



August 12, 2020

The Honorable Seema Verma
Administrator
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Baltimore, MD 21244

Re: CMS-1732-P: “End-Stage Renal Disease Prospective Payment System, Payment for Renal Dialysis Services Furnished to Individuals with Acute Kidney Injury, and End-Stage Renal Disease Quality Incentive Program”

Dear Administrator Verma:

Kidney Care Partners (KCP) appreciates the opportunity to provide comments on the “End-Stage Renal Disease Prospective Payment System, Payment for Renal Dialysis Services Furnished to Individuals with Acute Kidney Injury, and End-Stage Renal Disease Quality Incentive Program” (Proposed Rule). This letter outlines our support for the proposals related to the End-Stage Renal Disease (ESRD) Quality Incentive Program (QIP) and highlights concerns about the validity and reliability of some of the measures, as well as structural problems, including the impact of the pandemic on the QIP. Our comments on the prospective payment system will be shared in a separate letter.

KCP is an alliance of more than 30 members of the kidney care community, including patient advocates, health care professionals, providers, and manufacturers organized to advance policies that support the provision of high-quality care for individuals with chronic kidney disease (CKD), including those living with End-Stage Renal Disease (ESRD).

As described in more detail below, KCP strongly supports the four proposals CMS outlines in the Proposed Rule for the ESRD QIP:

- Updating the specifications used to calculate the Ultrafiltration Rate and Medication Reconciliation measures;
- Reducing the number of records facilities selected for the National Health Safety Network (NHSN) validation are required to submit;
- Clarifying the timeline for facilities to make changes to their NHSN Bloodstream Infection (BSI) clinical measure and NHSN Dialysis Event reporting measures; and

- Establishing the performance standards and payment reductions that would apply for PY 2023.

In addition, KCP is pleased that CMS has affirmed its “plan to re-evaluate our reporting measures for opportunities to more closely align them with NQF measure specifications.”¹ In light of this effort, KCP also offers suggestions with regard to specific measures that would allow the Agency to meet this goal. We also encourage CMS to evaluate the existing QIP measures consistent with the following principles and include those modifications in the final rule this year. As we noted in our 2019 comment letter on the ESRD QIP, we ask that CMS:

- Use valid and reliable measures as established through NQF endorsement;
- Adopt endorsed measures when they are available over measures that have not been endorsed;
- Not use or remove measures that NQF has rejected as part of its endorsement process from the ESRD QIP or that have been assigned to reserve status;
- Avoid modifying NQF-endorsed measures when adopting them for the ESRD QIP;
- Seek NQF endorsement for new measures prior to adopting them in the ESRD QIP or at least use them only as reporting measures while seeking NQF-endorsement;
- Honor its commitment to use rate measures in favor of ratio measures;
- Continue to work with stakeholders in a transparent process to identify and address the potential causes that could lead to the penalties increasing when actual performance has improved;
- Work with the community and NQF to develop a better approach to the small numbers problem; and
- Align the ESRD QIP and ESRD DFC/Five Star.

In addition to the comments on the specific measures, KCP provides suggestions to address the differential handling of Medicare Advantage patients in several measures in the ESRD QIP and how to address the pandemic in a manner that ensures the integrity of the ESRD QIP long-term.

We continue to support the two vascular access measures in the ESRD QIP. We also support the decision not to add any new measures to the ESRD QIP at this time. There are now 14 ESRD QIP measures (not counting the pooled measure for dialysis adequacy), which dilutes the impact of any one of these measures. As noted below, we propose reducing the current measure set by removing some of the measures. We look forward to working with CMS to make sure that there is a parsimonious set of measures reflecting the

¹CMS, “Medicare Program; End-Stage Renal Disease Prospective Payment System, Payment for Renal Dialysis Services Furnished to Individuals With Acute Kidney Injury, and End-Stage Renal Disease Quality Incentive Program” 85 *Fed. Reg.* 42132 (July 13, 2020).

most critical outcomes for patients and accurately reflecting the care actually provided by the facilities.

I. The use of valid and reliable measures that align with NQF-endorsed measures

Measures used in the ESRD QIP should be endorsed by NQF to be consistent with the statutory mandate, unless as the statute notes, there is no endorsed measure in a specific domain. Section 1890 of the Social Security Act (SSA) requires CMS to contract with a consensus-based entity for developing measures used in VBPs. The second statutory duty listed for the consensus-based entity, which is currently NQF, is to endorse measures for CMS' use. When the Congress established the ESRD PPS, it was even more specific in its mandate to use NQF endorsed measures. The Statute requires that "any measure specified by the Secretary under subparagraph (A)(iv) **must have been endorsed by the entity with a contract under section 1890(a).**"² Thus, KCP is pleased that the preamble states that CMS plans to more closely align the QIP measures with the NQF measure specifications. KCP recommends that to achieve this goal not only for reporting measures, but also clinical measures, CMS take the following steps outlined below. KCP also strongly opposes use of measures in the QIP that NQF has rejected through the endorsement evaluation process. Simply put, CMS should use valid and reliable measures as established through NQF endorsement.

A. KCP supports aligning the QIP Ultrafiltration and Medication Reconciliation Denominators with the NQF-Endorsed Specifications: CMS should avoid modifying NQF-endorsed measures when adopting them for the ESRD QIP.

KCP is pleased that CMS has proposed to update the specifications used for the **Ultrafiltration Rate (UFR; NQF 2701)** reporting measure by stating that it will use the patient-months construction that comports with the NQF-endorsed measure. We also appreciate the clarification that this reporting measure is based on the one for which the Kidney Care Quality Alliance (KCQA) is the steward. Similarly, we are pleased that the preamble also reaffirms that it will no longer use the "facility-months" construction for the **Medication Reconciliation (NQF 2988)**. Using the a "patient-months" denominator construction aligns both of these measures with the specifications submitted by the measure developer and steward (the KCQA), which were endorsed by NQF. KCP appreciates and concurs with the change to the "patient-months" construction for both measures.

²SSA § 1881(h)(2)(B) (emphasis added).

B. KCP supports the effort to reduce the burdens created by the NHSN validation study; KCP also reiterates concerns that the NHSN Bloodstream Infection measure is not valid and needs to be modified to provide accurate information to patients and actionable information to facilities.

CMS proposes to reduce the submission requirement for facilities selected to participate in the NHSN validation study from 40 to 20 patient records from any two quarters during the applicable calendar year. KCP concurs with this reduction and appreciates this revision, which will reduce facility burden. We also support the clarification for both the **NHSN Dialysis Event** and the **NHSN Bloodstream Infection** measures that “any changes that a facility makes to its data after the ESRD QIP deadline that applies to those data will not be included in the quarterly permanent data file that the CDC generates for purposes of creating the annual CMS ESRD QIP Final Compliance File.”³

While KCP continues to support the NHSN Dialysis Event measure as a reporting measure, we encourage CMS to submit this measure for NQF review, consistent with the statutory language indicating that CMS should use measures endorsed by the body selected to review them, which in this case is the NQF. Therefore, we also ask that CMS submit the measure to NQF for review in the next cycle. We reiterate our recommendation that the recent addition of a set of subjective factors (*e.g.*, redness, swelling) to the measure be eliminated because these factors do not support the purpose of the measure.

Consistent with our previous recommendations, KCP asks CMS to eliminate the **NHSN Blood Stream Infection (BSI)** measure while it determines how to revise the specifications so that the validity problems with the measure can be resolved and the NQF has the opportunity to review the measure. CMS has not identified data indicating that the problem that **as many as 60-80 percent of dialysis events may be under-reported** with the NHSN BSI measure has been resolved. The measure does not meet the criterion of validity for endorsement. Thus, patients who rely upon the information generated by this measure are, in many instances, relying on inaccurate data that suggest that a particular facility has a low number of blood stream infections when, in fact, the facility has a higher number. The importance of understanding how a facility manages bloodstream infections is critical for patient decision-making. A measure that fails to accurately represent the facility’s performance deprives patients of their ability to make informed health care decisions. It also unfairly penalizes facilities that diligently pursue and report the hospital infection data necessary for a full picture of infection rates.

Thus, we reiterate our request that CMS remove the NHSN BSI clinical measure immediately and use the Dialysis Event Reporting Measure alone. KCP strongly supports

³*Id.*

transparency and efforts to reduce bloodstream infections. Therefore, we ask CMS to work with the community to identify specific modifications to the NHSN BSI measure to address the validity concerns and submit that revised measure to the NQF for review.

C. KCP continues to support the conversion of the Standardized Transfusion Ratio (STrR) measures to a reporting measure, because of concerns about validity arising from the shift to ICD-10 coding, but urges CMS to replace it with a more appropriate anemia management measure and seek endorsement of the new measure.

KCP continues to support the statutory requirement that CMS adopt endorsed measures when they are available, but recognizes that there may be times when changing circumstances result in an endorsed measure no longer being appropriate. As we noted during last year's rulemaking, we support CMS addressing these problems as they arise.

For example, KCP continues to support the decision CMS made to convert the **Standardized Transfusion Ratio (STrR/NQF 2979)** to a reporting measure. Because it became clear after the ICD-9 to ICD-10 transition that the codes used in the STrR measure were not accurately capturing blood transfusions to ensure validity of the measure, CMS converted the measure to a reporting metric in the CY 2019 Final Rule to allow for an examination of the problem. Going forward, however, KCP recommends shifting away from the STrR measure and adopting a measure that more directly reflects patient quality of care, is more clearly actionable, and reduces burden. We again recommend that CMS replace the STrR with a low hemoglobin (Hgb) measure (*e.g.*, a Hgb <10 g/dL).

While it will be necessary to develop updated specifications, exclusions, testing, and business rules, KCP would welcome the opportunity to work with CMS on such a measure; we note that CMS developed a similar measure several years ago that would be an appropriate starting point. We are aware such a measure was not endorsed by NQF, but believe NQF's updated evidence algorithm provides a path for its consideration anew. A low Hgb measure would reduce burden, because any transfusion measure requires dialysis facilities to chase paperwork created by other providers. It also is a better measure than the STrR because facilities and physicians have access to patient Hgb data in the facility, whereas they do not have access to transfusion data. Moreover, it is actionable by physicians and will have a direct positive impact on an issue of critical import to patients. Additionally, as we note in the following section, KCP has significant concerns about the reliability of the STrR.

D. When NQF has rejected a measure or moved a measure to reserve status, CMS should not include it in the QIP to be consistent with the statute; thus, KCP asks CMS to remove the Prevalent Patients Waitlisted measure and retire the Hypercalcemia measure from the QIP and replace the Dialysis Adequacy Comprehensive Measure with the individual Kt/V measures that NQF has endorsed.

KCP strongly supports the President's initiative to increase the number of successful kidney transplants. To achieve the goal, it is important that patients are empowered by having accurate information to assess whether their providers are doing what they can and should be to help them qualify for a transplant. Having a valid and reliable measure in the ESRD QIP that supports transplants is a worthy goal.

Unfortunately, the **Percentage of Prevalent Patients Waitlisted (PPPW)** measure has been determined to lack validity by the NQF. Thus, it should not be included in the QIP, because it will mislead patients. While CMS has flexibility to adopt a measure when NQF has not endorsed a measure in a particular domain, it is a contorted reading to suggest that this flexibility means the Agency can or should use a measure that has failed to meet the scientific, consensus-based endorsement criteria.

Rather than continue with this measure, we encourage CMS to work with KCP and others in the community to address the problems underlying this measure so that there is a valid and reliable measure that will provide accurate information related to transplantation and empower patients in their decision-making.

Similarly, the NQF has concluded after extensive review that the **(Kt/V) Dialysis Adequacy Comprehensive Measure** does not meet the endorsement criteria, because it failed on measuring a performance gap, which is a threshold requirement for further discussion on factors such as validity and reliability. KCP is also concerned that a pooled measure fails to provide the transparency necessary to promote patient decision-making when it comes to home dialysis. By reporting all Kt/V scores, it hides from view how each facility performs when it comes to providing home dialysis. Given the Administration's emphasis on home dialysis, we urge CMS to remove the Dialysis Adequacy Comprehensive Measure from the QIP and replace it with the following measures that have meet the endorsement criteria:

- NQF #0249 Delivered Dose of HD Above Minimum;
- NQF #0318 Delivered Dose of PD Above Minimum;
- NQF #1423 Minimum spKt/V for Pediatric HD Patients
- NQF #2704, Minimum Delivered PD Dose;
- NQF #2706, Pediatric PD Adequacy—Achievement of Target Kt/V

This step would align with the statutory mandate and provide patients with the ability to understand each facility's actual performance on the different dialysis modalities.

KCP also recommends that CMS work with KCP to re-specify and test new individual PD measures so that facilities that provide home dialysis are not disadvantaged because of the differences in the frequency of testing home dialysis patients.

Finally, CMS should retire the **Hypercalcemia Measure** from use in the ESRD QIP, because it is based on NQF #1454, which the NQF has placed in reserve status because it has “topped-out” (*i.e.*, there is little room for additional improvement in this clinical area) and provides no significant benefit for patients. Therefore, CMS should retire the measure.

In addition, KCP reiterates that it would be appropriate, for purposes of having a bone mineral metabolism measure, to use the NQF serum phosphorus measure as a reporting measure in the QIP. Even though the measure is in reserve status, physicians rely upon the serum phosphorus measure to make clinical decisions. While work still needs to be done to identify the optimal phosphorus target, how to address the target for certain subpopulations, and when phosphorus should be assessed, a reporting measure emphasizes the need to monitor phosphorus levels while allowing time to address these unresolved issues.

E. KCP encourages CMS to address the reliability problems with the standardized ratio measures and to use risk-standardized rate measures instead.

KCP members believe that hospitalization and readmission rates are essential metrics that should be the core of any value-based purchasing program. However, for such metrics to be effective they must be reliable – meaning accurate and replicable in how they measure facility performance – and transparent. Unfortunately, the **Standardized Hospitalization Ratio, (SHR/NQF 1463)** and **Standardized Readmission Ratio (SRR/NQF 2496)** measures, as well as the STrR, do not meet these requirements, as CMS’s own data demonstrate.

CMS’ decision to provide only average reliability statistics across all facility sizes lacks transparency. To improve, a facility should be able to assess the degree to which its own SHR or SRR scores represent noise or actual quality results. While reliability data stratified by size may no longer be required by NQF for endorsement, it is critical data for facilities to understand their performance and improve upon it. KCP strongly recommends that CMS provide these data in its NQF submissions or make them publicly available elsewhere.

In the most recent iteration of the SRR, currently under review at NQF, the overall IUR was 0.35—a dramatic decline from the 2009 NQF submission value of 0.55. Statistical

literature traditionally interprets a reliability statistic <0.5 as “unacceptable”.⁴ A measure wherein 65 percent of a facility’s score is due to random noise and not a quality signal is inappropriate for use in the QIP. Moreover, the SRR’s reliability of 0.35 is the average across all facilities. The reliability for smaller facilities will be significantly less, as acknowledged by CMS’ contract developer.

Likewise, the overall IUR for the one-year SHR was 0.53-0.59 for 2015-2018; a “poor” reliability statistic that also represents a decline from the 2010-2013 IURs (0.7). Based on current CMS data, 41-47 percent of a facility’s SHR score is due to random noise, and smaller facilities again will have a significantly greater contribution of noise to their score.

Again, KCP also notes that CMS now declines to provide testing data stratified by facility size for any measures it submits to NQF because it is “not required” by NQF. As the most recent CMS reliability data stratified by size reveal, the IUR for small facilities (defined by CMS at the time as <50 for the SHR and <70 for the SRR) for both measures was 0.46 in 2009 (SRR) and 2013 (SHR)—*i.e.*, for approximately one third of all facilities (those meeting CMS’ own definition of “small”), 54 percent of the score they received on the SRR and SHR could be attributed to random noise and not signal.

Any score assigned to a facility for the SRR has no quality meaning based on CMS testing results, and the SRR should be removed from the QIP. The SHR should be deployed only for large facilities, as defined by CMS’ historical stratification results in its submissions to NQF. Finally, although the clinical version of the STrR is not yet proposed, KCP feels it is important also to emphasize its poor reliability, especially for small facilities. In the most recent iteration of the measure, the overall IUR for the one-year STrR was 0.63-0.68 across the years 2014-2017. Data from 2011-2014, for which there was a similar overall IUR, revealed values as low as 0.30 for small facilities—that is, for approximately one third of facilities, 70% of the score they received on the STrR could be attributable to random noise and not signal. While new details were not provided, CMS’ contract measure developer acknowledged that the STrR was less reliable in smaller facilities for the 2014-2017 data period.

Lastly, although not mentioned *per se* in the Proposed Rule, we note that CMS now relies on a novel, additional metric of reliability, referred to as the profile-IUR (PIUR).⁵ Per CMS, “The PIUR indicates the presence of outliers or heavier tails among the providers, which is not captured in the IUR itself. . . . [When] there are outlier providers, even measures with a low IUR can have a relatively high PIUR and can be very useful for

⁴ Adams JL. The Reliability of Provider Profiling: A Tutorial. Santa Monica, CA:RAND Corporation. TR-653-NCQA, 2009.

⁵ He K, Dahlerus C, Xia L, Li Y, Kalbfleisch JD. The profile inter-unit reliability. *Biometrics*. 2019 Oct 23. doi: 10.1111/biom.13167. [Epub ahead of print].

identifying extreme providers.”⁶ The PIUR was 0.61 for the SRR and 0.75-0.85 for the SHR, which CMS interprets as demonstrating the measures are “effective at detecting outlier facilities and statistically meaningful differences in performance scores across dialysis facilities.”⁷ KCP strongly concurs with the NQF’s Scientific Methods Panel (SMP) conclusion that the PIUR is not an appropriate measure of reliability for any QIP measure. QIP measures are used to distinguish performance along a continuum, in particular among providers falling in the *middle of the curve*, to determine penalties; the ability to reliably distinguish outliers for implementation of these measures is not the point. The IUR is and remains the appropriate measure of reliability for measures proposed for the QIP.

F. Modifying ICH-CAHPS measure to address validity problems and make it meaningful to patients and providers.

KCP continues to support patient satisfaction measures, such as the **ICH-CAHPS** measure. However, the low response rates threaten the validity of ICH-CAHPS as an accountability measure. In addition, the current measure does not allow for feedback from home dialysis patients. We appreciate the Technical Expert Panel that CMS convened earlier this year and support the close review of the measure. However, there are some immediate modifications CMS could adopt that would reduce the burden on patients asked to respond and address some of the response rate problems. Specifically, CMS could:

- Administer ICH-CAHPS to patients once a year (not twice) to reduce burden on patients; and
- Ask individual patients to complete only one of the three independently validated sections on the survey; thus, while facilities are subject to the entire survey instrument, no one patient will be asked to complete the more than 60 questions in a single response.

In addition, we reiterate our outstanding request that the survey be revised to include home dialysis patients and that CMS obtain NQF endorsement of the new measure, which MedPAC and others in the community also have consistently requested. It is also important that CMS allow facilities and patients to use the ICH-CAHPS survey results to improve care.

II. Differential Handling of Medicare Advantage Patients in QIP measures threatens the validity of several QIP measures.

The increasing numbers of MA patients in the ESRD program—and the unavailability of outpatient claims data for these patients—threaten the validity of several

⁶ Kalbfleisch JD, He K, Xia L, Li Y. Does the inter-unit reliability (IUR) measure reliability? *Health Services and Outcomes Research Methodology*. 2018;18(3):215-225. Doi: 10.1007/s10742-018-0185-4.

⁷ Citation: SHR measures submission materials to NQF.

QIP measures. Data provided by CMS indicate that at the end of 2017, 27 percent of dialysis patients had MA coverage (presumably higher now), and this varied widely across states—from about 2 percent in Wyoming to 34 percent in Rhode Island, and more than 44 percent in Puerto Rico. Such geographic variation compromises the validity of the measures if MA patients are not accurately accounted for in the QIP metrics. Specifically, without changes to the current specifications, the evolving patient mix will introduce significant bias into measure calculations that could affect results for facilities with either very low or high MA patient populations. Recognizing this, KCP concurs with the need to change specifications for several CMS measures to accommodate the increase in MA patients and to avoid disparities in performance due to geography. KCP strongly believes, however, that greater transparency is required by CMS as it updates the relevant measures.

While the approach to handling MA patients varies considerably across CMS' metrics (Table 1, Attachment B), KCP recognizes the difficulty CMS faces in addressing this issue across measures of varying construction and notes there appears to be a logical rationale for most of the decisions made because of the properties and intended purpose of each measure. Nevertheless, KCP strongly recommends that CMS perform a sensitivity analysis of performance with and without MA patients for each of the applicable QIP measures and make the results publicly available. Such data will provide an opportunity for KCP and others to offer potential, evidence-based mitigation strategies (*e.g.*, a model that accounts for both populations, use of risk coefficients as necessary).

We also ask CMS to perform and provide an analysis of risk model fit under the previous approach and the new in-patient-claims-only approach; currently we are unable to assess whether model fit improved or worsened with this approach. KCP is particularly concerned that limiting comorbidity data to inpatient claims might skew the models towards a sicker population, and that such a skew might reflect unfavorably on facilities that successfully keep hospitalization rates low. That is, because comorbidity adjusters developed exclusively from hospitalization data will necessarily underestimate the comorbidity profile of patients in facilities with low hospitalization rates, the “expected” hospitalization or mortality rates calculated for such facilities will be erroneously low, and the facilities' scores will be erroneously high. Only with transparency in these matters can the community assess the impact MA patient mix has on the QIP measures.

Finally, KCP notes that the SHR and SFR (and Standardized Mortality Ratio (SMR), which is not part of the QIP) obtain past-year comorbidity data from multiple Part A sources (inpatient, SNFs, home health, hospice). Conversely, the past-year comorbidity source for the SRR is limited to inpatient claims. We ask that CMS incorporate data from the multiple Part A sources used in the SMR, SHR, and SFR models—inpatient, as well as SNF, home health, and hospice data—to make the SRR adjustment potentially more robust. As a matter of measure construction, it also is a logical harmonization issue. We recommend CMS perform this analysis and make it publicly available or release existing data and justify the current approach.

III. KCP asks CMS to address the impact of COVID-19 measure performance.

The COVID-19 pandemic has presented unprecedented challenges to patients with ESRD and the dialysis community and has significantly affected patient care—and has the potential to impact the QIP. The pandemic will impact performance beyond the obvious outcome measures such as the SHR and SRR in areas with a heavy COVID burden, but also “upstream” process and intermediate-outcome metrics, even in relatively unaffected locales. For instance, to avoid or minimize potential exposure to the virus, patients and providers have postponed elective fistula placement and delayed routine lab draws, and adequacy targets have not been met in some cases as anxiety sometimes means an early end to a dialysis session.

A. KCP asks CMS to extend the nationwide Extraordinary Circumstances Exception for the ESRD QIP through the end of the public health emergency, plus a short grace period.

KCP appreciates CMS’ proactive granting of a universal Extraordinary Circumstance Exception (ECE) for the ESRD QIP in response to COVID-19. We likewise thank CMS for allowing facilities the flexibility to opt out of the ECE, at their discretion. We note, however, that the recently witnessed progressive and unpredictable regional spread of the virus now renders the current June 20 deadline for this decision obsolete. Previously unaffected facilities that chose to opt out of the ECE prior to June 20 may now be in the center of a new “hotspot”, no longer able to meet the required data submission that previously seemed feasible. KCP thus requests that CMS revisit the June 20 deadline, allowing facilities that previously opted out of the ECE to now opt-in, without penalty.

We believe that CMS has the authority to extend the flexibility provided in the universal ECE. CMS created the ECE policy through regulation. 42 C.F.R. § 413.178(d)(3) indicates that the timeframe for an ECE may be “for one or more calendar days, when there are certain extraordinary circumstances beyond the control of the facility.” The regulations also indicated that “CMS may grant exceptions to facilities without a request if it determines that one or more of the following has occurred: (i) An extraordinary circumstance affects an entire region or locale.”⁸ There is no other time restriction.

The statute governing the ESRD QIP does not prohibit CMS from extending exceptions to the reporting requirements. While the statute requires CMS to reduce payments to a dialysis facility that does not meet or exceed the total performance score with respect to the performance standards, this requirement is subject to the discretion of the Secretary as evidenced by the clause to which the requirement is subject “as

⁸42 C.F.R. § 413.178(d)(6).

determined appropriate by the Secretary.”⁹ This phrase gives the Secretary the authority to establish the ECE.

In addition, CMS indicated the statute clearly authorizes the ECE through the discretion the Secretary is provided to develop the methodology for setting the total performance score. There is no time limitation on this authority either.

Section 1881(h)(3)(A)(i) of the Act states, “[T]he Secretary shall develop a methodology for assessing the total performance of each provider of services and renal dialysis facility based on performance standards with respect to the measures selected under paragraph (2) for a performance period established under paragraph (4)(D).” Given the possibility that facilities could be unfairly penalized for circumstances that are beyond their control, we believe the best way to implement an extraordinary circumstances exception is under the authority of this section. We therefore proposed to interpret section 1881(h)(3)(A)(i) of the Act to enable us to configure the methodology for assessing facilities’ total performance such that we will not require a facility to submit, nor penalize a facility for failing to submit, data on any ESRD QIP quality measure data from any month in which a facility is granted an extraordinary circumstances exception.¹⁰

This authority and the rationale outlined when CMS finalized the ECE policy for CY 2015 supports extending the ECE period during the public health emergency, and we encourage CMS to do so immediately. We also encourage CMS to also consider extending the ECE for a grace period beyond once the public health emergency has ended (*e.g.*, 30-60 days) to provide time for providers to ramp back up, because areas/states will be hit unevenly.

B. KCP asks CMS to work with the KCP to address challenges the pandemic has created for the ESRD QIP.

As CMS has recognized through the nationwide ECE, the pandemic is an extraordinary circumstance over which we have no control. It has been devastating to providers and patients alike. The impact of the outbreaks in the United States has required an unprecedented response and changes in practice patterns that will remain with us throughout the duration of the public health emergency and, perhaps, even longer.

We note that, in addition to the short-term impact on patient care and outcomes, the COVID-19 pandemic will have effects on the QIP for several years after the pandemic ends. This is because the QIP relies on benchmarks set through

⁹42 U.S.C. § 1395rr(h)(1).

¹⁰CMS, “End-Stage Renal Disease Prospective Payment System, Quality Incentive Program, and Durable Medical Equipment, Prosthetics, Orthotics, and Supplies” *Display Copy* 240 (November 2014).

previous years' performance. To the extent that dialysis performance and measure reporting is anomalous due to COVID-19, those anomalies will affect the benchmarks in subsequent years.

KCP asked Discern Health to help us understand the potential impact of the disruptions created by the pandemic and how those disruptions could impact the accuracy and reliability of the ESRD QIP. Discern modeled three distinct scenarios to evaluate the impact of the ECE on QIP performance. It used the CY2019 QIP performance data to model each of the following scenarios.

- Scenario 1 – Baseline Scenario – The baseline scenario represents a “normal” QIP year, assuming no ECE and no impact from COVID.
- Scenario 2 – Current ECE Maintained – This scenario assumes that the terms of the ECE expiring in June are not amended. Accordingly, we are assuming smaller measure denominators as a result of the ECE, and poorer performance from July 2020 through December 2020. We also note that smaller denominators result in less reliability of the measure scores, but since we cannot estimate the impact of poor reliability on the distribution, the model does not account for this second-order effect.
- Scenario 3 – ECE Extended 3 Months – This scenario assumes that the ECE is extended another three months for measures reported through CROWNWeb or Claims.

The components of the model were:

- Measure threshold eligibility – Each measure included in QIP has a required denominator to evaluate a facility. For example, the Standardized Hospitalization Ratio (SHR) measure is not reported if there are fewer than 5 patient years at risk. A natural consequence of the ECE is a reduction in the denominator, which will push more facilities below that threshold.
- Impaired performance – The ECE was initially issued through the end of June. If the ECE is left to expire in June (Scenario 2) or is only extended another 3 months (Scenario 3), CY2020 will include data collected during the COVID-19 Pandemic. Even at low levels of community prevalence, the pandemic will likely affect measure performance. This interaction is dynamic and its direction and magnitude are not known. For example, the pandemic has been shown to discourage care seeking behavior, which may reduce hospitalizations measured by the Standardized Hospitalization Ratio (SHR)¹¹. On the other hand, some data illustrate the relatively

¹¹ <https://catalyst.nejm.org/doi/full/10.1056/CAT.20.0193>

high rate of hospitalization for those with ESRD¹². Accordingly, a specific impact on performance is not modeled, but potential implications of this variability are outlined as appropriate.

In addition, Discern considered how the ECE would affect the QIP during three calendar years, as outlined below:

- PY2022/CY2020 – Facilities with low volume in CY2020 as a result of ECE will be ineligible for performance scores.
- PY2023/CY2021 – Facilities with low volume in CT2020 as a result of ECE will be ineligible for improvement scores.
- PY2024/CY2022 – The ECE will affect the national performance standard used to calculate the Achievement Score.

Results for PY2022/CY2020 Measure Eligibility: Discern estimated the denominators for eight measures based on CY 2018 ESRD performance QIP data and the COVID-19 ECE FAQ. These eight measures were selected based on availability of data and denominators in the CY 2018 QIP dataset. Below, the number of facilities eligible for each measure (*Facilities*), and the percent of all facilities (*%Ttl*) they represent are shown:

Figure 3. ESRD Facilities Eligible for Each Measure

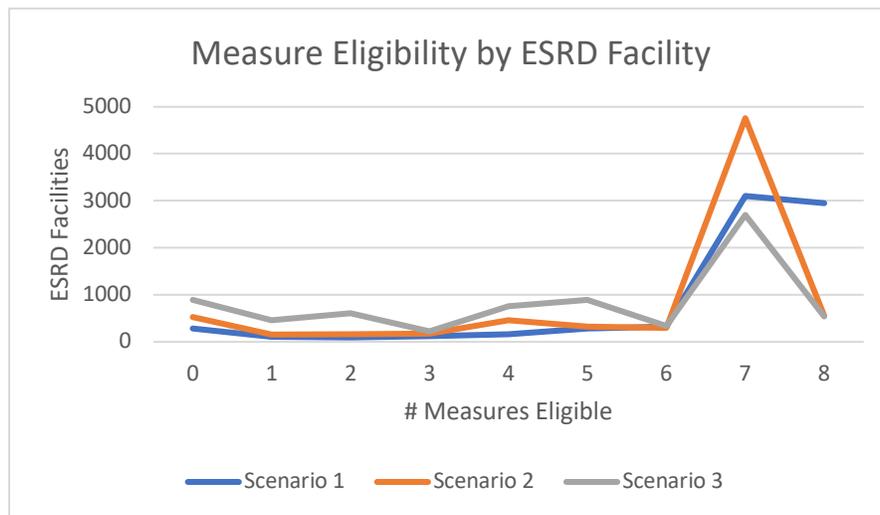
| | Scenario 1 | | Scenario 2 | | Scenario 3 | |
|--------------------|------------|------|------------|------|------------|------|
| | Facilities | %Ttl | Facilities | %Ttl | Facilities | %Ttl |
| Long Term Catheter | 6,475 | 87% | 5,673 | 76% | 3,598 | 48% |
| SFR | 6,442 | 87% | 5,579 | 75% | 3,341 | 45% |
| Kt/V | 7,055 | 95% | 6,620 | 89% | 5,403 | 73% |
| Hypercalcemia | 6,981 | 94% | 6,582 | 89% | 5,389 | 73% |
| ICH CAHPS Measures | 2,957 | 40% | 566 | 8% | 566 | 8% |
| SRR | 6,859 | 92% | 6,572 | 89% | 6,030 | 81% |
| STrR | 6,292 | 85% | 5,694 | 77% | 4,431 | 60% |
| SHR | 6,895 | 93% | 6,734 | 91% | 6,403 | 86% |

From this table, 6,895 facilities have sufficient volume to be eligible for the SHR in Scenario 1; 6,734 in Scenario 2; and 6,403 in Scenario 3. From this analysis, SRR and SHR measures retain fairly high coverage; Kt/V, Hypercalcemia, STrR, Long Term Catheter, and SFR have modest coverage; and the ICH CAHPS measure has poor coverage.

¹² <https://www.cms.gov/blog/medicare-covid-19-data-release-blog>

Discern also considered the number of measures for which each facility would be eligible. A significant number of facilities are still eligible for seven or eight of the eight analyzed measures. In Scenario 2, this is 72 percent of facilities, and 44 percent of facilities in Scenario 3, as compared to 82 percent in Scenario 1. Conversely, the number of facilities eligible for *no measures*, rises from 3.7 percent in Scenario 2, to 7.2 percent in Scenario 2, to 12.0 percent in Scenario 3.

Figure 4. Estimated Number of Eligible Measures by ESRD Facility



In addition to reduced eligibility, smaller denominators will increase the weight given to the national average through reliability adjustment.

Results for PY2023/CY2021: In PY2023, data from CY2020 will serve as the baseline for the improvement score. Technical guidance specifies that “If a facility does not have sufficient data to calculate a measure improvement rate... then the facility score for that measure is based solely on achievement,”¹³. We are assuming that the same threshold for measure eligibility is used for improvement score eligibility.

Because of the data excepted by the ECE, more facilities than usual will be ineligible for the improvement score. This is only an issue for facilities that would have otherwise received an improvement score. The table below estimates how many facilities would have received an improvement score if not for the ECE.

¹³ <https://www.cms.gov/files/document/esrd-measures-manual-v52.pdf>

Figure 5. Number of ESRD Facilities Receiving Lower Score Due to Ineligibility for Improvement Score

| | Scenario 2 | Scenario 3 |
|---|-------------------|-------------------|
| VAT Catheter Measure | 78 | 199 |
| VAT Fistula Measure | 73 | 170 |
| Kt/V Comprehensive Measure | 93 | 191 |
| Hypercalcemia Measure | 50 | 155 |
| ICH CAHPS Neph Comm and Caring Measure | 98 | 98 |
| ICH CAHPS Quality of Dialysis Care and Ops Measure | 71 | 71 |
| ICH CAHPS Providing Info to Patients Measure | 88 | 88 |
| ICH CAHPS Overall Rating of Neph Measure | 133 | 133 |
| ICH CAHPS Overall Rating of Dialysis Staff Measure | 108 | 108 |
| ICH CAHPS Overall Rating of Dialysis Facility Measure | 86 | 86 |
| SRR Measure | 20 | 51 |
| STrR Measure | 58 | 206 |
| SHR Measure | 20 | 64 |

For example, under Scenario 3, 199 facilities will be ineligible for an improvement score on the VAT Catheter Measure in PY2023, and will receive a lower score under the Achievement Score.

Results for PY 2024/CY2022: In PY2024/CY2022, data from CY2020 will serve as the national performance standard used to calculate the Achievement Score. Given that these targets are set nationally, even with a partial year of results, small number problems are unlikely for most measures. However, if the ECE continues through the end of 2020, the ICH CAHPS measures will have no data for 2020.

While the direction and magnitude of COVID-19's influence on measure performance is not known, the impact on the PY2024 national performance standard would counterbalance the effect on the PY2022 performance score. For example, if COVID-19 is a net harm to facility performance, more facilities would receive penalties in PY2022, but the Achievement targets would be lower in PY 2024. While these impacts counterbalance each other, their net effect is unclear.

As the Discern data show, there are several short- and long-term expected results of the ECE on the QIP and areas of uncertainty. Given the significant financial effect of the QIP and uncertainty around COVID-19's effect on the QIP, we ask that CMS:

- Perform an evidence-based impact assessment to determine the long-term effect of COVID-19 on measures used for QIP. Long-term consequences of COVID-19 are still being understood by the scientific community, and preliminary research suggests effects on multiple body systems.¹⁴ Other evidence suggests that COVID-19 leads to kidney damage, with 15 percent of those hospitalized requiring dialysis after discharge. The America that emerges from the PHE will be different from the one that enters it.
- Base Improvement and Achievement benchmarks upon the last full year of pre-COVID-19 performance, CY2019. Based upon the impact assessment, modification of these calendar year benchmarks may be needed.

COVID-19 presents a unique challenge for which there is little precedent, and there are likely no simple solutions (especially when we do not yet know the full impact). We believe these recommendations will stabilize the QIP into the future, and promote quality outcomes.

IV. KCP supports maintaining the structural aspects of the ESRD QIP for PY 2024, but encourages CMS to consider changes that will make payment reductions under the program more predictable.

As we have indicated in previous comment letters, we appreciate that CMS recognizes the importance of maintaining the structural aspects of the ESRD QIP year-to-year that allow for multi-year comparisons of providers. This consistency is appropriate and helpful. Thus, KCP the proposals for PY 2024 that maintain the performance period, performance standards, and scoring aspects of the program. We continue to urge CMS to weight certain measures, such as the reduction in catheter measure, more heavily than others.

A. Addressing unintended payment reductions.

We also reiterate our concerns that in past rulemakings the payment reduction scale has resulted in a substantial increase in the number of facilities being penalized under the ESRD QIP, even though the actual performance of the facilities was improving. We also reiterate our concerns that in past rulemakings the payment reduction scale has resulted in unpredictable percentages of facilities being penalized under the ESRD QIP, even though the actual performance of the facilities was improving.

¹⁴ <https://www.advisory.com/daily-briefing/2020/06/02/covid-health-effects>

Analysis by Discern has shown that any underlying changes in performance distribution could have large effects on this system. **Figure 6** shows the yearly changes in minimum TPS and payment reduction scales:¹⁵

Figure 6. Payment Reduction Scale PY 2020 - 2023

| Reduction % | PY 2020 | PY 2021 | PY 2022 | PY 2023 (estimated) |
|--------------------|----------------|----------------|----------------|----------------------------|
| 0.0% | 100-61 | 100-56 | 100-54 | 100-57 |
| 0.5% | 60-51 | 55-46 | 53-44 | 56-47 |
| 1.0% | 50-41 | 45-36 | 43-34 | 46-37 |
| 1.5% | 40-31 | 35-26 | 33-24 | 36-27 |
| 2.0% | 30-0 | 25-0 | 23-0 | 26-0 |

In the Proposed Rule, CMS projects that the number of facilities that fall under each payment reduction level each year as shown in **Figure 7**. More facilities are projected to receive payment reductions in PY 2021, but then decrease thereafter. Discern has performed previous analyses that suggest this is not an intentional policy decision, but rather a result of changes in the distribution of facility performance. This year’s projections appear to follow that trend.

Figure 7. Actual and Estimated Distributions of Payments Reductions PY 2018 – PY 2023

| Payment Reduction | PY 2020 Actual¹⁶ | | PY 2021 Projected¹⁷ | | PY 2022 Projected¹⁸ | | PY 2023 Projected¹⁹ | |
|---|------------------------------------|-------|---------------------------------------|-------|---------------------------------------|-------|---------------------------------------|-------|
| | Count | % | Count | % | Count | % | Count | % |
| 0.0% | 4481 | 60.4% | 3,802 | 56.0% | 5,293 | 73.9% | 5,490 | 76.8% |
| 0.5% | 1669 | 22.5% | 1,532 | 22.6% | 1,339 | 18.7% | 1,215 | 17.0% |
| 1.0% | 849 | 11.4% | 896 | 13.2% | 432 | 6.0% | 336 | 4.7% |
| 1.5% | 294 | 4.0% | 359 | 5.3% | 81 | 1.1% | 65 | 0.9% |
| 2.0% | 127 | 1.7% | 188 | 2.8% | 19 | 0.3% | 41 | 0.6% |
| Weighted Average Payment Penalty | 0.32% | | 0.38% | | 0.18% | | 0.16% | |

¹⁵ <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP/Downloads/ESRD-QIP-Summary-Payment-Years-2019-2024.pdf>

¹⁶ <https://data.medicare.gov/data/dialysis-facility-compare>

¹⁷ <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP/Downloads/ESRD-Final-Rule-2019.pdf>

¹⁸ <https://www.govinfo.gov/content/pkg/FR-2019-11-08/pdf/2019-24063.pdf>

¹⁹ <https://www.govinfo.gov/content/pkg/FR-2020-07-13/pdf/2020-14671.pdf>

Given that payment reductions shift based on underlying program performance trends, KCP has previously urged CMS to consider setting payment penalties at specific distribution points. This would create a more predictable model for facilities and CMS, while still incentivizing facilities to maximize their QIP performance.

KCP continues to believe that quality is not relative and that any program that requires public reporting and penalizes providers should reflect the actual quality of care being provided. To that end, KCP reiterates that we would prefer the Total Performance Score (TPS) cut points and the benchmarks and thresholds for attainment and improvement to be based objective goals. We remain concerned that setting a fixed number of facilities in any of the five TPS categories distorts quality and eliminates transparency. It results in a pre-determined number of facilities being labeled as providing poor quality, when in reality there may actually a greater or lesser number of facilities that should fall into the lowest quintile based on their actual performance. If this approach were taken, the results projected by earlier rulemakings should not have occurred. We would like to meet with CMS to discuss specific proposals for resolving this problem.

B. KCP continues to encourage CMS to work with the community and NQF to develop a better approach to the small numbers problem.

Another issue that we ask CMS to address relates to the small number problem. The decision to include facilities with 11 or more cases as the basis for measure applicability instead of the more widely accepted 25 or more cases that commercial insurers and other private quality programs typically apply undermines the statistical reliability of the measure results. We appreciate the work CMS has done on the small facility adjuster, but as Discern Health analyses have repeatedly shown (which we have provided in several of the previous KCP comment letters), the current policy unfortunately does not eliminate the random results associated with small numbers. We encourage CMS to review the work that the NQF has completed in relation to rural areas that identifies ways to developed measures that can be used without small numbers negatively impacting the outcomes reported, as well.²⁰

V. Alignment of ESRD Quality Programs

As a final issue, KCP would like to reiterate our commitment to work with CMS to eliminate the inconsistencies and conflicts that have arisen among the various Medicare ESRD quality programs. In previous comment letters, KCP has suggested a way to align the programs, both in terms of measures and structural scoring issues. We ask again that CMS review these recommendations and work with KCP to strengthen both programs to Dialysis Facility Compare (DFC) and the QIP to achieve the independent goals CMS has identified for each and that would preserve the Congressional intent for the ESRD QIP.

²⁰*Id.* at 6.

Figure 8 below outlines the suggestions of the members of KCP for focusing DFC on meaningful measures that are not used in the ESRD QIP and providing patients with the data about each measure on its website in a way that allows patients to prioritize the measure results they want to see. The ESRD QIP would be a smaller set of meaningful measures that ensure that each measure has substantial weight to avoid any one measure being diluted by the others. Because the Congress mandated that the QIP be a public reporting program, we suggested that CMS shift the star ratings to the QIP TPS scores.

Figure 8: KCP Recommendations for Distributing Measures Across the QIP and DFC

| ESRD QIP Measures | ESRD DFC Measures |
|--|---|
| Standardized hospitalization rate measure (current ratio measure modified to a true risk-standardized rate) | KCQA UFR Measure |
| Standardized readmissions rate measure (current ratio measure modified to a true risk-standardized rate) | KCQA Medication Reconciliation (MedRec) Measure |
| Catheter > 90 Days Clinical Measure | NHSN Healthcare Personnel Influenza Vaccination Reporting Measure |
| Bloodstream infection measure (not the current measures, but one that is valid and reliable and meets other NQF criteria) | Kt/V Dialysis Adequacy Comprehensive Clinical Measure (modified to return to individual dialysis adequacy measures) |
| Patient Experience of Care: In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) Survey Clinical Measure (modified per historic recommendations) | Fistula measures (Current AV measure; future standardized fistula rate) |
| Hgb < 10 g/dL | Clinical Depression Screening and Follow-Up Reporting Measure |
| Serum phosphorous | Standardized Mortality Rate measure (current ratio measure modified to a true risk-standardized rate) |
| Transplant referral measure, including assistance with first visit | Patient Reported Outcome Measure (when developed and endorsed) |

We also would ask that each of these measures be refined based on KCP recommendations for the specific measures. We have also suggested that CMS could align the two programs by ensuring that the DFC and QIP measures have the same specifications and the same scoring mechanism.

We encourage CMS to carefully review these proposals and would welcome the opportunity to identify ways of better aligning the ESRD QIP and DFC so that patients could

The Honorable Seema Verma
August 12, 2020
Page 21 of 39

use both programs for decision-making, but each one would be supportive of the other rather than conflicting as they are today.

V. Conclusion

KCP appreciates the opportunity to provide comments on the Proposed Rule. Kathy Lester, our counsel in Washington, will be in touch to schedule a meeting. However, please feel free to contact her at any time if you have questions about our comments or would like to discuss any of them in further details. She can be reached at klester@lesterhealthlaw.com or 202-534-1773. Thank you again for considering our recommendations.

Sincerely,

A handwritten signature in black ink, appearing to read 'J. Butler', with a long, sweeping horizontal flourish extending to the right.

John Butler
Chairman

Appendix A: KCP Members

Akebia Therapeutics
American Kidney Fund
American Nephrology Nurses' Association
American Renal Associates, Inc.
American Society of Pediatric Nephrology
Amgen
Ardelyx
American Society of Nephrology
AstraZeneca
Atlantic Dialysis
Baxter
BBraun
Cara Therapeutics
Centers for Dialysis Care
DaVita
DialyzeDirect
Dialysis Patient Citizens
Fresenius Medical Care North America
Fresenius Medical Care Renal Therapies Group
Greenfield Health Systems
Kidney Care Council
Nephrology Nursing Certification Commission
National Renal Administrators Association
Renal Physicians Association
Renal Support Network
Rockwell Medical
Rogosin Institute
Satellite Healthcare
U.S. Renal Care
Vertex
Vifor Pharma

Appendix B: Table 1: KCP Measure Summary and Recommendations Analysis

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---|---|
| 1 | <p>0258</p> <p>In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) Survey Administration (clinical measure): Measure assesses patients' self-reported experience of care through percentage of patient responses to multiple testing tools.</p> | <p>Measure Validity CMS' own data show that the ICH-CAHPS response rate is low and continues to drop, and that the increasingly lower response rates threaten the validity of ICH-CAHPS as an accountability measure. The Patient-Reported Outcomes TEP suggested that the low response rate is due to patient fatigue; the manner in which the measure is fielded exhausts patients and discourages them from completing the survey. Understanding the patient's perspective and incorporating it into health care decision-making is critical.</p> <p><i>Recommendation: KCP suggest maintaining the measure as a <u>reporting</u> measure until the response rate is improved. In previous letters, KCP has offered suggestions as to how to address the problem of fatigue by dividing the survey into the three validated section and fielding each one. Then, while a facility is surveyed on the complete tool, any one patient has to complete only a third of the questions.</i></p> <p>Home Dialysis Patients Despite requests from MedPAC and others in the community, the survey does not include home dialysis patients. Given the Administration's strong desire to incentivize home dialysis, having an in-center only tool seems to contradict that position.</p> <p><i>Recommendation: The survey should be revised to include home dialysis patients; NQF endorsement of the new measure should be sought.</i></p> <p>Homeless Patients The survey does not exclude the homeless. Because facilities are not allowed to provide the survey directly to patients, distribution to homeless patients is not possible.</p> <p><i>Recommendation: CMS should exclude the homeless to whom the survey cannot be distributed given that facilities are not allowed to provide it directly to patients.</i></p> <p>Burden Reduction Twice yearly fielding of the survey imposes substantial administrative burden on facilities and contributes to patient "survey-fatigue."</p> <p><i>Recommendation: CMS should field the survey once a year and not twice to reduce burden on facilities and patients.</i></p> <p>Patient Empowerment Facilities do not see and so cannot use survey results to improve care. The fact that facilities never see the survey results and cannot communicate with patients about the results leaves patients feeling unheard.</p> |

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---------------------------|--|
| | | <p><i>Recommendation: CMS should allow facilities to see the results of the surveys so they can respond to the specific patient concerns. Patient members of the TEPs have recommended this step. KCP has consistently recommended extending the survey to include questions related to home dialysis patients. Given the Administration’s Advancing Kidney Care Initiative, CMS should prioritize adding these questions to the survey and seeking NQF endorsement of the new measure.</i></p> |
| 2 | 2496 | <p>Standardized Readmission Ratio (SRR) (clinical measure): Ratio of the number of observed unplanned 30-day hospital readmissions to the number of expected unplanned 30-day readmissions.</p> <p>Overall Reliability CMS data have shown the SRR is not reliable. In the most recent iteration of the measure currently under review at NQF, the overall IUR was 0.35. Statistical literature traditionally interprets a reliability statistic of 0.5-0.6 as “unacceptable”.²¹</p> <p><i>Recommendation: We again recommend CMS implement the measure and/or adjustment to yield a reliability statistic ≥ 0.70, consistent with how NQF bases its evaluation of measures and more generous than the literature.²² This and/or an update to the SFA ranges is necessary to prevent small facilities from having scores highly subject to random variability.</i></p> <p>Reliability Not Stratified by Facility Size Testing data stratified by facility size were not provided for the measure iteration currently under review by NQF because it “is not required.” CMS data from 2009 revealed an IUR of 0.46 for small facilities—i.e., for approximately one-third of all facilities, 54 percent of the score they receive on the 2009 SRR could be attributable to random noise and not signal.</p> <p><i>Recommendation: KCP believes penalizing facilities for performance due to random chance is not appropriate and that it is imperative that CMS provide the most recent reliability results stratified by facility size. Absent that information, we submit that the demonstrably unreliable SRR, as currently specified, is particularly unreliable and unsuitable for use in small facilities. KCP maintains that until it is reliable for all facilities, the SRR should not be used in the ESRD QIP.</i></p> <p>PIUR is Not an Appropriate Measure of Reliability CMS/UM-KECC crafted an additional metric of reliability termed the profile-IUR (PIUR)²³ to “indicate the presence of outliers or heavier tails among the providers, which is not captured in the IUR itself. . . . [When] there are outlier providers, even measures with a low IUR can have a relatively high PIUR and can</p> |

²¹ Adams JL. The Reliability of Provider Profiling: A Tutorial. Santa Monica, California:RAND Corporation. TR-653-NCQA, 2009.

²² Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London: Routledge, p. 13; DeVellis, RF. (2012). *Scale development: Theory and applications*. Los Angeles: Sage. pp. 109–110; Adams, JL. (2009). The reliability of provider profiling. RAND Health.

²³ He K, Dahlerus C, Xia L, Li Y, Kalbfleisch JD. The profile inter-unit reliability. *Biometrics*. 2019 Oct 23. doi: 10.1111/biom.13167. [Epub ahead of print.]

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---------------------------|---|
| | | <p>be very useful for identifying extreme providers.”²⁴ The PIUR for the SRR was PIUR is 0.61, which CMS interprets as demonstrating that “the SRR is effective at detecting outlier facilities and statistically meaningful differences in performance scores across dialysis facilities.”²⁵ In its review of this measure, however, NQF’s Scientific Methods Panel (SMP), none of whom were familiar with the PIUR, disagreed that it is an appropriate measure of reliability for <i>any</i> QIP measure, which are used to distinguish performance between providers falling in the <i>middle</i> of the curve to determine penalties. The SMP concluded that the IUR is and remains the appropriate measure of reliability for this purpose.</p> <p><i>Recommendation: KCP strongly concurs with the NQF’s Scientific Methods Panel (SMP) conclusion that the PIUR is not an appropriate measure of reliability for any QIP measure. QIP measures are used to distinguish performance along a continuum, in particular among providers falling in the middle of the curve to determine penalties; the ability to reliably distinguish outliers for implementation of these measures is not the point. The IUR is and remains the appropriate measure of reliability for measures proposed for the QIP.</i></p> <p>Double Penalties</p> <p>There is unnecessary overlap with the SRR and the Standardized Hospitalization Ratio measure (SHR, NQF 1463), which results in a facility being twice penalized for a readmission occurring within 30 days of the index discharge. In response to stakeholders expressing this concern during NQF’s current review of the most recent iteration of the measures, CMS acknowledged that the same hospitalization event may indeed be counted twice, but believes “this is appropriate because it places additional emphasis on the importance of avoiding hospitalizations and re-hospitalization for dialysis patients . . . [and can] help reduce this major cost driver.”</p> <p><i>Recommendation: While KCP agrees reduction of hospitalizations and readmissions is paramount, we do not believe inflicting specious penalties on dialysis facilities is an appropriate or ethical solution and may ultimately limit access to care. To avoid this “double penalty”, we again ask that CMS include an exclusion in the SHR for hospitalizations that occur within 29 days of the index discharge. Incorporating this exclusion will avoid readmissions being captured as a hospitalization by the SHR, but it will be captured as a readmission by the SRR. This change prevents a facility from being penalized twice for each such readmission.</i></p> <p>Rates vs. Ratios</p> |

²⁴ Kalbfleisch JD, He K, Xia L, Li Y. Does the inter-unit reliability (IUR) measure reliability? *Health Services and Outcomes Research Methodology*. 2018;18(3):215-225. Doi: 10.1007/s10742-018-0185-4.

²⁵ Citation: SHR measures submission materials to NQF.

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---------------------------|--|
| | | <p>The QIP should use a true risk-standardized rate measure; the ratio measure has relatively wide confidence intervals that can lead to facilities being misclassified and their actual performance not being reported. A ratio that is then multiplied by a national median is not a true risk-standardized rate.</p> <p><i>Recommendations: CMS could use the underlying readmission rate and appropriately risk adjust it using race/ethnicity (as is done with the standardized mortality ratio). It should also build off of its contracted work with NQF and develop socio-demographic adjusters, consistent with KCP's 2018 comment letter recommendations. While CMS submits the new measure to the NQF for endorsement, it could use this improved readmissions rate measure in the QIP.</i></p> <p><i>CMS has acknowledged in previous rulemaking that rate measures are more transparent and easier for patients and caregivers to understand. CMS should act quickly to establish a meaningful readmissions measure for the QIP.</i></p> <p>SDS Factors CMS could use the underlying readmission rate and appropriately risk adjust it using race/ethnicity (as is done with the SMR). It should also build off of its contracted work with NQF and develop socio-demographic adjusters, consistent with KCP's 2018 comment letter recommendations. While CMS submits the new measure to the NQF for endorsement, it could use this improved transfusion rate measure as a reporting measure in the QIP.</p> <p><i>Recommendation: CMS should appropriately adjust the underlying transfusion rate using race/ethnicity.</i></p> <p>Burden Reduction Incorporation of a measure with scores known to be highly subject to random variability and double penalizes providers imposes an unnecessary burden on facilities, as well as patients who are interested in understanding the actual performance of facilities and cannot.</p> <p><i>Recommendation: As above, KCP believes ensuring that performance measures addressing this critical clinical topic are fair and reliable is vital and necessary to reduce facility and patient burden and confusion.</i></p> <p>Patient Empowerment Readmissions is an important factor in making health care decisions for patients.</p> <p><i>Recommendation: As above, KCP believes ensuring that performance measures addressing this critical clinical topic are reliable and a valid representation of performance for all facilities is vital and necessary to inform patients in making these weighty decisions.</i></p> |
| 3 | Based on NQF 2979 | <p>Standardized Transfusion Ratio (STrR) (a reporting measure): Dialysis facility</p> <p>Measure Validity</p> |

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---|---|
| | <p>reporting of data on Medicare claims and in CROWNWeb that are used to determine the number of eligible patient years at risk for calculating the risk adjusted facility level transfusion ratio (STrR) for adult Medicare dialysis patients.</p> | <p>The STrR measure lacks validity; KCP is pleased that CMS has acknowledged this concern and supports its decision to change the measure to a reporting metric while reviewing the problem.</p> <p>Insufficient Reliability for Small Facilities The STrR clinical measure has not been demonstrated reliable for small facilities. In the most recent iteration of the measure, currently under review at NQF, the overall IUR for the one-year STrR was 0.63-0.68 across the years 2014-2017. CMS did not provide testing data stratified by facility size to NQF because it “is not required”. Yet data from 2011-2014 for which there was a similar overall IUR revealed values as low as 0.30 for small facilities—that is, for approximately one third of facilities, 70 percent of the score they received on the STrR could be attributable to random noise and not signal. Absent this information for the new clinical measure iteration (currently under review at NQF), we submit that the STrR clinical measure remains unreliable and unsuitable for use in small facilities, and that until it is reliable for all facilities the measure should not be used in the ESRD QIP.</p> <p><i>Recommendation: KCP does not believe that penalizing facilities for performance due to random chance is appropriate and that it is imperative that CMS provide the most recent reliability results stratified by facility size. We again recommend that CMS implement the measure and/or adjustment to yield a reliable result (reliability statistic >=0.70), which is consistent with how the NQF bases its evaluation of measures and more generous than the literature.²⁶ This step is necessary to prevent small facilities from having scores that are highly subject to random variability and/or to update the SFA ranges. Until it is reliable for all facilities, the clinical measure should not be used in the ESRD QIP.</i></p> <p>The STrR is Not an Appropriate Measure of Anemia Management Given that physicians and hospitals, not dialysis facilities, control whether or not a patient receives a transfusion, KCP again recommends shifting away from the STrR to assess anemia management to a more appropriate measure that more directly reflects patient quality of care, is more clearly actionable, and reduces burden. The STrR should be replaced with low hemoglobin (Hgb) measure (e.g., a Hgb <10 g/dL). While it will be necessary to develop updated specifications, exclusions, testing, and business rules, KCP would welcome the opportunity to work with CMS on such a measure; we note that CMS developed a similar measure several years ago that would be an appropriate starting point. A low Hgb measure would reduce burden, because any transfusion measure requires dialysis facilities to chase paperwork created by other providers. It also is a better measure than the STrR because facilities and physicians have access to patient hemoglobin data in the facility, whereas they do not have access to STrR data. Moreover, it is actionable by physicians and will have a direct a positive impact on an issue of critical import to patients.</p> |

²⁶ Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London: Routledge, p. 13; DeVellis, RF. (2012). *Scale development: Theory and applications*. Los Angeles: Sage. pp. 109–110; Adams, JL. (2009). The reliability of provider profiling. RAND Health.

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---------------------------|---|
| | | <p><i>Recommendation: KCP again urges CMS to adopt a more appropriate anemia management measure, such as the Hgb <10 g/dL. We are aware that such a measure is not currently endorsed by NQF, but believe NQF's updated evidence algorithm would provide a path for its consideration anew, and that the Hgb <10 measure, stewarded by CMS, represents a framework to which updated specifications, exclusions, and business rules could be applied. KCP volunteers to work with CMS to develop such a measure. Once an appropriate measure is developed, KCP asks that CMS submit it to NQF for endorsement.</i></p> <p>Rates vs. Ratios The QIP should use true risk-standardized rate measures, because ratio measures have relatively wide confidence intervals that can lead to facilities being misclassified and their actual performance not being reported. A ratio that is then multiplied by a national median is not a true risk-standardized rate.</p> <p><i>Recommendation: The QIP should use true risk-standardized rate measures.</i></p> <p>SDS Factors CMS could use the underlying transfusion rate and appropriately risk adjust it using race/ethnicity (as is done with the SMR). It should also build off of its contracted work with NQF and develop socio-demographic adjusters, consistent with KCP's 2018 comment letter recommendations. While CMS submits the new measure to the NQF for endorsement, it could use this improved transfusion rate measure as a reporting measure in the QIP.</p> <p><i>Recommendation: CMS should appropriately adjust the underlying transfusion rate using race/ethnicity.</i></p> <p>Burden Reduction Shifting to a more appropriate anemia management measure for dialysis facilities would reduce burden, because any transfusion measure (including a rate measure) requires dialysis facilities to chase paperwork created by other providers who also experience the burden on having to provide the data/documentation of providing the transfusion.</p> <p><i>Recommendation: As above, KCP again urges CMS to adopt a more appropriate anemia management measure, such as the Hgb <10 g/dL, to minimize facility burden.</i></p> <p>Patient Empowerment Anemia management is an important factor in making health care decisions for dialysis patients. Transfusions also place patients at risk of becoming ineligible for transplant. CMS has acknowledged in previous rulemaking that rate measures are more transparent and easier for patients and caregivers to understand. CMS should act quickly to establish a meaningful transfusion rate measure for the QIP.</p> <p><i>Recommendation: The QIP should use true risk-standardized rate measures to make the metrics more meaningful to patients.</i></p> |

| | NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|---|--|--|---|
| 4 | NQF endorsed different measure and has rejected the pooled measure | <p>(Kt/V) Dialysis Adequacy Comprehensive (clinical measure): A measure of dialysis adequacy where K is dialyzer clearance, it is dialysis time, and V is total body water volume. Percentage of all patient months for patients whose delivered dose of dialysis (either hemodialysis or peritoneal dialysis) met the specified threshold during the reporting period.</p> | <p>Lack of NQF Endorsement CMS should remove measures that NQF has rejected as part of its endorsement process. Although NQF had endorsed a distinct composite dialysis adequacy measure, the NQF Renal Standing Committee has since reviewed the (Kt/V) Dialysis Adequacy Comprehensive measure and recommended against endorsement.</p> <p><i>Recommendation: CMS should adopt endorsed measures when they are available over measures that have not been endorsed. NQF has endorsed other measures in the domain of dialysis adequacy: NQF #0249 Delivered Dose of HD Above Minimum; NQF #0318 Delivered Dose of PD Above Minimum; NQF #1423 Minimum spKt/V for Pediatric HD Patients; NQF #2704, Minimum Delivered PD Dose; NQF #2706, Pediatric PD Adequacy—Achievement of Target Kt/V.</i></p> <p>Pooled Measure Using a pooled measure approach results in all patients from the four dialysis populations (adult and pediatric/peritoneal and hemodialysis) to be pooled into a single denominator and in scores being calculated as would be done for a single measure. This approach eliminates the ability to determine performance on any specific patient population or dialysis modality.</p> <p>The pooled measure also disincentivizes home dialysis. Home facilities will have lower adequacy scores under the pooled measure, which will make them more likely to be penalized.</p> <p><i>Recommendation: To promote transparency in dialysis performance and the adoption of home dialysis by patients in their facilities, KCP suggests using the distinct adult HD and PD adequacy adult and pediatric measures endorsed by the NQF. KCP volunteers to work with CMS to address the small numbers problem for pediatric facilities and suggests building on the lessons learned from the NQF's rural health project in which small numbers were addressed through other means than pooling measures.</i></p> <p>Burden Reduction The confusion created by pooling the adequacy measures creates an unnecessary burden on facilities, as well as on patients who are interested in understanding the actual performance of facilities and cannot.</p> <p><i>Recommendation: To reduce both facility and patient burden, KCP again urges CMS to replace the pooled Kt/V Comprehensive Measure with the individual NQF-endorsed adequacy measures, as above.</i></p> <p>Patient Empowerment To make informed decisions about modality choice, patients need to understand a facility's actual performance on the different modality types. The pooled measure hides this information from patients.</p> |

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---|--|
| | | <p><i>Recommendation: To facilitate the conveyance of actionable, meaningful information to patients, KCP again urges CMS to replace the pooled Kt/V Comprehensive Measure with the individual NQF-endorsed adequacy measures, as above.</i></p> |
| 5 | 2977 Hemodialysis Vascular Access: Standardized Fistula Rate (clinical measure): Measures the use of an AV fistula as the sole means of vascular access as of the last hemodialysis treatment session of the month. | <p>Rates vs. Ratios The QIP should use true risk-standardized rates because the ratio measures have relatively wide confidence intervals that can lead to facilities being misclassified and their actual performance not being reported. A ratio that is then multiplied by a national median is not a true risk-standardized rate. <i>Recommendation: CMS could use the underlying fistula rate measure. While CMS submits the new measure to the NQF for endorsement, it could use the current measure in the QIP.</i></p> <p>Insurance Status KCP notes CMS may wish to work with the community to determine if insurance status prior to receiving dialysis should be a risk adjuster for this measure. <i>Recommendation: CMS should consider working with the community to determine if insurance status prior to receiving dialysis should be a risk adjuster for this measure.</i></p> <p>Patient Empowerment Vascular access may be the most important measure for patients making decisions about dialysis facilities in the ESRD QIP, with catheter reduction being the most important of the two access measures. CMS has acknowledged in previous rulemaking that rate measures are more transparent and easier for patients and caregivers to understand. CMS should act quickly to make this a rate measure. <i>Recommendation: The QIP should use true risk-standardized rate measures to make the metrics more meaningful to patients.</i></p> <p>Burden Reduction The confusion around the ratio measure and misclassification of facilities create an unnecessary burden on facilities, as well as patients who are interested in understanding the actual performance of facilities and cannot. <i>Recommendation: The QIP should use true risk-standardized rate measures to reduce facility and patient burden and confusion.</i></p> |

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---|---|
| 6 | 2978 Hemodialysis Vascular Access: Long-Term Catheter Rate (clinical measure): Measures the use of a catheter continuously for 3 months or longer as of the last hemodialysis treatment session of the month. | <p>Insurance Status Generally, KCP supports this measure, but notes CMS may wish to work with the community to determine if insurance status prior to receiving dialysis should be a risk adjuster for this measure. <i>Recommendation: CMS should consider working with the community to determine if insurance status prior to receiving dialysis should be a risk adjuster for this measure.</i></p> <p>Patient Empowerment Vascular access may be the most important measure for patients making decisions about dialysis facilities in the ESRD QIP, with catheter reduction being the most important of the two access measures.</p> |
| 7 | Based on 1454,(NQF reserve status); the Measure Applications Partnership (MAP) did not support the measure in its 2016 report | <p>The Measure is “Topped Out” The measure is not used to make clinical decisions and is topped out. <i>Recommendation: CMS should retire the measure.</i></p> <p>Burden Reduction Reporting a measure that has provides neither clinical value nor differentiates among facilities imposes a burden without providing benefit. <i>Recommendation: CMS should retire the Hypercalcemia measure.</i></p> <p>Patient Empowerment Given the topped-out nature of this measure, there is no significant benefit for patients. <i>Recommendation: CMS should retire the Hypercalcemia measure.</i></p> |
| 8 | 1463 Standardized Hospitalization Ratio (SHR) (clinical measure): Risk-adjusted SHR of the number of observed hospitalizations to the number of expected hospitalizations. | <p>Overall Reliability CMS data have shown that the SHR measure is not reliable. In the most recent iteration of the measure currently under review at NQF, the overall IUR for the one-year SHR was 0.53-0.59 for 2015-2018. We note that this value represents a <u>decline</u> from the 2010-2013 IURs (0.7), and that statistical literature traditionally interprets a reliability statistic of 0.50-0.60 as “poor”.²⁷ <i>Recommendation: We again recommended that CMS implement the measure and/or adjustment to yield a reliable result (reliability statistic >=0.70), which is consistent with how the NQF bases its evaluation of</i></p> |

²⁷ Adams JL. The Reliability of Provider Profiling: A Tutorial. Santa Monica, California:RAND Corporation. TR-653-NCQA, 2009.

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---------------------------|---|
| | | <p><i>measures and more generous than the literature.²⁸ This step is necessary to prevent small facilities from having scores that are highly subject to random variability and/or to update the SFA ranges.</i></p> <p>Reliability Not Stratified by Facility Size While testing data stratified by facility size were not provided for the measure iteration currently under review by NQF because it “is not required”, 2010-2013 data revealed an IUR as low as 0.46 for small facilities—that is, for approximately one-third of facilities, 54 percent of the score they received on the SHR could be attributable to random noise and not signal. We believe it's disingenuous, at best, not to provide reliability based on facility size merely because NQF “does not require” it.</p> <p><i>Recommendation: KCP believes penalizing facilities for performance due to random chance is not appropriate and that it is imperative that CMS provide the most recent reliability results stratified by facility size. Absent that information, we submit that the demonstrably unreliable SHR, as currently specified, is particularly unreliable and unsuitable for use in small facilities. Until it is reliable for all facilities, the measure should not be used in the ESRD QIP.</i></p> <p>PIUR is Not an Appropriate Measure of Reliability To assess more directly the value of SHR in identifying facilities with extreme outcomes, CMS and UM-KECC crafted an additional metric of reliability termed the profile-IUR (PIUR).²⁹ Per CMS, “The PIUR indicates the presence of outliers or heavier tails among the providers, which is not captured in the IUR itself. . . . [When] there are outlier providers, even measures with a low IUR can have a relatively high PIUR and can be very useful for identifying extreme providers.”³⁰ The PIUR for the SHR was PIUR is 0.75-0.85 for 2015-2018, which CMS interprets as demonstrating that “the SHR is effective at detecting outlier facilities and statistically meaningful differences in performance scores across dialysis facilities.”³¹ We note that in its review of this measure, however, NQF’s Scientific Methods Panel (SMP), none of whom were familiar with the PIUR, disagreed that it is an appropriate measure of reliability for <i>any</i> QIP measure, which are used to distinguish performance between providers falling in the <i>middle</i> of the curve to determine penalties. The SMP concluded that the IUR is and remains the appropriate measure of reliability for this purpose.</p> <p><i>Recommendation: KCP strongly concurs with the NQF’s Scientific Methods Panel (SMP) conclusion that the PIUR is not an appropriate measure of reliability for any QIP measure. QIP measures are used to distinguish</i></p> |

²⁸ Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London: Routledge, p. 13; DeVellis, RF. (2012). *Scale development: Theory and applications*. Los Angeles: Sage. pp. 109–110; Adams, JL. (2009). The reliability of provider profiling. RAND Health.

²⁹ He K, Dahlerus C, Xia L, Li Y, Kalbfleisch JD. The profile inter-unit reliability. *Biometrics*. 2019 Oct 23. doi: 10.1111/biom.13167. [Epub ahead of print]

³⁰ Kalbfleisch JD, He K, Xia L, Li Y. Does the inter-unit reliability (IUR) measure reliability? *Health Services and Outcomes Research Methodology*. 2018;18(3):215-225. Doi: 10.1007/s10742-018-0185-4.

³¹ Citation: SHR measures submission materials to NQF.

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---|---|
| | | <p><i>performance along a continuum, in particular among providers falling in the middle of the curve to determine penalties; the ability to reliably distinguish outliers for implementation of these measures is not the point. The IUR is and remains the appropriate measure of reliability for measures proposed for the QIP.</i></p> <p>Rates vs. Ratios The QIP should use true risk-standardized rates because the ratio measures have relatively wide confidence intervals that can lead to facilities being misclassified and their actual performance not being reported. A ratio that is then multiplied by a national median is not a true risk-standardized rate. <i>Recommendation: The QIP should use true risk-standardized rate measures.</i></p> <p>SDS Factors CMS could use the underlying hospitalization rate and appropriately risk adjust it using race/ethnicity (as is done with the SMR). It should also build off of its contracted work with NQF and develop socio-demographic adjusters, consistent with KCP’s 2018 comment letter recommendations. While CMS submits the new measure to the NQF for endorsement, it could use this improved hospitalization rate measure in the QIP. <i>Recommendation: CMS should appropriately adjust the underlying hospitalization rate using race/ethnicity.</i></p> <p>Burden Reduction The confusion around the ratio measure and misclassification of facilities create an unnecessary burden on facilities, as well as patients who are interested in understanding the actual performance of facilities and cannot. <i>Recommendation: The QIP should use true risk-standardized rate measures to reduce facility and patient burden and confusion.</i></p> <p>Patient Empowerment Hospitalization rates are critical indicators of quality performance for both patients and providers. The lack of reliability for the SHR means that the measure is not accurately reflecting the performance of small facilities and provides inaccurate information upon which patients are then asked to make health care decisions. <i>Recommendation: The QIP should use true risk-standardized rate measures to make the metrics more meaningful to patients.</i></p> |
| 9 | Based on NQF #0418 Clinical Depression Screening and Follow-Up (reporting measure): Facility reports in CROWNWeb one of six conditions for each | <p>CMS Should Implement Measures as Endorsed by NQF CMS has changed the specifications making the measure different than the one that NQF endorsed. These changes mean that the QIP measure has not been reviewed or endorsed by NQF.</p> |

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---|---|
| | qualifying patient treated during performance period. | <p><i>Recommendation: If it were to remain in the QIP, KCP continues recommending that CMS use it as a reporting measure but encourages CMS to work with the kidney care community to establish a standardized ESRD-specific tool.</i></p> <p>Burden Reduction When CMS changes the specifications of an NQF-endorsed measure, it creates a burden on facilities because they are reporting a measure that may or may not meet measure development criteria, and if it does not, reporting the information does not provide any value. Patients are burdened by having to figure out on their own whether or not the measure is accurately reporting on a facility's performance.</p> <p><i>Recommendation: To reduce both facility and patient burden, KCP again urges CMS to implement only NQF-endorsed measure specifications in the QIP.</i></p> <p>Patient Empowerment: Clinical Depression is an important component in managing patients living with kidney failure. However, this measure is better suited for the Dialysis Facility Compare program so that a facility's performance on the measure is not diluted by other measures, making it difficult for patients to use it to make decisions. CMS has indicated that the purpose of DFC is specific to this task.</p> <p><i>Recommendation: To facilitate patient usability, the Clinical Depression Screening and Follow-Up measure should be limited to use in the Dialysis Facility Compare program.</i></p> |
| 10 | Based on NQF 2701: Avoidance of Utilization of High Ultrafiltration Rate (>13 ml/kg/hour) | <p>Ultrafiltration Rate (reporting measure): Number of months for which a facility reports elements required for ultrafiltration rates for each qualifying patient.</p> <p>Patient-Months Construction KCP appreciates that CMS now concurs with our longstanding position that the NQF-endorsed UFR measure specifications should be used and has revised the specifications to the patient-months construction. As we have previously noted, KCP strongly objected to the change to "facility-months"; the patient-months measure construction was carefully and deliberately selected by KCQA when developing the measure so that patients receiving care at a given facility for fewer than 12 months would still be captured and counted in measure calculations and would contribute to the facility score in accordance with the number of months they received care there. This specific—and intended—construction was supported by the NQF Renal Standing Committee when it endorsed the measure in 2017. The calculation using the patient-months construction now comports with the NQF-endorsed measure and should be used.</p> <p>Based on NQF 2701 KCP also appreciates CMS' explicit notation in this Proposed Rule that the UFR Reporting Measure is based on the Kidney Care Quality Alliance's NQF-endorsed Avoidance of Utilization of High UFR, NQF #2701.</p> <p>Patient Empowerment</p> |

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------------------------|--|---|
| | | <p>KCP continues to believe that fluid management is an important quality area, which is why it funded the Kidney Care Quality Alliance (KCQA) to undertake such measure development. KCP members identified addressing fluid management as the highest priority in KCP’s Strategic Blueprint for Kidney Care Quality.</p> |
| <p>1 1</p> | <p>Based on NQF 1460</p> <p>NHSN Bloodstream Infection (BSI) in Hemodialysis Patients (clinical measure): The Standardized Infection Ratio (SIR) of BSIs will be calculated among patients receiving hemodialysis at outpatient hemodialysis centers.</p> | <p>NHSN Validation Study KCP appreciates CMS’ proposal to reduce the submission requirement for facilities selected to participate in the NHSN validation study from 40 to 20 patient records from any two quarters during the year for the applicable calendar year. We concur that this revised approach will reduce facility burden while maintaining an adequate sample size for the measure validation analysis.</p> <p>The Measure is Not Reliable or Valid The measure is not meeting the rigorous criteria of reliability and validity; as a result, the measure is not reporting accurate data to patients or providers. Research conducted by the CDC (the measure’s developer) and others, including CMS, show that the measure is not valid or reliable. CMS data shows that as many as 60-80 percent of dialysis events may be under-reported with the NHSN BSI measure.³² In a follow-up TEP, CMS and other HHS agency officials indicated that the percentage was slightly lower, but TEP members raised concerns that the percentage remains unacceptably high. In light of these data, it is clear that the measure does not meet the criterion of validity for endorsement. This means that the measure in many instances may incorrectly report that a facility has a low number of blood stream infections when, in fact, the facility has a higher number. Given the understandable importance that patients place on a facility’s ability to manage blood stream infections, a measure that fails to accurately represent the facility’s performance deprives patients of their ability to make informed health care decisions. It also unfairly penalizes facilities that diligently pursue and report the hospital infection data necessary for a full picture of infection rates.</p> <p><i>Recommendation: In the short-term, removing the clinical measure and using the Dialysis Event Reporting Measure alone would let patients know whether a facility is reporting such infections while allowing CMS and the community to fix the problems. In previous comments, KCP has suggested that CMS convert the NHSN BSI measure to a reporting measure while it convenes a TEP to identify the problem with the measure and propose solutions. Once a new measure is specified, CMS should submit it to NQF for endorsement before adopting it as a clinical measure for the ESRD QIP.</i></p> <p>CMS Should Implement Measures as Endorsed by NQF CMS should avoid modifying NQF-endorsed measures when adopting them for the ESRD QIP; the NHSN Bloodstream Infection (BSI) in Hemodialysis Patients is noted to be “based on” NQF 1460 but does not fully comport with the endorsed specifications.</p> |

³²2018 Proposed Rule Display Copy 90.

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|-------------------------------------|---|
| | | <p><i>Recommendation: As described above, CMS should eliminate the NHSN BSI measure and rely upon the NHSN dialysis event reporting measure while CMS convenes a TEP to identify the problems with the BSI measure. Once it has revised the measure, CMS should submit the revised measure [to NQF], which would meet the validity requirements of endorsements, to the NQF.</i></p> <p>Burden Reduction Research suggests that the underreporting identified with this measure may be due to the fact that hospitals, not dialysis facilities, have the requisite data. It is a burden on hospitals to provide the data to facilities and on facilities to chase hospitals for the data. Addressing this problem through a valid measure would reduce unnecessary burden on the hospitals and facilities.</p> <p><i>Recommendation: To minimize facility burdens, KCP again urges CMS to eliminate the NHSN BSI measure and rely upon the NHSN dialysis event reporting measure while CMS explores and identifies the problems with the BSI measure.</i></p> <p>Patient Empowerment This measure topic area is critically important to patients. A measure that incorrectly reports a facility as having a low number of BSI when in fact it does not distorts the care being provided and misleads patients in a way that disrupts their ability to make an informed health care decision.</p> <p><i>Recommendation: KCP again urges CMS to eliminate the NHSN BSI measure and rely upon the NHSN dialysis event reporting measure while CMS explores and identifies the problems with the BSI measure.</i></p> |
| 1 2 | Never submitted for NQF endorsement | <p>NHSN Dialysis Event (reporting measure): Number of months for which facility reports NHSN Dialysis Event data to CDC.</p> <p>CMS Has Not Submitted this Measure to NQF for Endorsement This is inconsistent with the intent of the Congress for CMS to use NQF endorsed measures in the QIP (see SSA § 1881(h)(2)(B)). Without the rigor of endorsement, the reliability and validity of the measure remain uncertain and the specification have been allowed to morph so that there are now several subjectively interpreted signs of infection (<i>e.g.</i>, swelling, redness) included.</p> <p><i>Recommendation: CMS should remove the subjective factors and seek NQF endorsement of the measure.</i></p> <p>Burden Reduction The expansion of the reporting protocol to be highly subjective is extremely burdensome and does not contribute to the measure’s underlying premise—to identify BSIs verified by positive blood cultures. Eliminating the subjective factors would help reduce the burden of this measure.</p> <p><i>Recommendation: CMS should remove the subjective factors specified in the measure.</i></p> <p>Patient Empowerment</p> |

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------------------------|--|--|
| | | <p>It is important to patients and KCP that facilities are appropriately monitoring BSI. However, the information reported should be objective and serve the purpose of identifying patients at risk for BSI so they can receive appropriate treatment. The subjective factors added to the measure specifications last year do not achieve this goal.</p> <p><i>Recommendation: CMS should remove the subjective factors specified in the measure.</i></p> |
| <p>1 3</p> | <p>Rejected by NQF</p> <p>Percentage of Prevalent Patients Waitlisted (PPPW) (clinical measure): Percentage of patients at each dialysis facility who were on the kidney or kidney - pancreas transplant waitlist averaged across patients prevalent on the last day of each month during the performance period.</p> | <p>CMS Should Remove Measures NQF has Rejected from the QIP NQF has rejected the PPPW measure as lacking validity.</p> <p><i>Recommendation: CMS should remove the PPPW from the QIP. KCP stands ready to develop an appropriate transplant-related measure with CMS and others in the kidney care community that meets the endorsement criteria of NQF and the intent of the Congress.</i></p> <p>KCP Does Not Support Attribution to Dialysis Facilities of Successful/Unsuccessful Waitlisting KCP believes that while a referral to a transplant center, initiation of the waitlist evaluation process, or completion of the waitlist evaluation process may be appropriate facility-level measures that could be used in ESRD quality programs, the PPPW is not. Waitlisting per se is a decision made by the transplant center and is beyond a dialysis facility's locus of control. In reviewing these measures, we offer the following comments.</p> <p><i>Recommendation: CMS should remove the PPPW from the QIP. KCP stands ready to develop an appropriate transplant-related measure with CMS and others in the kidney care community that meets the endorsement criteria of NQF and the intent of the Congress.</i></p> <p>Stratification of Reliability Results by Facility Size CMS has provided no stratification of reliability scores by facility size for either measure; we are thus unable to discern how widely reliability varies across the spectrum of facility sizes. We are concerned that the reliability for small facilities might be substantially lower than the overall IURs, as has been the case, for instance, with other CMS standardized ratio measures.</p> <p><i>Recommendation: KCP believes it is incumbent on CMS to demonstrate reliability for all facilities by providing data by facility size.</i></p> <p>Burden Reduction Collecting and submitting data on the PPPW measure when it does not provide an accurate view of dialysis facility quality is a burden without benefit.</p> |

| NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|------------|---|--|
| | | <p><i>Recommendation: CMS should remove the PPPW from the QIP. KCP stands ready to develop an appropriate transplant-related measure with CMS and others in the kidney care community that meets the endorsement criteria of NQF and the intent of the Congress.</i></p> <p>Patient Empowerment Making sure that facilities are doing everything within their scope to promote transplants (<i>e.g.</i>, educating patients about transplant options, protecting patients from infection, referring patients to transplant centers, etc.) is important to patients, the community, and the Administration. However, using a measure that is not accurately reporting on facility action misleads patients and forces them to make health care decisions based on false data.</p> <p><i>Recommendation: CMS should remove the PPPW from the QIP. KCP stands ready to develop an appropriate transplant-related measure with CMS and others in the kidney care community that meets the endorsement criteria of NQF and the intent of the Congress.</i></p> |
| 1 4 | Based on NQF 2988 Medication Reconciliation for Patients Receiving Care at Dialysis Facilities (MedRec) (reporting measure): Percentage of patient-months for which medication reconciliation was performance and documented by an eligible professional. | <p>Patient-Months Construction KCP appreciates that CMS now concurs with our longstanding position that the NQF-endorsed Medication Reconciliation measure specifications should be used and has reverted to the patient-months construction. As we have previously noted, KCP strongly objected to the change to “facility-months”; the measure was deliberately constructed and endorsed using patient-months to address the fact that patients may contribute varying amounts of time to the annual denominator population. The calculation using the patient-months construction now comports with the NQF-endorsed measure and should be used.</p> <p>Discrepancies Between the Published and the Endorsed Specifications Remain CMS has changed the specifications from those that NQF endorsed. Specifically, the QIP revisions delete specific items that must be addressed in the medication reconciliation (<i>e.g.</i>, medication name, dosage, etc.). These changes mean that NQF has not reviewed or endorsed the new measure. <i>Recommendation: KCP supports using the Medication Reconciliation measure in the QIP and asks that CMS uses the specifications as endorsed by the NQF.</i></p> <p>Burden Reduction When CMS changes the specification of an NQF-endorsed measure, it creates a burden on facilities because they are reporting a measure that may or may not meet measure development criteria and, if it does not, reporting information that has questionable value. Patients are burdened by having to figure out on their own whether or not the measure is accurately reporting a facility’s performance.</p> <p><i>Recommendation: KCP supports using the Medication Reconciliation measure in the QIP and asks that CMS uses the specifications as endorsed by the NQF.</i></p> |

| | NQF NUMBER | MEASURE TITLE/DESCRIPTION | KCP CONCERNS AND RECOMMENDATION |
|--|---------------|------------------------------|---|
| | | | <p>Patient Empowerment TEPs have consistently endorsed the adoption of a medication reconciliation measure. To be consistent with CMS' own principles and those of experts like NQF, the measure used should be reliable and valid so that patients can use the information to make informed decisions. Changing the specifications calls the new, revised measure's validity and reliability into question.</p> <p><i>Recommendation: KCP supports using the Medication Reconciliation measure in the QIP and asks that CMS uses the specifications as endorsed by the NQF.</i></p> |