Measure Justification Form and Instructions

Project Title: ***Practitioner Level Measurement of Effective Access to Kidney Transplantation.***

Date:

Information included is current on *February 9, 2022*.

Project Overview:

The Centers for Medicare & Medicaid Services (CMS) has the University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) to develop practitioner-level measures in the area of access to kidney transplantation for dialysis patients. The contract name is Kidney Disease Quality Measure Development, Maintenance, and Support. The contract number is 75FCMC18D0041, task order number 75FCMC18F0001.

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

First Year Standardized Waitlist Ratio (FYSWR).

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 of placement and maintenance on the kidney or kidney-pancreas transplantation waitlist in active status (meaning the patient is ready to be transplanted immediately if a suitable organ becomes available), with the intended objective of improving the overall health of patients on dialysis.*

☐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)

This measure tracks the outcome of placement and maintenance on the kidney or kidney-pancreas transplantation waitlist in active status (meaning the patient is ready to be transplanted immediately if a suitable organ becomes available), with the intended objective of improving the overall health of patients on dialysis. Being waitlisted is an outcome as it represents a desirable change in health status for patients on dialysis, indicating achievement of a health condition conducive to kidney transplantation. This outcome results from specific activities directed by dialysis practitioners with the particular goal of achieving suitability for kidney transplantation by addressing the specific healthcare needs of patients on dialysis. These activities can include, but are not limited to, ensuring an ideal dialysis prescription and care, correction and optimization of common underlying chronic health conditions such as heart failure, coronary artery disease, diabetes mellitus and obesity, and as needed, optimizing mental health and social support systems. In addition, dialysis practitioners support the path for patients towards waitlisting or living donor transplantation through proper education about the transplantation option, referral to a transplant center and assistance with completion of the transplant evaluation process. The logic model for the steps involved is diagrammed below (with the outcome measure in bold):

Patients with ESRD are initiated on dialysis -> Patients not already on the wait list are assessed for eligibility for transplant referral by a dialysis practitioner -> Patients are referred to a transplant center for evaluation of candidacy for kidney or kidney-pancreas transplantation -> Dialysis practitioner assists patient with completion of the transplant evaluation process and in optimizing their health and functional status -> Patients deemed to be candidates for transplantation who have compatible living donors receive living donor transplant; otherwise they are placed on the waitlist -> ***Dialysis practitioner helps patient maintain active status on the wait list through involvement in ongoing evaluation activities and by optimizing health and functional status, with possibility to receive a deceased donor kidney transplant****.*

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

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

Two previous Technical Expert Panels (TEP) have been convened to discuss potential measures directed at improving access to kidney transplantation, in 2015 and most recently, in 2021 (2015 TEP Report: https://dialysisdata.org/sites/default/files/content/ESRD_Measures/Access_To_Kidney_Transplantation_TEP_Summary_Report.pdf; 2021 TEP Report: https://dialysisdata.org/content/esrd-measures, please see Practitioner Level Measurement of Effective Access to Kidney Transplantation under Ongoing Technical Expert Panels section). Both were comprised of relevant stakeholders, including dialysis

nephrologists, transplant nephrologists, transplant surgeons, social workers, researchers, and notably, patient representatives with a history of end-stage kidney disease. Discussions during both TEPs revealed broad support for the importance of waitlisting, and formal voting demonstrated a majority of TEP members were in favor of the development of quality measures targeting waitlisting (at the dialysis facility level for the 2015 TEP, and the practitioner level for the 2021 TEP).

In addition to the above, empirical support for the value of waitlisting to patients comes from a published study reporting on a large survey of 409 patients or family members who agreed to receiving emails from the National Kidney Foundation (Husain S.A. et al, Am. J. Transplant 2018;18(11):2781-2790). Participants include both patients with advanced chronic kidney disease prior to transplant, and recipients of transplants, and were asked about their priorities in choice of a transplant center. Notably, participants were most likely (a plurality of participants) to rank waitlisting characteristics (such as ease of getting on the waitlist) as the most important feature, in contrast to other transplant center characteristics such as post-transplant outcomes and practical considerations (e.g. distance to center).

National or large regional studies provide strong empirical support for the association between processes under dialysis practitioner control and subsequent waitlisting. In one large regional study conducted on facilities in the state of Georgia, a standardized dialysis facility referral ratio was developed, adjusted for age, demographics and comorbidities (Paul S. et al, Clin J Am Soc Nephrol 2018;13:282-289). There was substantial variability across dialysis facilities in referral rates, and a Spearman correlation performed between ranking on the referral ratio and dialysis facility waitlist rates was highly significant ($r=0.35$, $p<0.001$). A national study using registry data (United States Renal Data System) from 2005-2007 examined the association between whether patients were informed about kidney transplantation (based on reporting on the Medical Evidence Form 2728) and subsequent access to kidney transplantation (waitlisting or receipt of a live donor transplant) (Kucirka LM et al. Am J Transplant 2012;12:351-357). Approximately 30% of patients were uninformed about kidney transplantation, and this was associated with half the rate of access to transplantation compared to patients who were informed. In a related survey study of 388 hemodialysis patients, whether provision of information about transplantation by nephrologists or dialysis staff occurred was directly confirmed with patients (Salter ML et al, J Am Soc Nephrol 2014;25:2871-2877). Patient report of provision of such information was associated with a three-fold increase in likelihood of waitlisting. Finally, a large survey study of 170 dialysis facilities in the Heartland Kidney Network (Iowa, Kansas, Missouri and Nebraska) was conducted to examine transplant education practices (Waterman AD et al, Clin J Am Soc Nephrol 2015;10:1617-1625). Facilities employing multiple (>3) transplant education strategies (e.g. provision of brochures, referral to formal transplant education program, distribution of transplant center contact information) had 36% higher waitlist rates compared to facilities employing fewer strategies.

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)

N/A

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)

A measure focusing on waitlisting is appropriate for several reasons. First, in preparing patients for suitability for waitlisting, dialysis practitioners optimize their health and functional status, improving their overall health state. Second, waitlisting is a necessary step prior to potential receipt of a kidney transplant, which is known to be beneficial for survival and quality of life [1]. Third, dialysis practitioners exert substantial control over the processes that result in waitlisting. This includes proper education of dialysis patients on the option for transplant, referral of appropriate patients to a transplant center for evaluation, and assisting patients with completion of the transplant evaluation process, in order to increase their candidacy for transplant waitlisting. These types of activities are included as part of the conditions for coverage for Medicare certification of ESRD dialysis facilities. Finally, wide regional and facility variations in waitlisting rates highlight substantial room for improvement for this measure [2-5].

This measure focuses specifically on the prevalent dialysis population, examining waitlisting monthly in active status for each patient. As this measure assesses monthly waitlisting in active status of patients, it also evaluates and encourages maintenance of patients on the waitlist which is important given the long duration most patients have to wait to eventually access a deceased donor transplant (national median of roughly 4 years) [6]. In particular, maintenance of active status requires ongoing attention by dialysis practitioners to optimizing the health of patients, to ensure sustained suitability for transplant waitlisting. Maintenance of active status on the waitlist is additionally important given demonstrated disparities [7] and positive association with subsequent transplantation [8]. This is an important area to which dialysis practitioners can contribute through ensuring patients remain healthy, and complete any ongoing testing activities required to remain active on the wait list. In contrast to this measure, the First Year Standardized Waitlist Ratio focuses solely on new waitlistings and living donor kidney transplants to incentivize early action, rather than ongoing maintenance on the waitlist, as this measure does.

1. Tonelli M, Wiebe N, Knoll G, et al. Systematic review: kidney transplantation compared with dialysis in clinically relevant outcomes. American Journal of Transplantation 2011;11:2093-2109.

Abstract: Individual studies indicate that kidney transplantation is associated with lower mortality and improved quality of life compared with chronic dialysis treatment. We did a systematic review to

summarize the benefits of transplantation, aiming to identify characteristics associated with especially large or small relative benefit. Results were not pooled because of expected diversity inherent to observational studies. Risk of bias was assessed using the Downs and Black checklist and items related to time-to-event analysis techniques. MEDLINE and EMBASE were searched up to February 2010. Cohort studies comparing adult chronic dialysis patients with kidney transplantation recipients for clinical outcomes were selected. We identified 110 eligible studies with a total of 1 922 300 participants. Most studies found significantly lower mortality associated with transplantation, and the relative magnitude of the benefit seemed to increase over time ($p < 0.001$). Most studies also found that the risk of cardiovascular events was significantly reduced among transplant recipients. Quality of life was significantly and substantially better among transplant recipients. Despite increases in the age and comorbidity of contemporary transplant recipients, the relative benefits of transplantation seem to be increasing over time. These findings validate current attempts to increase the number of people worldwide that benefit from kidney transplantation.

2. Ashby VB, Kalbfleisch JD, Wolfe RA, et al. Geographic variability in access to primary kidney transplantation in the United States, 1996-2005. American Journal of Transplantation 2007; 7 (5 Part 2):1412-1423.

Abstract: This article focuses on geographic variability in patient access to kidney transplantation in the United States. It examines geographic differences and trends in access rates to kidney transplantation, in the component rates of wait-listing, and of living and deceased donor transplantation. Using data from Centers for Medicare and Medicaid Services and the Organ Procurement and Transplantation Network/Scientific Registry of Transplant Recipients, we studied 700,000+ patients under 75, who began chronic dialysis treatment, received their first living donor kidney transplant, or were placed on the waiting list pre-emptively. Relative rates of wait-listing and transplantation by State were calculated using Cox regression models, adjusted for patient demographics. There were geographic differences in access to the kidney waiting list and to a kidney transplant. Adjusted wait-list rates ranged from 37% lower to 64% higher than the national average. The living donor rate ranged from 57% lower to 166% higher, while the deceased donor transplant rate ranged from 60% lower to 150% higher than the national average. In general, States with higher wait-listing rates tended to have lower transplantation rates and States with lower wait-listing rates had higher transplant rates. Six States demonstrated both high wait-listing and deceased donor transplantation rates while six others, plus D.C. and Puerto Rico, were below the national average for both parameters.

3. Satayathum S, Pisoni RL, McCullough KP, et al. Kidney transplantation and wait-listing rates from the international Dialysis Outcomes and Practice Patterns Study (DOPPS). Kidney Intl 2005 Jul; 68 (1):330-337.

Abstract: BACKGROUND: The international Dialysis Outcomes and Practice Patterns Study (DOPPS I and II) allows description of variations in kidney transplantation and wait-listing from nationally representative samples of 18- to 65-year-old hemodialysis patients. The present study examines the health status and socioeconomic characteristics of United States patients, the role of for-profit versus not-for-profit status of dialysis facilities, and the likelihood of transplant wait-listing and transplantation rates. METHODS: Analyses of transplantation rates were based on 5267 randomly selected DOPPS I patients in dialysis units in the United States, Europe, and Japan who received chronic hemodialysis therapy for at least 90 days in 2000. Left-truncated Cox regression was used to assess time to kidney

transplantation. Logistic regression determined the odds of being transplant wait-listed for a cross-section of 1323 hemodialysis patients in the United States in 2000. Furthermore, kidney transplant wait-listing was determined in 12 countries from cross-sectional samples of DOPPS II hemodialysis patients in 2002 to 2003 (N= 4274). RESULTS: Transplantation rates varied widely, from very low in Japan to 25-fold higher in the United States and 75-fold higher in Spain (both P values <0.0001). Factors associated with higher rates of transplantation included younger age, nonblack race, less comorbidity, fewer years on dialysis, higher income, and higher education levels. The likelihood of being wait-listed showed wide variation internationally and by United States region but not by for-profit dialysis unit status within the United States. CONCLUSION: DOPPS I and II confirmed large variations in kidney transplantation rates by country, even after adjusting for differences in case mix. Facility size and, in the United States, profit status, were not associated with varying transplantation rates. International results consistently showed higher transplantation rates for younger, healthier, better-educated, and higher income patients.

4. Patzer RE, Plantinga L, Krisher J, Pastan SO. Dialysis facility and network factors associated with low kidney transplantation rates among United States dialysis facilities. Am J Transplant. 2014 Jul; 14(7):1562-72.

Abstract: Variability in transplant rates between different dialysis units has been noted, yet little is known about facility-level factors associated with low standardized transplant ratios (STRs) across the United States End-stage Renal Disease (ESRD) Network regions. We analyzed Centers for Medicare & Medicaid Services Dialysis Facility Report data from 2007 to 2010 to examine facility-level factors associated with low STRs using multivariable mixed models. Among 4098 dialysis facilities treating 305 698 patients, there was wide variability in facility-level STRs across the 18 ESRD Networks. Four-year average STRs ranged from 0.69 (95% confidence interval [CI]: 0.64-0.73) in Network 6 (Southeastern Kidney Council) to 1.61 (95% CI: 1.47-1.76) in Network 1 (New England). Factors significantly associated with a lower STR (p<0.0001) included for-profit status, facilities with higher percentage black patients, patients with no health insurance and patients with diabetes. A greater number of facility staff, more transplant centers per 10,000 ESRD patients and a higher percentage of patients who were employed or utilized peritoneal dialysis were associated with higher STRs. The lowest performing dialysis facilities were in the Southeastern United States. Understanding the modifiable facility-level factors associated with low transplant rates may inform interventions to improve access to transplantation.

5. Melanson TA, Gander JC, Rossi A, et al. Variation in Waitlisting Rates at the Dialysis Facility Level in the Context of Goals for Improving Kidney Health in the United States. Kidney International Reports 2021;6:1965-1968. No abstract.

6. United States Renal Data System. 2020 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, 2020.

7. Kulkarni S, Ladin K, Haakinson D, et al. Association of Racial Disparities With Access to Kidney Transplant After the Implementation of the New Kidney Allocation System. JAMA Surg 2019; 154(7):618-625.

8. Grams, M. E., Massie, A. B., Schold, J. D., Chen, B. P., & Segev, D. L. (2013). Trends in the inactive kidney transplant waitlist and implications for candidate survival. American Journal of Transplantation, 13(4), 1012-1018.

Abstract: In November 2003, OPTN policy was amended to allow kidney transplant candidates to accrue waiting time while registered as status 7, or inactive. We evaluated trends in inactive listings and the association of inactive status with transplantation and survival, studying 262,824 adult first-time KT candidates listed between 2000 and 2011. The proportion of waitlist candidates initially listed as inactive increased from 2.3% prepolicy change to 31.4% in 2011. Candidates initially listed as inactive were older, more often female, African American, and with higher body mass index. Postpolicy change, conversion from initially inactive to active status generally occurred early if at all: at 1 year after listing, 52.7% of initially inactive candidates had been activated; at 3 years, only 66.3% had been activated. Inactive status was associated with a substantially higher waitlist mortality (aHR 2.21, 95%CI:2.15-2.28, p<0.001) and lower rates of eventual transplantation (aRR 0.68, 95%CI:0.67-0.70, p<0.001). In summary, waitlist practice has changed significantly since November 2003, with a sharp increase in the number of inactive candidates. Using the full waitlist to estimate organ shortage or as a comparison group in transplant outcome studies is less appropriate in the current era.

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

After applying all exclusion criteria, we evaluated the aPPPW performance scores for all dialysis practitioner group practices that had at least 11 patients in 2019. The mean value of aPPPW was 12.3%. The interquartile range (Q3-Q1) is 7.3%, with the bottom quartile of practitioner group practices having 7.3% or less of prevalent patients waitlisted vs. the top quartile of practitioner group having 15.6% or more of their prevalent patients waitlisted.

N=dialysis practitioner groups=2276; N of patients=280,855; N of patient-months=2,541,229.

Table 1: Descriptive statistics of aPPPW (%), overall and by decile, 2019

	Mean	*Std Dev*	*Minimum*	*Maximum*	*Median*	*Lower Quartile*	*Upper Quartile*
Overall	*	*	*	*	*	*	*
*	*12.3*	*6.2*	*0.0*	*70.4*	*11.7*	*8.3*	*15.6*
Decile	*	*	*	*	*	*	*
1	*2.8*	*1.8*	*0.0*	*5.4*	*3.2*	*1.0*	*4.4*
2	*6.5*	*0.6*	*5.4*	*7.5*	*6.5*	*6.0*	*7.0*
3	*8.3*	*0.4*	*7.5*	*9.1*	*8.3*	*7.9*	*8.6*
4	*9.8*	*0.4*	*9.1*	*10.4*	*9.8*	*9.5*	*10.1*
5	*11.0*	*0.4*	*10.4*	*11.7*	*11.0*	*10.7*	*11.3*
6	*12.4*	*0.4*	*11.7*	*13.1*	*12.4*	*12.0*	*12.8*
7	*13.8*	*0.5*	*13.1*	*14.7*	*13.7*	*13.4*	*14.2*
8	*15.6*	*0.5*	*14.7*	*16.5*	*15.6*	*15.1*	*16.1*
9	*17.9*	*0.9*	*16.5*	*19.6*	*17.8*	*17.2*	*18.7*
10	*24.7*	*5.8*	*19.6*	*70.4*	*22.7*	*20.7*	*26.9*

**Cells intentionally left blank.*

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)

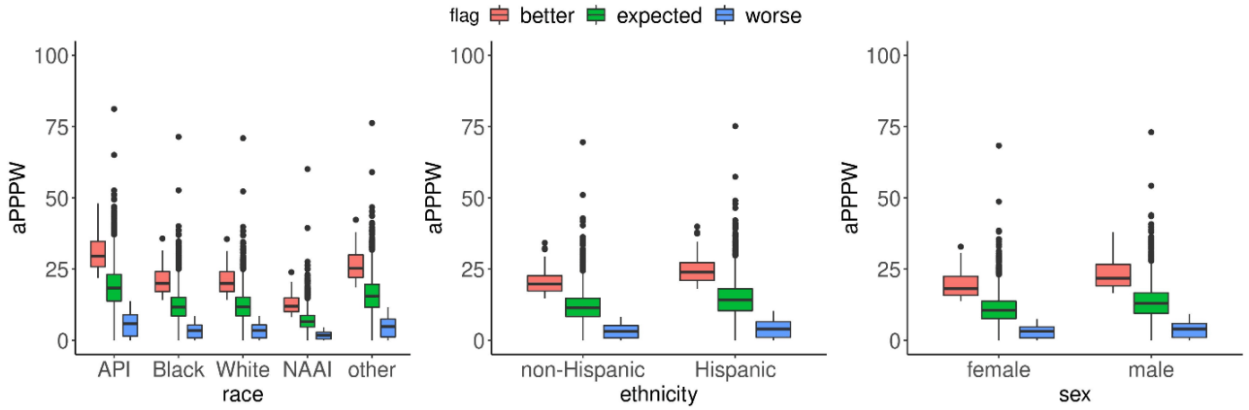
N= dialysis practitioner groups=2276; N of patients=280,855; N of patient-months=2,541,229.

Table 2: Descriptive statistics of aPPPW (%), by race, ethnicity and sex, 2019

	Mean	*Std Dev*	*Minimum*	*Maximum*	*Median*	*Lower Quartile*	*Upper Quartile*
Race	*	*	*	*	*	*	*
Asian/Pacific Islander	*18.4*	*8.4*	*0.0*	*81.2*	*18.0*	*13.1*	*23.3*
Black	*12.0*	*6.1*	*0.0*	*71.4*	*11.4*	*8.1*	*15.2*
White	*12.0*	*6.1*	*0.0*	*70.9*	*11.5*	*8.1*	*15.3*
Native American/Alaskan Indian	*6.9*	*4.0*	*0.0*	*60.1*	*6.4*	*4.4*	*8.8*
"Other" race	*15.7*	*7.4*	*0.0*	*76.3*	*15.2*	*11.0*	*19.9*
Ethnicity	*	*	*	*	*	*	*
Non-Hispanic	*11.8*	*6.0*	*0.0*	*69.5*	*11.3*	*8.0*	*14.9*
Hispanic	*14.5*	*7.1*	*0.0*	*75.1*	*13.9*	*10.0*	*18.3*
Sex	*	*	*	*	*	*	*
Female	*10.9*	*5.7*	*0.0*	*68.3*	*10.4*	*7.3*	*13.8*
Male	*13.3*	*6.6*	*0.0*	*73.0*	*12.7*	*9.0*	*16.8*

**Cells intentionally left blank.*

Figure 1: Performance of aPPPW (%), by race, ethnicity and sex, 2019



The data presented in Table 2 and Figure 1 above demonstrate wide variation and performance gaps within strata of race, ethnicity and sex categories.

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)

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

☐abstracted from paper record
☒administrative claims
☒clinical database/registry
☐abstracted from electronic health record (EHR)
☐electronic clinical quality measure (eCQM) Health Quality Measure Format (HQMF) implemented in EHRs
☐other (specify) Click or tap here to enter text.

Measure tested with data from

☐abstracted from paper record
☒administrative claims
☒clinical database/registry
☐abstracted from EHRs
☐eCQM (HQMF) implemented in EHRs
☐other (specify) Click or tap here to enter text.

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

2019 data derived from a combination of CROWNWeb, the Nursing Home Minimum Dataset, transplant registries (OPTN, SRTR), the CMS Medical Evidence Form (CMS Form 2728), Medicare claims from CMS, and the monthly capitation payment (MCP) from the Integrated Data Repository (IDR).

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

01-01-2019 – 12-31-2019

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

Provide testing for all levels specified and intended for measure implementation (e.g., individual clinician, hospital, health plan).

Measure specified to measure performance of (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)

Using 2019 data, there were 2,276 dialysis practitioner groups included in these analyses, after restricting to dialysis practitioner group practices that had at least 11 eligible patients.

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 are 2,541,229 patient-months (280,855 patients) in total. Among all patients-months in 2019, the average age was 57.4 years old, 41.6% of patient-months were female, 54.8% were White, 37.9% were Black, 5.2% were Asian/Pacific Islander, 1.6% were American Indian/Alaskan Native, 0.4% were Other/Multi-racial/Unknown/missing and 18.0% were Hispanic.

At the patient-level, the mean age was 57.2 years old and 41.5% were female. Of these 56.2% were White, 36.5% were Black, 5.2% were Asian/Pacific Islander, 1.6% were American Indian/Alaskan Native, and 0.4% were other/Multi-racial/Unknown/missing and 17.6% were Hispanic.

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:

- *Sex (we acknowledge that sex is less recognized as a social risk factor but it is being increasingly considered as such especially given its relationship to gender [see for example, O'Neil et al. Gender/Sex as a social determinant of cardiovascular risk. Circulation 2018;137:854], and have therefore chosen to include an assessment of it in our analysis)*
- *Race*
- *Ethnicity*

- *Medicare-Medicaid dual eligibility*

Data on patient level factors obtained from Medicare claims and administrative data.

Zipcode level – Area Deprivation Index from 2015 Census data.

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

☐critical data elements used in the measure (e.g., inter-abstractor reliability; data element reliability must address all critical data elements)
☒performance measure score (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 2019 data to calculate dialysis practitioner group practice annual performance scores. Our approach for determining measure reliability aligns with one-way analysis of variance (ANOVA), in which the between dialysis practitioner group practice variation (σ_b^2) and the within- dialysis practitioner group practice variation ($\sigma_{t,w}^2$) in the measure is determined. The inter-unit reliability (IUR) measures the proportion of the total variation of the measure (i.e., $\sigma_b^2 + \sigma_{t,w}^2$) that is attributed to the between – dialysis practitioner group practice variation, the true signal reflects the differences across dialysis practitioner group practices. We assessed reliability by calculating inter-unit reliability (IUR) for the annual performance scores. If the measure were an average of the individuals' measurements under the care of one dialysis practitioner group practice, the usual ANOVA approach would be used. The yearly based measure, however, is not a simple average and we instead estimate the IUR using a bootstrap approach, which uses a resampling scheme to estimate the within dialysis practitioner group practice variation that cannot be directly estimated by ANOVA. A small IUR (near 0) reveals that most of the variation of the measures between dialysis practitioner group practices is driven by random noise, indicating the measure would not be a good characterization of the differences among dialysis practitioner group practices. A large IUR (near 1) indicates that most of the variation between dialysis practitioner groups practices is due to true differences between dialysis practitioner group practices.

Below is our approach to calculate IUR.

Let $T_1, \ldots, T_N$ be the Percentage of Prevalent Patients Waitlisted in Actives (aPPPW) for N practitioner groups. Within each practitioner group, select at random and with replacement B = 100 bootstrap samples. That is, if the i^{th} practitioner group has n_i subjects, randomly draw with replacement n_i subjects from those in the same practitioner group, find their corresponding aPPPW and repeat the process 100 times. Thus, for the ith practitioner group, we have bootstrapped aPPPWs of $T_{i1}^*, \ldots, T_{i100}^*$. Let S_i^* be the sample variance of this bootstrap sample. From this it can be seen that

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

$is\ a\ bootstrap\ estimate\ of\ the\ within-facility\ variance\ in\ the\ aPPPW, namely\ \sigma^2_{t,w}$. Calling on formulas from the one-way analysis of variance, an estimate of the overall variance in aPPPW can be estimated by

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

$where\ ni\ is\ the\ number\ of\ subjects\ in\ the\ i^{th}$ practitioner group, $\bar{T} = \sum n_i T_i / \sum n_i$, and

$$n' = \frac{1}{N-1} (\sum n_i - \sum n_i^2 \ / \sum n_i)$$

is approximately the average dialysis practitioner group practice size (number of patients per dialysis practitioner group practice).Note that s_t^2 is an estimate of $\sigma_b^2 + \sigma_{t,w}^2$ where σ_b^2 is the between-group variance, the true signal reflecting the differences across practitioner groups. Thus, the IUR, which is defined by IUR =$\sigma_b^2/(\sigma_b^2 + \sigma_{t,w}^2)$ can be estimated by $(s_t^2 - s_{t,w}^2)/s_t^2$.

The reliability of aPPPW calculation only included dialysis practitioner group practices with at least 11 patients during the entire year.

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

The IUR is 0.93. Dialysis practitioner group practices with <11 eligible patients were excluded from this calculation.

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

The value of IUR indicates that about 93% of the variation in the aPPPW measure can be attributed to the between-dialysis practitioner group practice differences (signal) and about 7% of variation to within-dialysis practitioner group practice variation (noise). The value of IUR implies 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)

☐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)

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

Validity of the measure was tested by evaluating the association between the dialysis practitioner group level measure performance, and mortality and overall transplant rates among all patients attributed to the practitioner groups. We hypothesized that practitioner groups with higher performance on the aPPPW measure would have higher transplant rates among their patients. This would be expected to follow from activities these practitioner groups conducted to improve the health and therefore suitability of their patients for transplant candidacy. Along similar lines, we hypothesized that practitioner groups with higher performance on the aPPPW measure would demonstrate lower mortality among their patients. However, we expected this to be a more modest association given the many other factors that can affect mortality within the dialysis population.

To evaluate the associations, we first divided dialysis practitioner groups, into 3 tertiles (T1 to T3) based on their performance on the aPPPW (T1 to T3, from highest to lowest waitlisting). Tertiles were chosen in order to evaluate a gradient in effect, but still maintain sufficient numbers within each group for statistical precision. We also computed the corresponding mortality rate and transplant rate for each practitioner group in the same year. We then tested the trend between the tertile grouping and these practitioner group-level outcomes. Finally, we examined the Spearman correlations between the practitioner group measure value and each of the outcomes respectively.

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

The tertile groups based on the performance scores were defined as:

T1 (best performance): 14.1% - 70.4%

T2: 9.6% - 14.1%

T3 (worst performance): 0% - 9.6%

The dialysis practitioner group level average mortality rates are 17.8, 18.3 and 19.2 deaths per 100 patient-years for T1, T2 and T3 respectively (trend test p=0.002). The Spearman correlation coefficient is -0.083 (p<0.0001). Average transplant rates are 5.0, 4.2 and 3.1 transplants per 100 patient-years for T1, T2 and T3 respectively (trend test p=0.002). The Spearman correlation coefficient is 0.279 (p<0.0001).

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

As expected, higher aPPPW performance correlated with higher transplant rate, with clear separation of transplant rates across dialysis practitioner group tertiles of performance. The direction of the relationship with mortality was also as expected, and statistically significant, with numerically lower mortality with higher performance on the measure although the magnitude of the association was smaller than for transplant rate.

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)

In order to evaluate the exclusion criteria, the differences in the number of patients with and without excluding age >= 75, nursing home patients, hospice patients, and dementia, were compared. We show the frequency of patients excluded due to each criteria. Additionally, we compared the performance

scores before and after exclusions. We do not exclude patients from dialysis practitioner groups with fewer than 11 attributed events.

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

Table 5: Overall number and percentage of patient months excluded

*	Before age, nursing home, hospice, and dementia exclusion	After age, nursing home, hospice, and dementia exclusion	Percentage excluded
Number of patient-months	3,561,019	2,541,229	28.6%

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

Table 6: Frequency distribution of patient-months excluded based on each exclusion criteria

Variable excluded	Frequency (%)
Age >= 75	766,648 (21.5)
Nursing home from CMS-2728	26,618 (0.8)
Nursing home from Nursing home history file	302,227 (8.5)
Hospice	14,581 (0.4)
Dementia	152,951 (4.3)

Table 7: Distribution of performance scores (aPPPW) before and after exclusions

aPPPW (%)	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*	*Q1*	*Median*	*Q3*
Before exclusion	*9.0*	*4.6*	*0.0*	*53.2*	*6.1*	*8.5*	*11.3*
After exclusion	*12.3*	*6.2*	*0.0*	*70.4*	*8.3*	*11.7*	*15.6*

Figure 2: Distribution of aPPPW (%) before exclusions

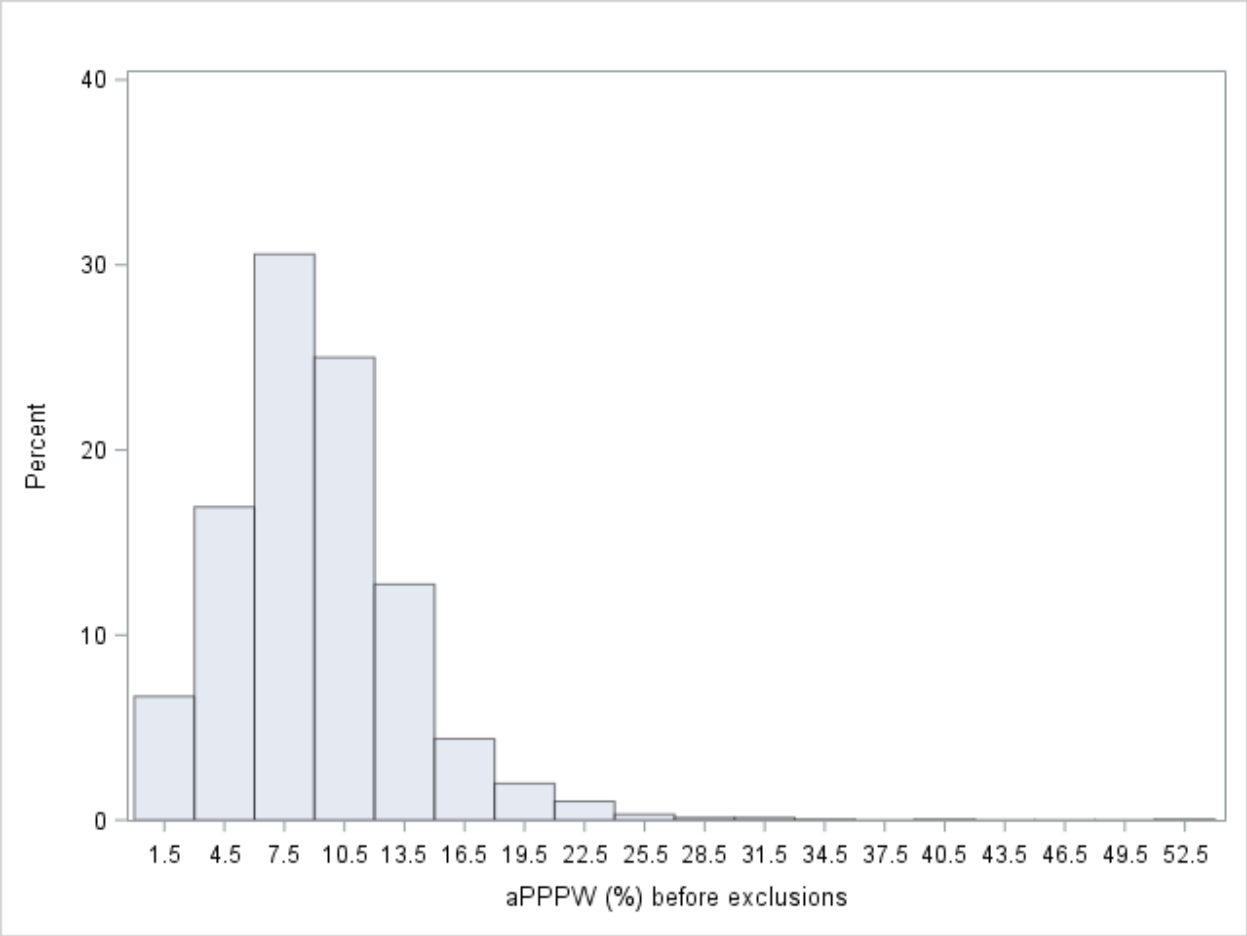


Figure 3: Distribution of aPPPW (%) after exclusions

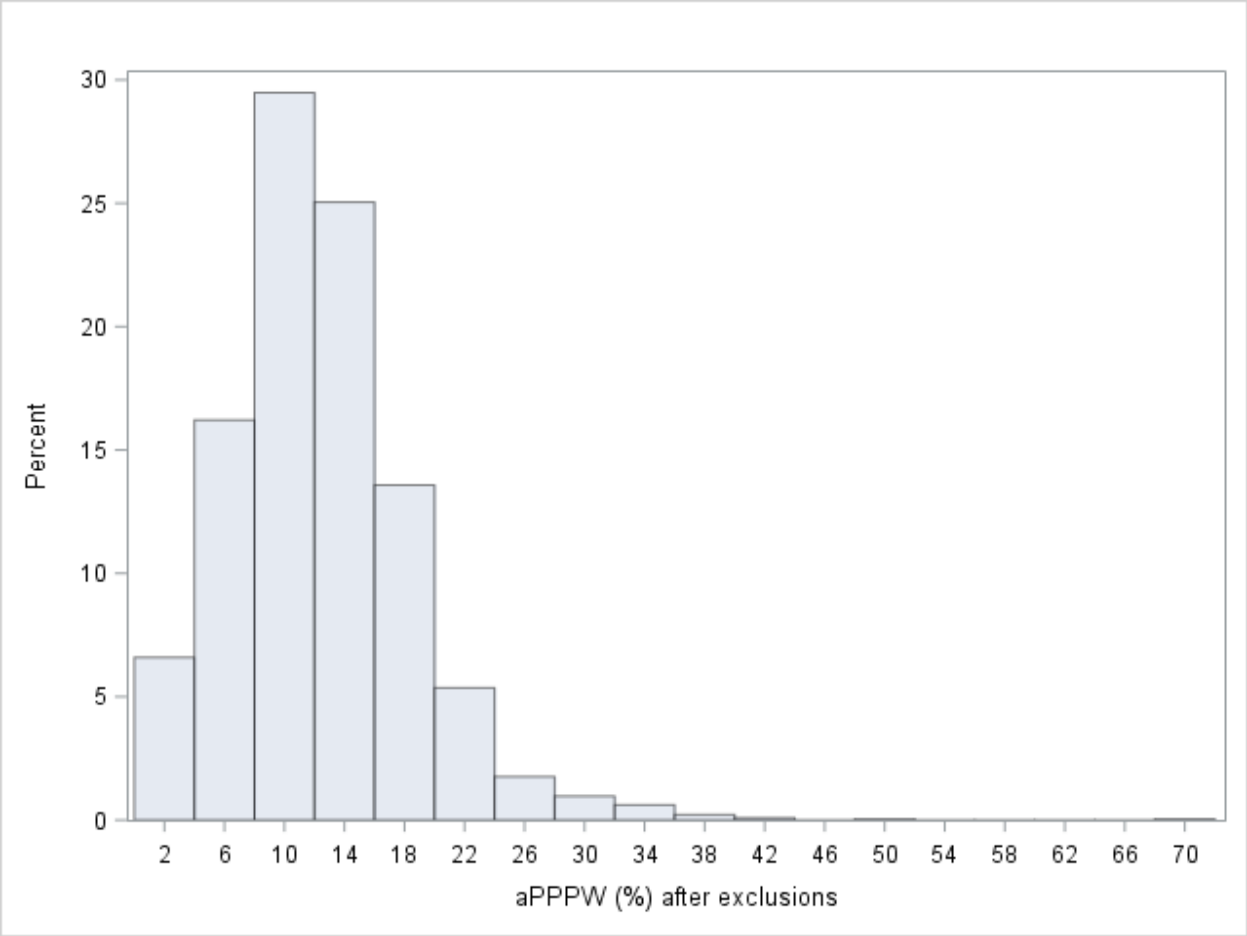
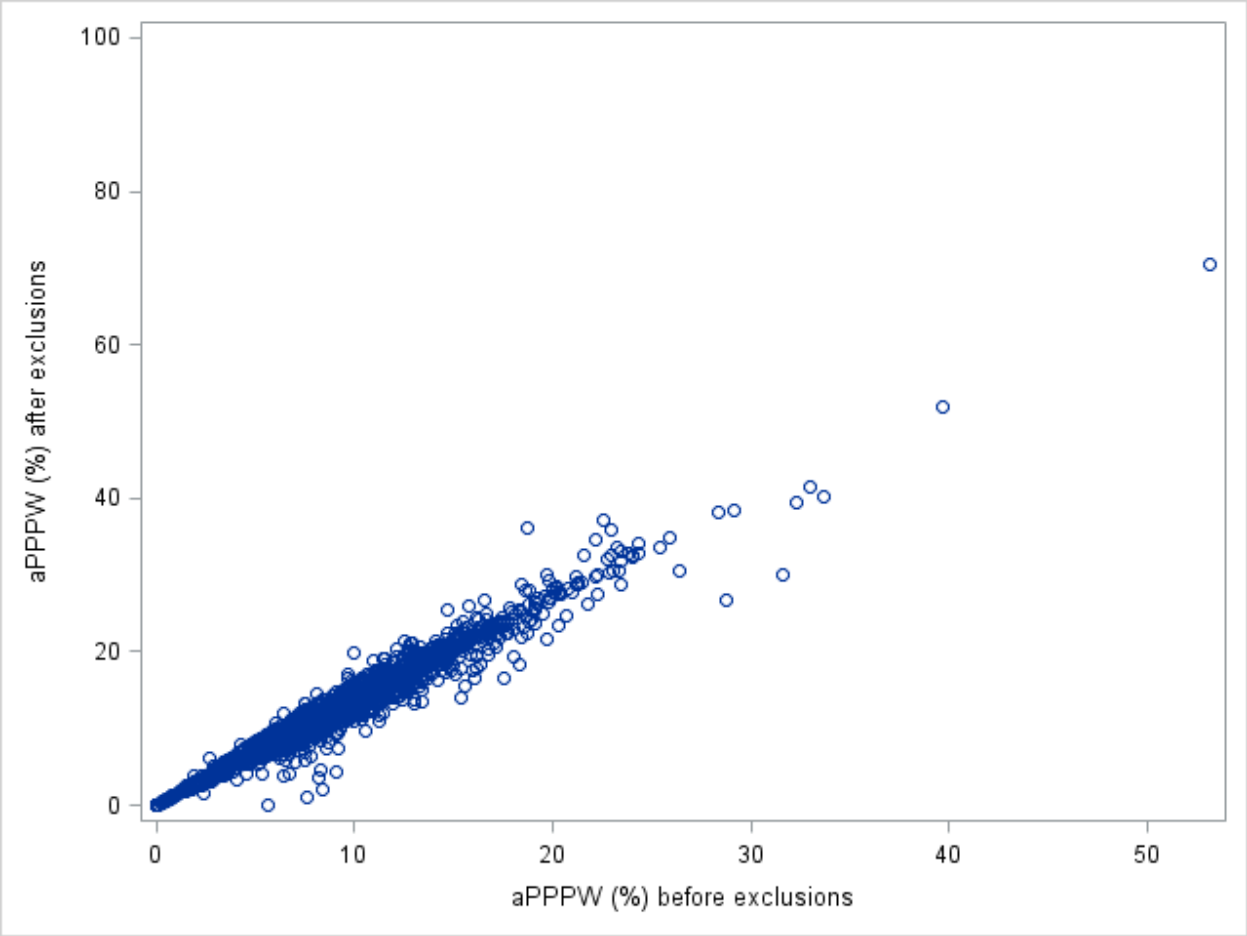


Figure 4: Scatterplot of aPPPW (%) with and without exclusions



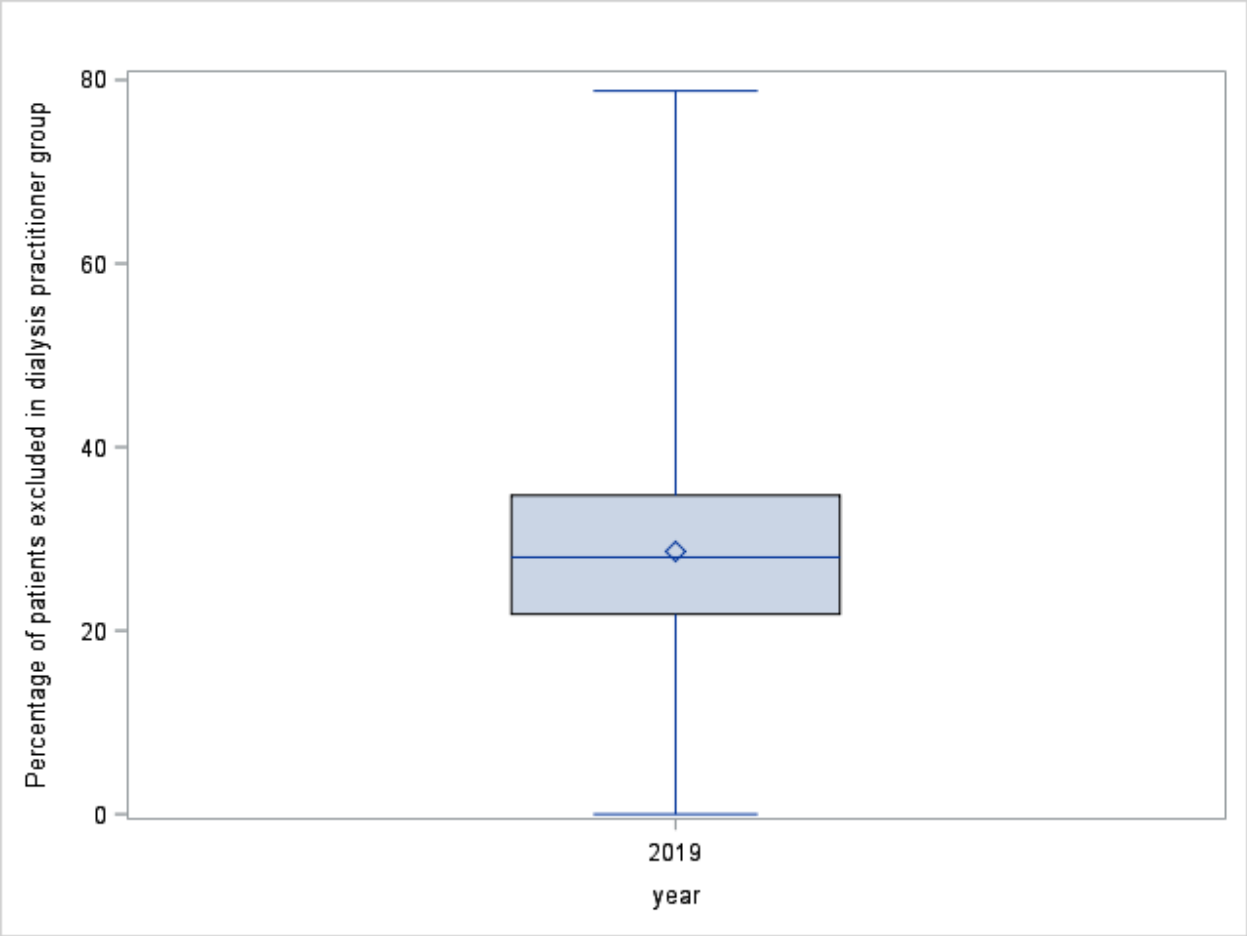
The correlation coefficient is 0.977 (p<0.001).

Table 8: Comparison of performance scores with and without excluded patients

*	*	*aPPPW without patient-level exclusion*	*aPPPW without patient-level exclusion*	*aPPPW without patient-level exclusion*	*aPPPW without patient-level exclusion*
*	*	*Better than Expected*	*As Expected*	*Worse than Expected*	*Total*
PPPW with patient-level exclusion	*Better than Expected*	*54*	*5*	*0*	*59 (2.6)*
PPPW with patient-level exclusion	*As Expected*	*5*	*2069*	*30*	*2104 (92.4)*
PPPW with patient-level exclusion	*Worse than Expected*	*0*	*12*	*101*	*113 (5.0)*
PPPW with patient-level exclusion	*Total*	*59 (2.6)*	*2086 (91.7)*	*131 (5.8)*	*2276*

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

Figure 5: Distribution of Excluded patients at dialysis practitioner group practice



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

Although overall measure scores are changed moderately by the exclusions (see Table 7, figure 2-3), practitioner group performance rankings are minimally affected (Table 8). Nevertheless, the exclusions are deemed important on clinical grounds as they represent a group of patients highly unlikely to be suitable for transplant waitlisting. Furthermore, there is a fair degree of variation in the percentage of patients excluded across dialysis practitioner groups, as shown in Figure 5. Finally, as the data to determine the exclusions is readily available, there is minimal additional burden for analysis anticipated by using these exclusion criteria.

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)

Covariates in the model are listed below:

- ***Age***
    - ***Age is included as continuous variable as well as age spline knots at 15, 55, and 70***
- ***ADI***
- ***Dual eligibility***
    - ***Dual Eligible***
    - ***Not Dual Eligible***
- ***Diabetes, primary cause of ESRD***
- ***Comorbidities at ESRD incidence:***
    - ***Congestive heart failure***
    - ***Atherosclerotic heart disease and other cardiac disease***
    - ***Cerebrovascular disease, CVA, TIA***
    - ***Peripheral vascular disease***
    - ***Diabetes other than as primary cause of ESRD (all types including diabetic retinopathy)***
    - ***Chronic obstructive pulmonary disease***
    - ***Inability to ambulate***
    - ***Inability to transfer***
    - ***Malignant neoplasm, cancer***
    - ***Tobacco use (current smoker)***
    - ***Drug dependence***
    - ***No Medical Evidence (CMS-2728) Form***
    - ***At least one of the comorbidities listed***

- ***A set of prevalent comorbidities based on either Medicare inpatient or outpatient claims (individual comorbidities categorized into 64 categories – see below)***
- ***Transplant center fixed characteristics and random effect***

To estimate the probability that a prevalent patient is waitlisted in active status, we use a mixed-effects logistic regression model, in which dialysis practitioner groups are modeled as fixed effects and transplant centers are modeled as random effects. The expected number of prevalent patients waitlisted in active status for the dialysis practitioner group under evaluation is estimated as the sum of the probabilities of prevalent patients waitlisted across all dialysis practitioner groups and assuming their effects are the same as the dialysis practitioner group under evaluation.

Consider patient k at dialysis practitioner group practice i and transplant center j during calendar month l; we set the response variate to Y_{ijkl} =1 if the patient is on the wait list in active status and Y_{ijkl} =0 if not. The model and methods are described in some additional detail below:

- ***To estimate the probability that a prevalent patient is waitlisted, we use a mixed-effects logistic regression model:***

$$log(\frac{p_{ijkl}}{1-p_{ijkl}}) = \gamma_i + \alpha_j + \beta^T Z_{ijkl}, \text{ (1)}$$

where p_{ijkl} represents the probability that patient k at dialysis practitioner group practice i and transplant center j during calendar month l is waitlisted in active status, and Z_{ijkl} represents the set of patient-level characteristics, including age (coded as a linear spline with empirically determined knots at ages 15, 55 and 70), incident comorbidities, prevalent comorbidities, ADI, and dual eligibility and i and the dialysis practitioner group practice indicators. In this mixed-effect model, γ_i is the fixed effect for dialysis practitioner groups and α_j is the random effect for transplant center j . It is assumed that the α_js arise as independent normal variables (i.e., $\alpha_j \sim N(0, \sigma^2)$).

- ***We then compute aPPPW$_m$ for each dialysis practitioner group practice m as follows: PPPW$_m$ = $\sum_i\sum_j\sum_k\sum_l$ exp($\gamma_m + \alpha_j + \beta^T Z_{ijkl}$) / {1 + exp($\gamma_m + \alpha_j + \beta^T Z_{ijkl}$)} / n, where n = total number of patient-months included in the overall study sample.***

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)

Table 9: Model statistics for risk factors in aPPPW model

Covariate	*odds ratio*	*95% CI*
Age	*	*
Continuous (years)	*1.033*	*1.021, 1.046*
Spline at 15	*0.950*	*0.939, 0.962*
Spline at 55	*0.976*	*0.975, 0.978*
Spline at 70	*0.796*	*0.789, 0.803*
Area Deprivation Index (ADI), per 10% increase on percentile scale	*0.916*	*0.991, 0.992*
Dual eligibility	*0.618*	*0.613, 0.624*
Diabetes, primary cause of ESRD	*0.643*	*0.632, 0.654*
Comorbidities at incidence		
Heart disease	*0.902*	*0.885, 0.920*
Other cardiac disease	*0.902*	*0.887, 0.916*
Congestive heart failure	*0.624*	*0.615, 0.632*
Chronic obstruction pulmonary disease	*0.608*	*0.588, 0.628*
Inability to ambulate	*0.362*	*0.342, 0.384*
Inability to transfer	*0.730*	*0.667, 0.798*
Cancer	*0.732*	*0.711, 0.753*
Peripheral vascular disease	*0.794*	*0.776, 0.812*
Cerebrovascular disease	*0.748*	*0.731, 0.765*
Tobacco use	*0.494*	*0.482, 0.505*
Drug use	*0.420*	*0.397, 0.444*
Diabetes, non-primary	*0.764*	*0.749, 0.779*
At least one incident comorbidity listed	*0.983*	*0.967, 1.000*
No Medical Evidence (CMS-2728 Form)	*0.510*	*0.490, 0.531*

Covariate	*odds ratio*	*95% CI*
At least 6 months of Medicare Coverage in prior year	*0.873*	*0.863, 0.884*
Prevalent comorbidities	*	*
Candidal esophagitis	*0.656*	*0.587, 0.733*
Sarcoidosis	*1.235*	*1.161, 1.314*
Cancer of Liver	*0.555*	*0.486, 0.635*
Cancer of Lung	*0.232*	*0.199, 0.270*
Cancer of Bladder	*0.569*	*0.501, 0.647*
Cancer of Bone	*0.175*	*0.139, 0.222*
Other Neoplasm	*0.867*	*0.805, 0.934*
Non-Hodgkins Lymphoma	*0.525*	*0.462, 0.596*
Multiple Myeloma	*0.299*	*0.274, 0.327*
Myelodysplastic Syndrome	*0.653*	*0.585, 0.727*
Diabetes without complications	*1.146*	*1.131, 1.161*
Diabetes with complications	*1.133*	*1.116, 1.149*
Glucocorticoid deficiency	*0.999*	*0.939, 1.064*
Malnutrition / Cachexia	*0.933*	*0.916, 0.951*
Disorders of urea cycle metabolism	*1.057*	*0.934, 1.195*
Other amyloidosis	*1.069*	*0.961, 1.188*
Other specified disorders of metabolism	*0.792*	*0.756, 0.831*
Sickle-cell Anemia	*0.851*	*0.774, 0.936*
Pancytopenia	*0.837*	*0.806, 0.869*
Neutropenia	*0.920*	*0.856, 0.988*
Substance Related Disorders	*0.525*	*0.472, 0.583*
Opioid Dependance	*0.659*	*0.628, 0.692*

Covariate	*odds ratio*	*95% CI*
Schizophrenia	*0.303*	*0.272, 0.338*
Peripheral autonomic neuropathy in disorders classified elsewhere	*0.903*	*0.824, 0.991*
Epilepsy	*0.728*	*0.710, 0.746*
Bipolar Disorder	*0.745*	*0.710, 0.781*
Major depressive affective disorder	*0.711*	*0.699, 0.723*
Alcohol Related Disorders	*0.889*	*0.835, 0.947*
Coma	*0.899*	*0.825, 0.981*
Cerebral edema	*1.158*	*1.029, 1.303*
Myocardial Infarction	*0.651*	*0.633, 0.670*
Coronary Atherosclerosis	*0.874*	*0.857, 0.892*
Pulmonary embolism and infarction	*0.881*	*0.839, 0.926*
Primary pulmonary hypertension	*0.816*	*0.762, 0.873*
Pulmonary Heart Disease	*0.793*	*0.776, 0.810*
Cardiomyopathy	*0.780*	*0.766, 0.796*
Atrioventricular block, complete	*0.712*	*0.662, 0.765*
Paroxysmal Tachycardia	*0.836*	*0.802, 0.871*
Atrial fibrillation	*0.816*	*0.802, 0.831*
Atrial flutter	*0.893*	*0.860, 0.928*
Acute Cerebrovascular Disease	*0.788*	*0.766, 0.811*
Peripheral and Visceral Atherosclerosis	*0.876*	*0.863, 0.890*
Venous Thromboembolism	*0.769*	*0.744, 0.795*
Esophageal varices	*2.239*	*2.042, 2.454*
Chronic Obstructive Pulmonary Disease	*0.607*	*0.594, 0.619*
Aspiration Pneumonitis	*0.976*	*0.921, 1.035*

Covariate	*odds ratio*	*95% CI*
Other Lower Respiratory Diseases	*1.062*	*0.995, 1.134*
Respiratory Failure	*0.648*	*0.635, 0.661*
Cirrhosis of Liver	*0.820*	*0.792, 0.848*
Other Liver Disease	*1.195*	*1.134, 1.259*
Pancreatitis	*0.724*	*0.683, 0.768*
Chronic Skin Ulcer	*0.604*	*0.592, 0.617*
Systemic lupus erythematosus and connective tissue disorders	*1.184*	*1.153, 1.215*
Rheumatoid Arthritis	*1.010*	*0.969, 1.053*
Pathologic Fracture	*0.798*	*0.722, 0.883*
Gangrene	*0.657*	*0.630, 0.685*
HIV	*0.522*	*0.497, 0.548*
Gastrostomy status	*0.920*	*0.842, 1.005*
Other artificial opening of urinary tract status	*0.536*	*0.468, 0.615*
Dependence on respirator, status	*1.024*	*0.944, 1.110*
Below knee amputation status	*0.544*	*0.523, 0.565*
Above knee amputation status	*0.483*	*0.444, 0.526*
Long-term (current) use of insulin	*1.056*	*1.042, 1.071*
Inflammatory polyarthropathy	*0.950*	*0.849, 1.062*
Weighted transplant center waitlist mortality ratio	*1.180*	*1.079, 1.291*
Weighted transplant center rate ratio	*0.705*	*0.676, 0.735*

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)

As noted in section 2b.23, we included Medicare-Medicaid dual eligibility and ADI as social risk factors in the model on a clinical and conceptual basis, and as supported by an expert panel. Both factors were

significantly associated with the outcome of waitlisting (see Table 9 in 2b.24).

We additionally examined selected variables, including sex, race and ethnicity, fitting models including covariates from the original model and adding each selected variable one at a time.

Table 10: *Odds Ratio and 95% Confidence Interval in the aPPPW model including race*

Race	***Odds Ratio***	***95%CI***
Asian/Pacific Islander	*1.190*	*1.119, 1.265*
Black	*0.885*	*0.876, 0.894*
White	*Reference*	*Reference*
Native American/Alaskan Indian	*0.602*	*0.558, 0.651*
"Other" race	*1.133*	*1.067, 1.202*

Table 11: *Odds Ratio and 95% Confidence Interval in the aPPPW model including ethnicity*

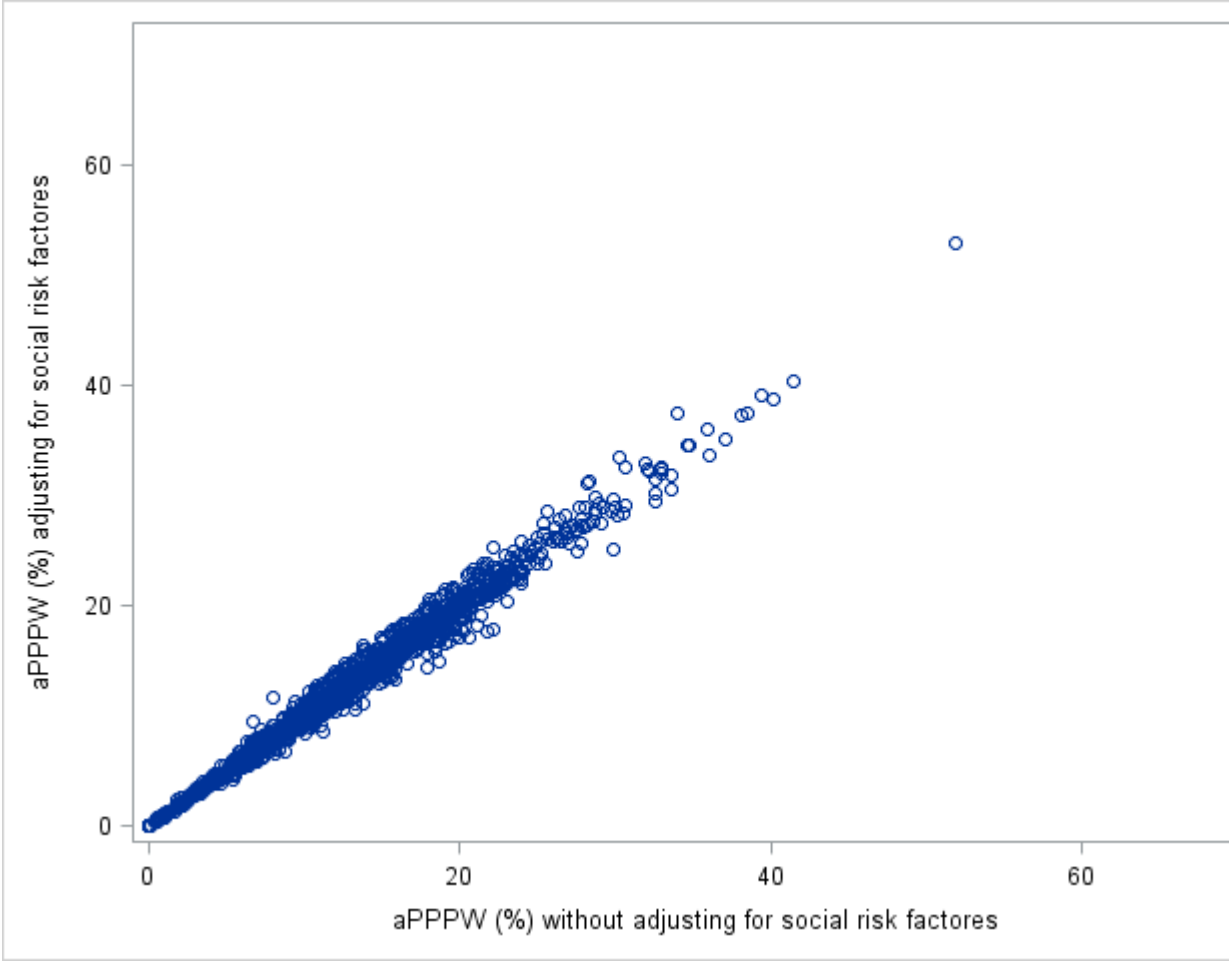
Ethnicity	***Odds Ratio***	***95%CI***
Non-Hispanic	*1.162*	*1.148, 1.176*
Hispanic	*Reference*	*Reference*

Table 12: *Odds Ratio and 95% Confidence Interval in the aPPPW model including sex*

Sex	***Odds Ratio***	***95%CI***
Female	*0.836*	*0.828, 0.843*
Male	*Reference*	*Reference*

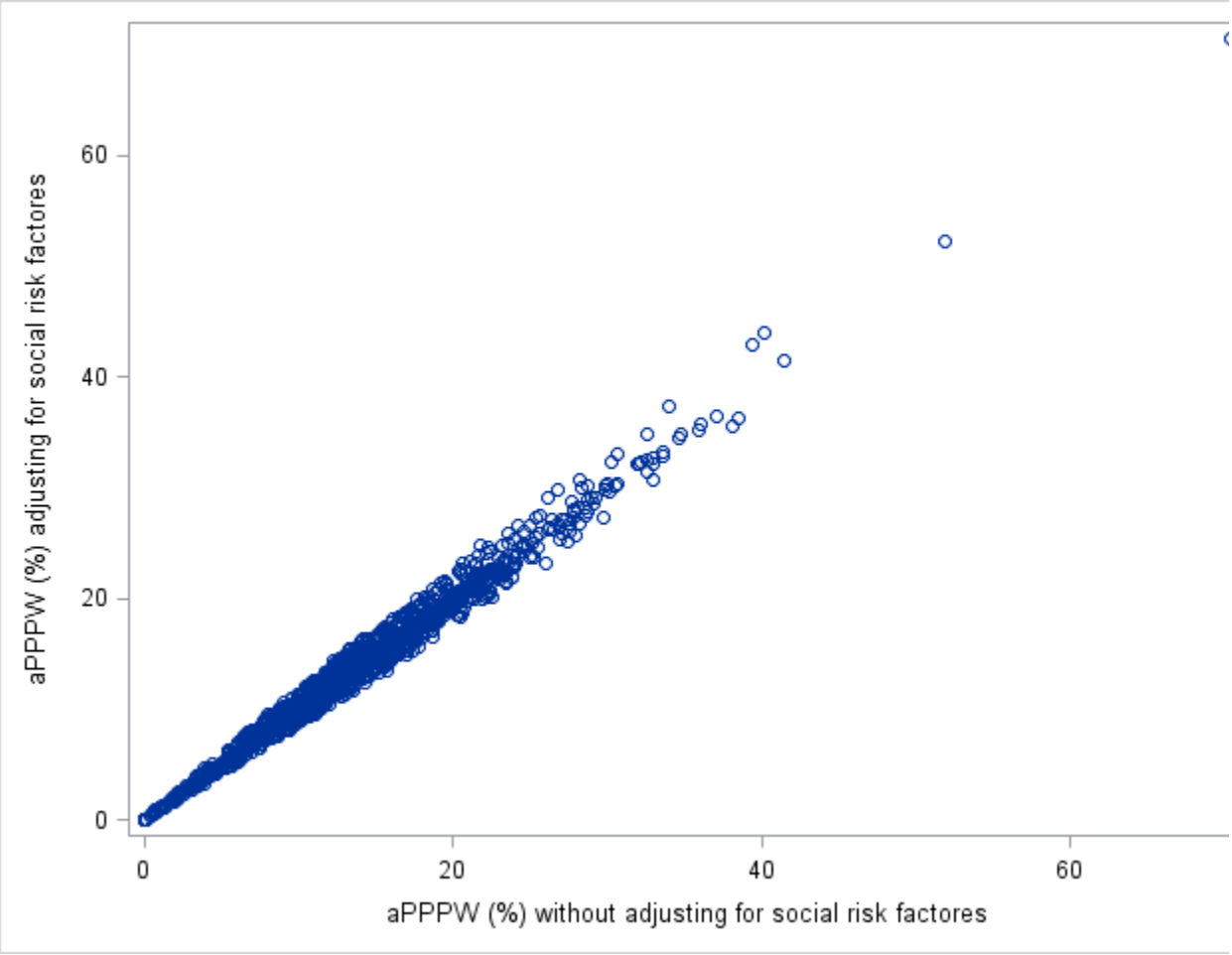
Figure 6: *Correlation between aPPPW with and without risk factors*

Race



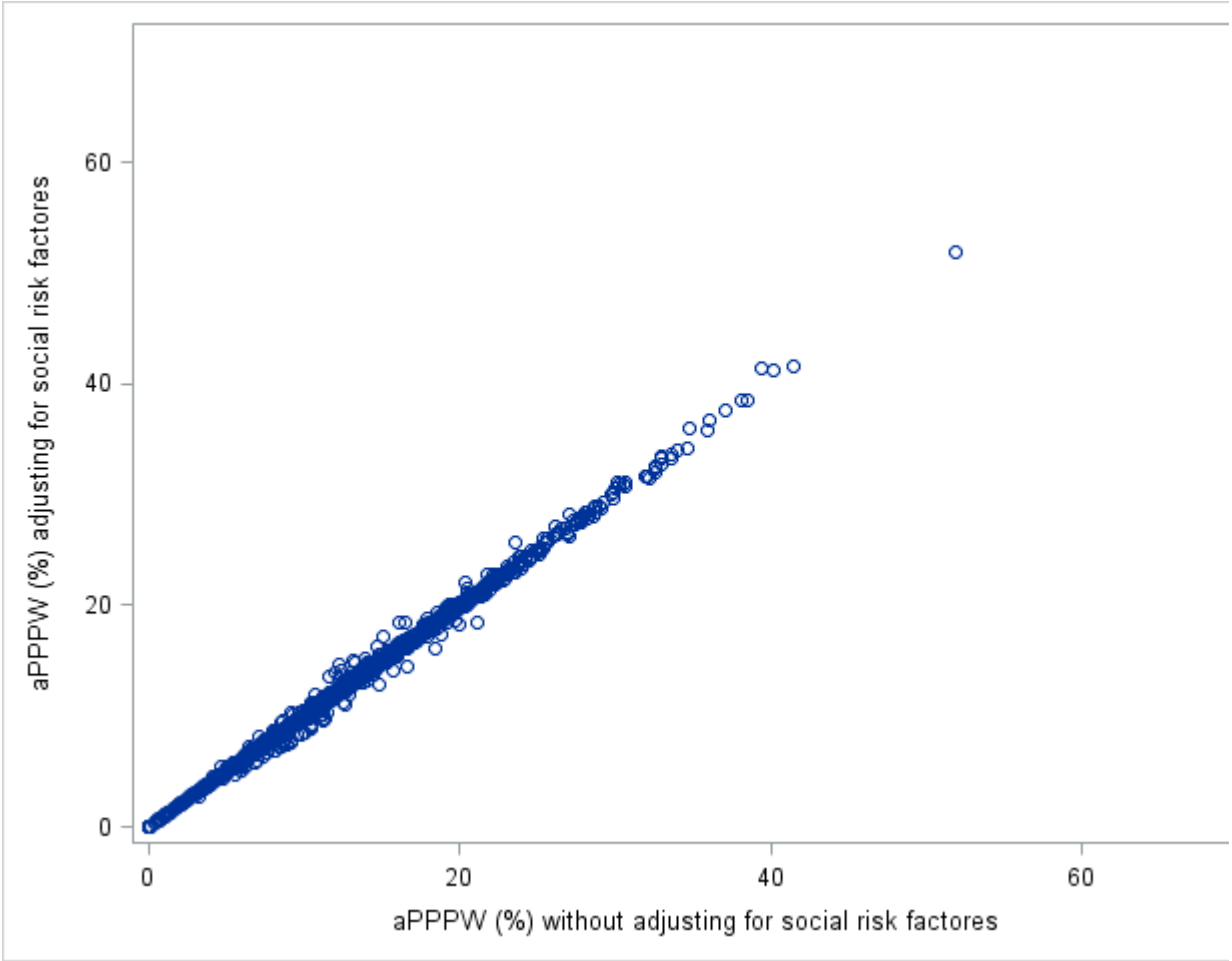
Correlation coefficient=0.992, p<0.0001

Ethnicity



Correlation coefficient=0.993, p<0.0001

Sex



Correlation coefficient=0.998, p<0.0001

Table 13: *Comparison of performances with and without adjusting for risk factors*

Race

*	*	***aPPPW without race***	***aPPPW without race***	***aPPPW without race***	***aPPPW without race***
*	*	***Better than Expected***	***As Expected***	***Worse than Expected***	***Total***
aPPPW with ethnicity	***Better than Expected***	*53*	*4*	*0*	*57 (2.5)*
aPPPW with ethnicity	***As Expected***	*6*	*2091*	*2*	*2099 (92.2)*
aPPPW with ethnicity	***Worse than Expected***	*0*	*9*	*111*	*120 (5.3)*
aPPPW with ethnicity	***Total***	*59 (2.6)*	*2104 (92.4)*	*113 (5.0)*	*2276*

**This cell is intentionally left blank.*

Ethnicity

*	*	***aPPPW without ethnicity***	***aPPPW without ethnicity***	***aPPPW without ethnicity***	***aPPPW without ethnicity***
*	*	***Better than Expected***	***As Expected***	***Worse than Expected***	***Total***
aPPPW with ethnicity	***Better than Expected***	*58*	*1*	*0*	*59 (2.6)*
aPPPW with ethnicity	***As Expected***	*1*	*2100*	*3*	*2104 (92.4)*
aPPPW with ethnicity	***Worse than Expected***	*0*	*3*	*110*	*113 (5.0)*
aPPPW with ethnicity	***Total***	*59 (2.6)*	*2104 (92.4)*	*113 (5.0)*	*2276*

**This cell is intentionally left blank.*

Sex

*	*	***PPPW without sex***	***PPPW without sex***	***PPPW without sex***	***PPPW without sex***
*	*	***Better than Expected***	***As Expected***	***Worse than Expected***	***Total***
aPPPW with sex	***Better than Expected***	*56*	*0*	*0*	*56 (2.5)*
aPPPW with sex	***As Expected***	*3*	*2100*	*6*	*2109 (92.7)*
aPPPW with sex	***Worse than Expected***	*0*	*4*	*107*	*111 (4.9)*
aPPPW with sex	***Total***	*59 (2.6)*	*2104 (92.4)*	*113 (5.0)*	*2276*

**This cell is intentionally left blank.*

Although there are differences in waitlisting by sex, ethnicity and race, it is unclear whether these associations are due to underlying biological or other patient factors, or represent disparities in care. Adjusting for these factors could have the unintended consequence of creating or reinforcing disparities. Furthermore, Tables 13 and Figure 6 show that adjustment for these factors had minimal impact on dialysis practitioner group performance. Therefore, these risk factors were not included in the final risk adjusted model.

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)

Risk factors were selected for the final model based on the magnitude of the coefficients, evaluation of their statistical significance, and the model C-statistic. The C-statistic measures the discriminative power of the regression model with considered risk factors. Two-way interactions were examined and selected for the final model based on both the magnitude and statistical significance of the estimates.

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)

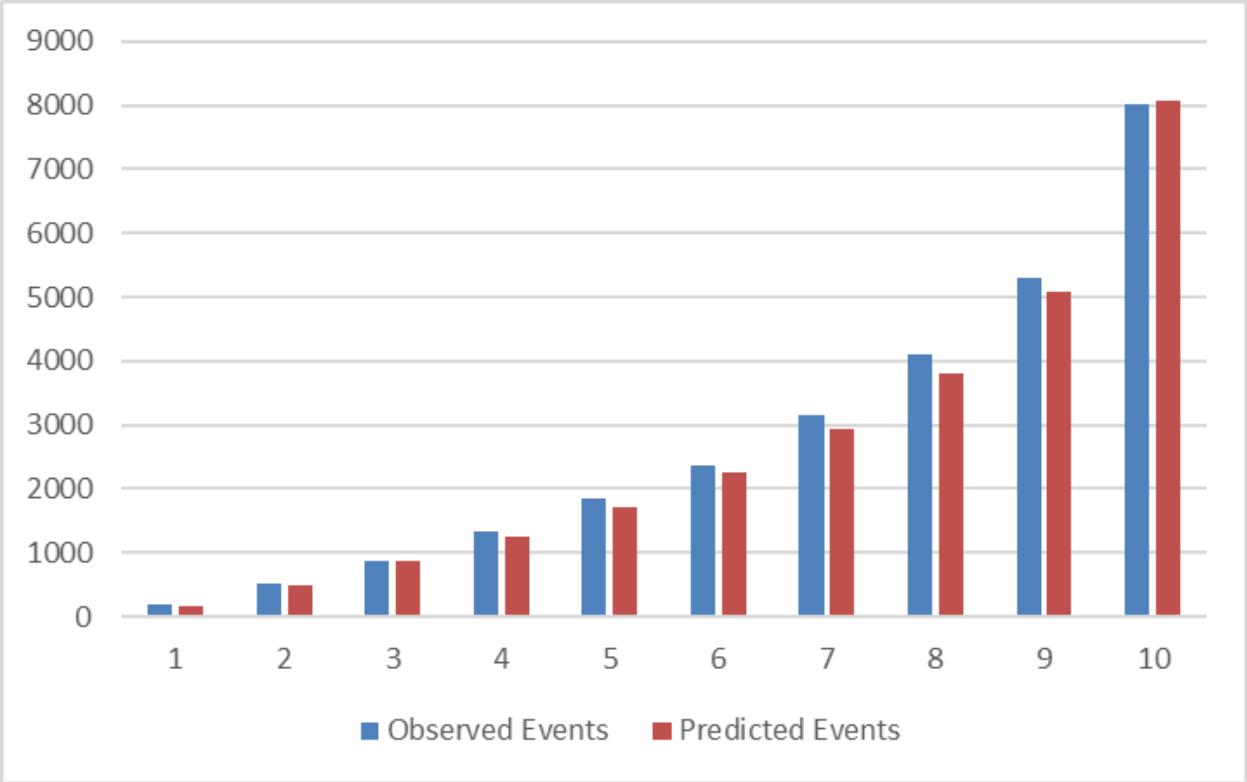
The C-statistic (also known as the Index of Concordance) was 0.763, meaning that the model correctly ordered 76.3% of the pairs of patient-months that were discordant with respect to the response variate. Month-specific C-statistics were computed in order to identify any trends by month in the model's discriminatory ability.

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-Lemshow (H-L) statistic is defined strictly for independent trials, and months within-patient are expected to be highly correlated. We therefore chose to perform the H-L statistic in a month-specific fashion, with the p values being low (p=0.0065 for January). However, in very large samples such as this even relatively small departures from the model will lead to significant results. While the p-value is significant, based on the decile plot in Figure 6 below, the observed and expected values by decile appear to be stable.

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)

Figure 7: *Observed and expected number of patients waitlisted by risk decile.*



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

N/A

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

Figure 7, above in section 2b.29, shows that in no decile is there an important discrepancy between the observed number of waitlisted patients in a decile and that predicted by the model.

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)

To test the null hypothesis that the aPPPW for a given dialysis practitioner group is statistically different from the national average, we use a simulation method to calculate the nominal p-value as the probability that the observed number of events (a binary outcome of 0 indicates that the patient is not on the waitlist in active status in during that month and a binary outcome of 1 indicates that the patient is on the waitlist in active status during that month) should be at least as extreme as that expected. This calculation is based on the supposition that, having adjusted for case mix, this practitioner group has a true event rate corresponding to the average practitioner groups. We then converted the p-values to z-scores. Using robust estimates of location and scale based on the normal curve fitted to the center of the z-scores, we derive the mean and variance of a normal empirical null distribution. The empirical null distribution is then used to calculate the p-value for each dialysis practitioner. Finally, dialysis practitioner group practices are flagged if they have outcomes that are extreme when compared to the variation in the national waitlist rate.

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)

***Table 3:** Count (%) of dialysis practitioner group practices and median aPPPW, stratified by classification category*

Classification category	***N (%)***	***Median aPPPW***
Better than Expected	*59 (2.6)*	*19.9*
As Expected	*2104 (92.4)*	*11.9*
Worse than Expected	*113 (5.0)*	*3.4*
Total	*2276 (100)*	*11.7*

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

As shown in Table 2, most dialysis practitioner group practices (92.4%) had an aPPPW that was "As Expected". Approximately 2.6% of dialysis practitioner group practices has a aPPPW that was "Better than Expected", while approximately 5.0% were "Worse than Expected". Across these categories,

performance on waitlisting in active status varied widely (from 3.4% of patients waitlisted in the worse than expected category, to nearly 20% in the better than expected category), suggesting that differences are also clinically meaningful.

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)

Many data elements can be obtained from multiple sources and missing data occurs rarely for covariates included in this measure.

Age is calculated using the date of birth and reporting month. Date of birth is required in our Standard Analysis Data Files, therefore no missing values were identified in the patient population. We assessed missing data for the CMS-2728 form which is used to determine incident comorbidities (i.e. at the time of dialysis initiation).

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)

Table 4: Distribution of missing data among 280,855 patients

Data element	***Missing (%)***
Patients with missing CMS-2728	***3,125 (1.11)***

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

There is a low percentage of patients with missing CMS-2728 Forms. Missing CMS-2728 was accounted for with a category for missingness in the model. As shown in Table 9 in section 2b.24, patients with

missing CMS-2728 form have a lower odds of waitlisting compared to those without a missing CMS-2728 form (OR = 0.510 ; 95% CI = 0.490, 0.531).

Note: This item is directed to measures that are risk-adjusted (with or without social risk factors) OR to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eCQMs). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator). Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.

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.)

To what extent are the data elements needed for the measure available electronically (i.e., needed elements to compute quality measure scores are in defined, computer-readable fields)?

- ☐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. One potential application for the measure is in the Quality Payment Program where it would be one of several optional measures that a group practice could select in their evaluation.

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

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)

N/A

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

The measure is not yet implemented in a public reporting program, so improvement could not be evaluated. CMS currently anticipates implementation of this waitlisting measure. Once implemented dialysis practitioner group practice performance on the measure can be evaluated to determine if the measure has supported and detected quality improvement in waitlisting rates among the target population.

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)

N/A

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

N/A

Additional Information (NQF Measure Submission Form, Additional)

Appendix

Available in attached files.

Other Additional Information

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

David Axelrod, MD, MBA
Transplant Surgeon, University of Iowa

Amy Waterman, PhD
Professor of Medicine, Nephrology, UCLA Nephrology

Bobby Howard
Patient, Director, Multicultural Donation Education Program
LifeLink of Georgia
Association of Organ Procurement

Jesse Schold, Mstat, PhD
Research Director, Cleveland Clinic

Emily Watson, MSW, LCSW
Social Worker, Satellite Healthcare, LLC

Krista Lentine, MD, PhD Professor of Medicine
American Society of Nephrology Policy & Advocacy Committee
Saint Louis University ASN Alliance for Kidney Health

Bryan N. Becker, MD, MMM,
Physician, DaVita, Inc.

John T. Ducker, MD, Transplant Nephrologist
Nephrology Associates of Northern Illinois and Indiana
Renal Physicians Association

Teri Browne, PhD, MSW
Associate Dean and Professor
University of South Carolina College of Social Work

Rachel Patzer, PhD, MPH,
Director, Health Services Research Center
Emory University School of Medicine

Della Major, MA
Patient, National Forum of ESRD Networks, member of the Kidney Patient Advisory Council

Sumit Mohan, MD, MPH
Physician and Epidemiologist, Columbia University
American Society of Nephrology Alliance for Kidney Health

Dawn P. Edwards
Patient, National Forum of ESRD Networks Kidney Patient Advisory Council

Geraldine Zingraf, DNP, MBA, RN, CNN, CCTC
Transplant Administrator, Edward Hines, Jr. VA Hospital

Sasha Couch
Patient, Renal Support Network

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2. First Year of Measure Release

2022

Ad.3. Month and Year of Most Recent Revision

01/2022

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

Annual

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

4/2023

Ad.6. Copyright Statement

N/A

Ad.7. Disclaimers

N/A

Ad.8. Additional Information/Comments

N/A

REFERENCES

Ashby VB, Kalbfleisch JD, Wolfe RA, et al. Geographic variability in access to primary kidney transplantation in the United States, 1996-2005. American Journal of Transplantation 2007; 7 (5 Part 2):1412-1423.

Grams, M. E., Massie, A. B., Schold, J. D., Chen, B. P., & Segev, D. L. (2013). Trends in the inactive kidney transplant waitlist and implications for candidate survival. American Journal of Transplantation, 13(4), 1012-1018.

Husain SA, Brennan C, Michelson A, Tsapepas D, Patzer RE, Schold JD, Mohan S. Patients prioritize waitlist over posttransplant outcomes when evaluating kidney transplant centers. Am J Transplant. 2018 Nov;18(11):2781-2790.

Kucirka LM, Grams ME, Balhara KS, Jaar BG, Segev DL. Disparities in provision of transplant information affect access to kidney transplantation. Am J Transplant. 2012 Feb;12(2):351-7.

Kulkarni S, Ladin K, Haakinson D, et al. Association of Racial Disparities With Access to Kidney Transplant After the Implementation of the New Kidney Allocation System. JAMA Surg 2019; 154(7):618-625.

Melanson TA, Gander JC, Rossi A, et al. Variation in Waitlisting Rates at the Dialysis Facility Level in the Context of Goals for Improving Kidney Health in the United States. Kidney International Reports 2021;6:1965-1968. No abstract.

Paul S, Plantinga LC, Pastan SO, Gander JC, Mohan S, Patzer RE. Standardized Transplantation Referral Ratio to Assess Performance of Transplant Referral among Dialysis Facilities. Clin J Am Soc Nephrol. 2018 Feb 7;13(2):282-289.

Patzer RE, Plantinga L, Krisher J, Pastan SO. Dialysis facility and network factors associated with low kidney transplantation rates among United States dialysis facilities. Am J Transplant. 2014 Jul; 14(7):1562-72.

Salter ML, Orandi B, McAdams-DeMarco MA, Law A, Meoni LA, Jaar BG, Sozio SM, Kao WH, Parekh RS, Segev DL. Patient- and provider-reported information about transplantation and subsequent waitlisting. J Am Soc Nephrol. 2014 Dec;25(12):2871-7.

Satayathum S, Pisoni RL, McCullough KP, et al. Kidney transplantation and wait-listing rates from the international Dialysis Outcomes and Practice Patterns Study (DOPPS). Kidney Intl 2005 Jul; 68 (1):330-337.

Tonelli M, Wiebe N, Knoll G, et al. Systematic review: kidney transplantation compared with dialysis in clinically relevant outcomes. American Journal of Transplantation 2011;11:2093-2109.

Waterman AD, Peipert JD, Goalby CJ, Dinkel KM, Xiao H, Lentine KL. Assessing Transplant Education Practices in Dialysis Centers: Comparing Educator Reported and Medicare Data. Clin J Am Soc Nephrol. 2015 Sep 4;10(9):1617-25.

United States Renal Data System. 2020 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, 2020.