report

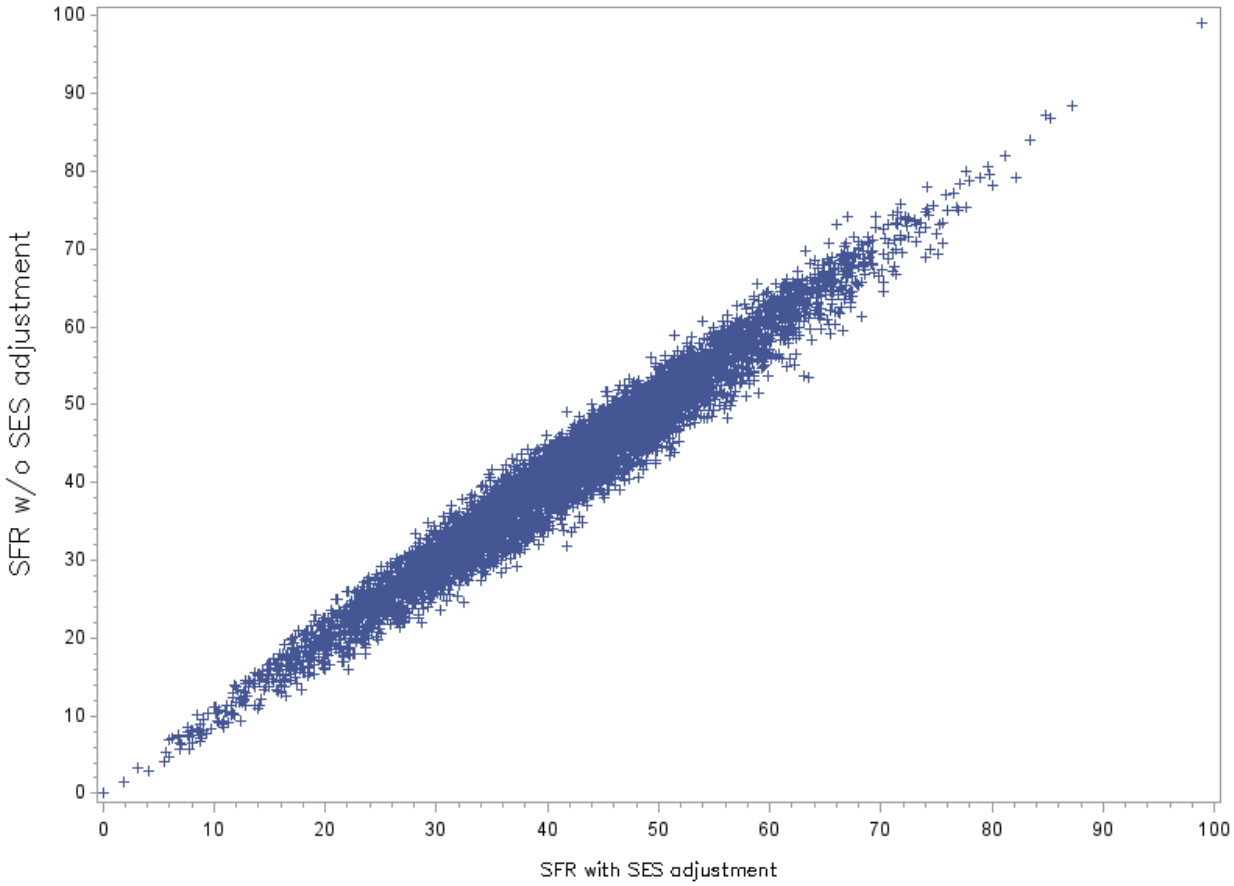
results for the SDS/SES factors.

Table 5. Coefficients and odds ratios for SDS/SES variables, using 2018-2019 data

Variable	Estimate	Odds Ratio	P-value
Sex	*	*	*
Female	-0.506	0.603	<0.001
Male	Reference	*	*
Ethnicity	*	*	*
Hispanic	0.100	1.105	0.019
Non-Hispanic	Reference	*	*
Race	*	*	*
White	Reference	*	*
Black	-0.241	0.786	<0.001
Other	0.122	1.130	0.030
Employment Status (2728)	*	*	*
Employed	Reference	*	*
Unemployed	-0.188	0.829	<0.001
Other	-0.214	0.807	<0.001
Medicare Coverage	*	*	*
Dual eligible	-0.021	0.979	0.462
Non dual eligible	Reference	*	*
ADI (zipcode_level)	*	*	*
National percentile ADI score	-0.001	0.999	0.288

*This cell is intentionally left blank.

Figure 3. Correlation between Incident SFRs with and without SDS/SES adjustment, Data year 2018-2019



The standard and SDS/SES-adjusted Incident SFRs were highly correlated at 0.98 (p<.001).

Table 6. Comparison of performances with vs. without adjusting for SDS/SES factors, using 2018-2019 data

*	*	**Incident SFR with SDS/SES**	**Incident SFR with SDS/SES**	**Incident SFR with SDS/SES**	**Incident SFR with SDS/SES**
*	*	**Better than expected**	**As expected**	**Worse than expected**	**Total**
Incident SFR w/o SDS/SES	**Better than expected**	161 (2.4%)	31 (0.5%)	0	192
Incident SFR w/o SDS/SES	**As expected**	43 (0.7%)	6083 (91.3%)	35 (0.5%)	6161
Incident SFR w/o SDS/SES	**Worse than expected**	0	36 (0.5%)	275 (4.1%)	311
Incident SFR w/o SDS/SES	**Total**	204	6150	310	6664

*This cell is intentionally left blank.

After adjustment for SDS/SES, 145 facilities (2.2%) changed performance categories. 66 (1.0%) facilities were down-graded and 79 (1.2%) facilities were upgraded.

Black race, female sex, unemployment at ESRD incidence, and dual-eligible status were all associated with lower odds of having an AVF indicating that patient-level, but not area-level, variables for SDS/SES have some impact on expected performance of the incident SFR. Furthermore, we observed that adjustment for SDS/SES only minimally shifted facility performance, with slightly more facilities declining in performance ranking with SDS/SES adjustment than improved.

Patient-level SDS/SES variables are not included as adjustments in the measure due to the absence of a

convincing biological or clinical rationale that warrant accounting for different outcomes on the basis of race, sex, or socioeconomic status. Adjusting for these factors could have the unintended effect of masking or reinforcing disparities in vascular access outcomes. Some providers in the dialysis community believe that women are less likely to have AVF due to smaller vessels and hypothesize that this may be a biologic explanation for subsequent higher primary failure rates seen in women. While several studies have reported that women have smaller vasculature than men [1,2] this has not been a consistent finding, and may be isolated to forearm vessels. Studies that have focused on upper arm AVF have demonstrated similar AVF rates between men and women, suggesting a lack of sufficient biological or clinical support for different outcomes in fistula rates between female and males [3].

Finally, area-level factors are not included as an adjustment due to the absence of clinically meaningful or statistically observed differences on the fistula rate with these adjustments.

1. Jemcov, TK Morphologic and functional vessel characteristics assessed by ultrasonography for prediction of radiocephalic fistula maturation. J Vasc Access 2013; 14(4):356-363

2. Allon, M et al. Effect of preoperative sonographic mapping on vascular access outcomes in hemodialysis patients. Kidney International, Vol. 60 (2001), pp. 2013–2020

3. Caplin, N et al. Venous Access:Women Are Equal. Am J Kidney Dis 2003. 41:429-432.

3.5.7 Method Used to Develop the Statistical Model or Stratification Approach (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.26)

See sections 2b.20 and 2b.27-2b.29, which describe the statistical methods used to develop and validate the model. Risk factors were selected for the final model based on the magnitude of the coefficients, evaluation of their statistical and clinical significance, and the model c-statistic.

3.5.8 Statistical Risk Model Discrimination Statistics (e.g., c-statistic, R^2) (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.27)

The C-statistic was 0.748. This indicates that the model correctly ordered 75% of the pairs of patient-months that were discordant with respect to the response variate.

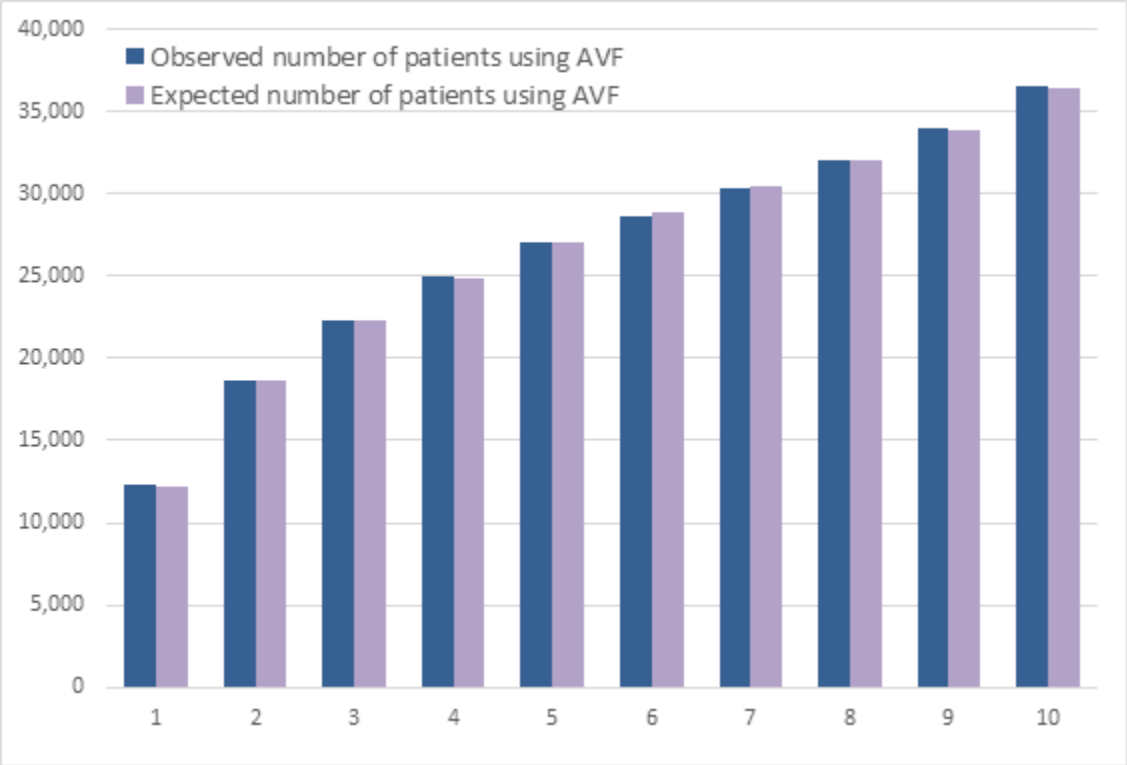
3.5.9 Statistical Risk Model Calibration Statistics (e.g., Hosmer-Lemeshow statistic) (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.28)

The Hosmer-Lemeshow test statistic based on deciles of risk is 9.7 with p-value=0.29. The c-statistic and risk decile plot show that the model provides an overall good fit to the data.

3.5.10 Statistical Risk Model Calibration—Risk decile plots or calibration curves (NQF Measure Submission Form: Other Threats to Validity [Exclusions, Risk Adjustment] 2b.29)

The preferred file format is .png, but most image formats are acceptable.

Figure 4: Decile plots for the number of patients using AVF, using 2018-2019 data



3.5.11 Results of Risk Stratification Analysis (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity (Exclusions, Risk Adjustment) 2b.30)

The decile plot (Figure 4) shows that the risk factors in the model are discriminating well between patients.

3.5.12 Interpretation (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.31)

There is good separation among all 10 groups by risk scores, and the ordering is as predicted by the model (i.e., patients predicted to have a lower probability of AVF use actually do have a lower percentage of AVF use). The absolute differences between the risk groups are also large, with patients predicted to have the highest likelihood of AVF use (Group 10) having 3 times higher AVF rate than those predicted to have the lowest likelihood (Group 1). This means that the model fit is good and therefore adequately adjusts for patient characteristics (case mix).

3.5.13 Optional Additional Testing for Risk Adjustment (NQF Measure Submission Form, Scientific Acceptability: Validity - Other Threats to Validity [Exclusions, Risk Adjustment] 2b.32)

N/A

3.6 Identification of Meaningful Differences in Performance **(for reference only)** (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b)

3.6.1 Method (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b.05)

Differences in measure performance were evaluated separately for each facility, where the incident standardized fistula rate (SFR) of each facility was compared to the overall national distribution.

Here we describe our approach for testing of statistical significance. Let $T_1,\ldots,T_N$ be the Incident Standardized Fistula Rate (SFR) for N facilities. Since the variation in $T_1,\ldots,T_N$ is mainly driven by the estimates of facility-specific intercepts ($\alpha_1,\ldots, \alpha_N$), we use their asymptotic distributions and apply the delta method to estimate the standard errors of SFRs. Let S_i denote the standard error estimate of T_i. The test-statistic is then calculated by (T_i - national average of SFR)/ S_i, which asymptotically follows the standard normal distribution under the null hypothesis. A two-sided test with significant level 0.05 was used. As the reference null distribution, we used Efron's empirical null distribution in lieu of the theoretical null distribution since the empirical null method is more robust approach that takes account of the national random variation among facilities not accounted for in the model (Efron, 2004; Kalbfleisch and Wolfe, 2013). It essentially rescales the critical value for the test statistic. The rescaling multiple is estimated by the slope (estimated via robust regression) correlating the empirical and theoretical Z-score quantiles (e.g., with a multiple of 1 indicating that in fact no rescaling is required). In this approach, facilities are flagged if they have outcomes that are more extreme when compared to the variation in national AVF rate for the incident ESRD population.

Efron, Bradley. Large-Scale Simultaneous Hypothesis Testing: The Choice of a Null Hypothesis. Journal of the American Statistical Association. Vol. 99, No. 465 (Mar., 2004), pp. 96-104

Kalbfleisch JD, Wolfe RA. On monitoring outcomes of medical providers. Statistics in the Biosciences. November 2013, Volume 5, Issue 2, pp 286-302

3.6.2 Statistical Results (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b.06)

Examples may include number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined.

Table 7. Proportion of facilities with statistically significant differences (p-values < 0.05), using 2018-2019 data

Category	Number of facilities	Percent of facilities
Better than expected	192	2.88
As expected	6161	92.45
Worse than expected	311	4.67

3.6.3 Interpretation (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b.07)

For the 2018-2019 Incident SFR, 6161 (92%) facilities have achieved expected performance, 311 (4.7%) facilities have performed worse than expected (lower fistula rate), 192 (2.9%) facilities have performed better than expected (higher fistula rate). In general, a higher rate of fistula use represents better quality of care. This analysis demonstrates both practical and statistically significant differences in performance across facilities based on their adjusted proportion of patient months with a fistula in use.

3.7 Comparability of Multiple Data Sources/Methods **(for reference only)** (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b)

3.7.1 Method (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b.12)

N/A

3.7.2 Statistical Results (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b.13)

N/A

3.7.3 Interpretation (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b.14)

N/A

3.8 Missing Data Analysis and Minimizing Bias **(for reference only)** (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data])

3.8.1 Method (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b.08)

The Standardized Fistula Rate measure includes all patients regardless of Medicare coverage. The measure is based on data from CROWNWeb (representative of all ESRD dialysis patients) and Medicare claims. Missing data for vascular access type occurs rarely. We report the frequency of missing for the data below (patient-month level).

- Missing CMS 2728 – All ESRD patients are required to have a CMS 2728 form submitted to CMS regardless of the patient's Medicare status. The 2728 certifies the patient has ESRD. The 2728 data are in CROWNWeb.

- Missing BMI from the 2728 – see above. This is part of the required fields in the 2728.

- The overall percentage of patient months with missing vascular access type – In CROWNWeb, "Access Type IDs" were reported for each patient at each month. A vascular access type is considered missing if no access type ID was found at the given month.

- Patient months where we are unable to determine presence of comorbidities for the limited life expectancy exclusions. For each month we search patients' Medicare claims for the presence of any comorbidity exclusion conditions in the past 12 months. If no claims were found, we consider the months as "unable to determine the presence of comorbidities".

  Incident SFR uses CROWNWeb clinical data along with other CMS administrative data for several important components of the measure calculation. These include comorbidities at ESRD incidence for risk adjustment from the CMS 2728, and claims-based conditions for the limited life expectancy exclusions. The source of vascular access type and several clinical risk factors are from CROWNWeb (registry that includes all ESRD dialysis patients).

3.8.2 Missing Data Analysis (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b.09)

Summary findings:

1. Patient-months with missing CMS 2728 is 1.11% and missing BMI on the CMS 2728 is 0.01% from the measure.

2. 3.6% of overall patient months are missing vascular access type.

Table 8. Frequency of missing data elements, 2018-2019 data

Data Element	Missing (n)	Total (n)	Missing (%)
Patient months with missing CMS 2728	20,038	1,864,647	1.07%
Patient months without BMI reported on 2728	235	1,844,609	0.01%
Vascular access type	57,801	1,864,647	3.1%
Patient months where we are unable to determine the presence of limited life expectancy exclusion conditions*	614,367	1,864,647	32.95%

3.8.3 Interpretation (NQF Measure Submission Form: Scientific Acceptability: Validity - Threats to Validity [Statistically Significant Differences, Multiple Data Sources, Missing Data] 2b.10)

N/A

4. Feasibility (NQF Feasibility Criterion 3)

4.1 Data Elements Generated as Byproduct of Care Processes (NQF Measure Submission Form, Feasibility 3.01)

Data used in the measure are (check all that apply)

☒generated or collected by and used by healthcare personnel during provision of care (e.g., blood pressure, laboratory value, diagnosis, depression score)
☒coded by someone other than the person obtaining original information (e.g., Diagnosis-Related Group [DRG], International Classification of Diseases, 10th Revision, Clinical Modification/Procedure Coding System [ICD-10-CM/PCS] codes on claims)
☐abstracted from a record by someone other than the person obtaining original information (e.g., chart abstraction for quality measure or registry)
☐other (specify) Click or tap here to enter text.

4.2 Electronic Sources

4.2.1 Data Elements Electronic Availability (NQF Measure Submission Form, Feasibility 3.02.)

☐All data elements are in defined fields in EHRs.
☐All data elements are in defined fields in electronic claims.
☐All data elements are in defined fields in electronic clinical data such as clinical registry, nursing home MDS, and home health OASIS.
☒All data elements are in defined fields in a combination of electronic sources.
☐Some data elements are in defined fields in electronic sources.
☐No data elements are in defined fields in electronic sources.
☐Data are patient/family reported information; may be electronic or paper.

4.2.2 Path to Electronic Capture (NQF Measure Submission Form, Feasibility 3.03)

N/A

4.2.3 eCQM Feasibility (NQF Measure Submission Form, Feasibility 3.05)

N/A

4.3 Data Collection Strategy

4.3.1 Data Collection Strategy Difficulties (optional) (Measure Submission Form, Feasibility 3.06)

None Identified

4.3.2 Fees, Licensing, Other Requirements (NQF Measure Submission Form, Feasibility 3.07)

N/A

5. Usability and Use (NQF Usability and Use Criterion 4)

5.1 Use (NQF Measure evaluation criterion 4a)

5.1.1 Current and Planned Use (NQF Measure Submission Form, Use 4a.01 and 4a.02)

☐public reporting
☐public health or disease surveillance
☐payment program
☐regulatory and accreditation programs
☐professional certification or recognition program
☐quality improvement with external benchmarking to multiple organizations
☐quality improvement internal to a specific organization
☒not in use
☐use unknown

5.1.1.1 Reasons for Not Publicly Reporting or Use in Other Accountability Application (NQF Measure Submission Form, Use 4a.03)

The measure is undergoing initial endorsement review.

5.1.1.2 Plan for Implementation (NQF Measure Submission Form, Use 4a.04)

CMS will determine if/when to report this measure in a public reporting/payment program. Potential applications for the measure include the ESRD Quality Incentive Program (ESRD QIP) or the Dialysis Facility Care Compare website.

5.1.2 Feedback on the Measure by Those Being Measured or Others (NQF Measure Submission Form, Use 4a.05)

N/A

5.1.2.1 Technical Assistance Provided During Development or Implementation (NQF Measure Submission Form, Use 4a.06)

N/A

5.1.2.2 Technical Assistance with Results (NQF Measure Submission Form, Use 4a.06)

N/A

5.1.2.3 Feedback on Measure Performance and Implementation (NQF Measure Submission Form, Use 4a.07)

N/A

5.1.2.4 Feedback from Measured Entities (NQF Measure Submission Form, Use 4a.08)

N.A

5.1.2.5 Feedback from Other Users (NQF Measure Submission Form, Use 4a.09)

N/A

5.1.2.6 Consideration of Feedback (NQF Measure Submission Form, Use 4a.10)

N/A

5.2 Usability (NQF Measure evaluation criterion 4b)

5.2.1 Improvement (NQF Measure Submission Form, Usability 4b.01)

See Importance to Measure and Report for data on performance gap and disparities.

The incident SFR is not yet implemented in a public reporting program, so improvement could not be evaluated. CMS currently anticipates implementation of this measure after endorsement review. Once implemented, facility performance on this measure can be evaluated to determine if the measure has supported and detected quality improvement in incident fistula rates.

5.2.2 Unexpected Findings (NQF Measure Submission Form, Usability 4b.02)

N/A

5.2.3 Unexpected Benefits (NQF Measure Submission Form, Usability 4b.03)

N/A

6. Related and Competing Measures (NQF Related and Competing Criterion 5)

6.1 Relation to Other NQF-Endorsed Measures (NQF Measure evaluation criterion 5)

Are there related measures or competing measures?

☐yes
☒no

6.2 Harmonization (NQF Measure Submission Form, Related and Competing 5.04 and 5.04)

6.3 Competing Measures (NQF Measure Submission Form, Related and Competing 5.06)

Additional Information (NQF Measure Submission Form, Additional)

Appendix

Other Additional Information

Ad.1. Working Group/Expert Panel Involved in Measure Development

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2. First Year of Measure Release

Ad.3. Month and Year of Most Recent Revision

Ad.4. What is your frequency for review/update of this measure?

Ad.5. When is your next scheduled review/update for this measure?

Ad.6. Copyright Statement

Ad.7. Disclaimers

Ad.8. Additional Information/Comments

[Please delete this list of references and replace with your own references before submission]

REFERENCES

Agency for Healthcare Quality and Research. (n.d.). *Evidence-based practice center (EPC) program overview*. Retrieved December 8, 2020, from https://www.ahrq.gov/research/findings/evidence-based-reports/overview/index.html

Centers for Medicare & Medicaid Services. (n.d.-a). *CMS Measures Management System Blueprint*. Retrieved May 11, 2021, from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/MMS-Blueprint

Centers for Medicare & Medicaid Services. (n.d.-b). *Creating accessible products*. Retrieved January 11, 2021, from https://www.cms.gov/research-statistics-data-systems/section-508/public/creating-accessible-products

Grading of Recommendations Assessment, Development and Evaluation Working Group. (n.d.). *GRADE*. Retrieved July 20, 2020, from http://www.gradeworkinggroup.org/

Institute of Medicine. (2011). *Finding what works in health care: Standards for systematic reviews.* The National Academies Press. https://doi.org/10.17226/13059

National Quality Forum. (n.d.). *Submitting standards*. Retrieved July 20, 2020, from http://www.qualityforum.org/Measuring_Performance/Submitting_Standards.aspx

National Quality Forum. (2010). *Measurement framework: Evaluating efficiency across patient-focused episodes of care.* Retrieved July 20, 2020, from https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=25912

National Quality Forum. (2019). *Measure evaluation criteria and guidance for evaluating measures for endorsement*. Retrieved May 11, 2021, from http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=88439

National Quality Forum. (2021a). *NQF composite measure submission form*. Retrieved June 8, 2021, from https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=83206

National Quality Forum. (2021b). *NQF measure submission form*. Retrieved June 7, 2021, from https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=86103

U.S. Preventive Services Task Force. (n.d.). *Grade definitions*. Retrieved July 20, 2020, from https://www.uspreventiveservicestaskforce.org/Page/Name/grade-definitions