



August 11, 2017

Seema Verma
Administrator
Centers for Medicare & Medicaid Services
Department of Health and Human Services
7500 Security Boulevard
Baltimore, MD 21244

Re: CMS-1674-P: 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

Dear Administrator Verma:

Kidney Care Partners (KCP) appreciates the opportunity to provide comments on the “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” (Proposed Rule). This letter addresses the proposals related to the Request for Information. We have provided our comments on the ESRD Prospective Payment System, payments for Acute Kidney Injury patients, and Quality Incentive Program in separate letters.

KCP appreciates the invitation to provide “ideas for regulatory, subregulatory, policy, practice, and procedural changes to better accomplish” the goals of reducing unnecessary burdens, while increasing quality of care, lowering costs, improving program integrity, and “making the health care system more effective, simple, and accessible.”¹ In the Medicare ESRD Program, we recommend making changes to the ESRD PPS patient-level adjusters, as well as relying upon a parsimonious set of measures for the ESRD QIP to help patients with kidney disease live “Life to the Fullest” and eliminating unnecessary and problematic duplications and overlap among the various ESRD quality programs, in particular the QIP and Five-Star.

¹82 *Fed. Reg.* 31190, 31224 (July 5, 2017).

I. CMS can reduce unnecessary burdens for dialysis facilities, clinicians, and patients by eliminating/revising the ESRD PPS case-mix adjusters.

Generally speaking, CMS successfully implemented the transition from the composite rate payment to the ESRD PPS with few problems. However, there is one glaring exception to this general rule, which has created unnecessary burdens on dialysis facilities, health care professionals, and patients. Despite numerous analyses and recommendations from the kidney care community, as well as recommendations from the Medicare Payment Advisory Commission (MedPAC), CMS continues to rely on contractor data analyses and recommendations for patient-level adjusters for comorbidities, age, and body size/weight that result in an underpayment of \$1.76 per treatment. Additionally, the documentation requirements for the comorbidity adjusters would result in requiring patients to undergo clinically unnecessary medical procedures that create an additional financial burden on the system and patients.

The purpose of patient-level, or case-mix adjusters, when appropriately defined, is to protect beneficiaries by ensuring that all beneficiaries, regardless of the cost of providing care to them, have access to health care services. CMS has also stated that “[t]he purpose of the co-morbidity adjustments is to provide added payment for those co-morbid diseases that result in higher dialysis costs.”² Thus, adjusters counterbalance the incentive inherent in any prospective payment system to treat only the healthiest patients. Unfortunately, the chronic underpayment for the adjusters since the inception of the ESRD PPS is strong evidence that the current set of adjusters does not serve the policy purpose of ensuring that beneficiaries have access to the services for which they are eligible. In fact, it creates an unnecessary burden on dialysis facilities and patients.

In our comment letter for the CY 2017 PPS, KCP recommended ways to address the problems The Moran Company and MedPAC have identified and which we describe in detail in this section once again. Specifically, we encourage the new Administration to work with the kidney care community and establish an interim set of adjusters until the Agency can identify a more appropriate data source, as well as address the problems with the regression analysis, which has remained basically unchanged since it was first used to establish the ESRD PPS in 2010. In determining which adjusters are appropriate, CMS should ask: Does the adjuster center on a patient or facility characteristic that, without additional dollars, would limit access to patients? Additionally for patient-level adjusters, CMS should evaluate the burden imposed by these adjusters both in terms of the clinical relevance and the cost of documenting a patient’s condition to claim the adjuster.

²CMS, “End-Stage Renal Disease Prospective Payment System Final Rule” 75 *Fed. Reg.* 49029, 49098 (Aug. 12, 2010).

To the extent that CMS adopts these recommendations, it will be necessary to adjust the standardization factor used to achieve budget neutrality, so that dollars taken out of the base rate historically to account for the adjusters are re-calculated to reflect the actual frequency and size of any modified adjusters, and returned to the base rate if their value is reduced.

A. The analysis underlying the case-mix adjusters is deeply flawed.

In its most detailed letter on the ESRD PPS case-mix adjusters, MedPAC raised concerns about the methodology the CMS contractor uses to estimate the adjustment factors. Specifically, MedPAC identified several technical problems with the methodology, including:

- “If separately billable services are included in the dependent variable for both regressions, the weights will not accurately distinguish the relative cost or payment addressed by each regression.”
- “Multiplying factors from the facility- and patient-level regressions (with different bases) may diminish the accuracy of the combined factors.”
- “Through various revisions of the model, the empirically-determined lowest-cost reference population for age category variables has shifted from ages 45 to 59 in the proposed rule for the CY 2011 PPS, to ages 60 to 69 in the final rule for the CY 2011 PPS, and now ages 70 to 79 in the proposed rule for the CY 2016 PPS. We would expect the relative cost of dialysis treatment across age categories to remain relatively stable over time, and are concerned that such shifts indicate that the estimated factors are highly sensitive to the model’s specification and that the model lacks robustness. The two-equation approach might contribute to these results.”³

These concerns echo those outlined in the previous KCP letters, including most recently for the CY 2016 ESRD PPS Proposed Rule. In our previous comment letters, KCP has suggested that the problems with the adjusters stem from the fact that: (1) facility cost reports are inappropriate data sources for patient-level adjusters; (2) analysis of cost report data shows that control variables are not valid; and (3) payment variables are not independent of each other and, therefore, result in regression results that are not accurate. This also leads to problems with the standardization factor, which we also have recommended that CMS update using the most current data available.

³MedPAC, “Letter to Acting Administrator Andrew Slavitt Re: CMS-1628-P” 3 (Aug. 6, 2015).

Like KCP, MedPAC recommended that CMS develop the payment adjustment factors using “a single-equation methodology that accounts for variation in the cost of providing the full PPS payment bundle.”⁴ MedPAC noted:

Given the availability of cost data for the full PPS payment bundle, it is no longer necessary to use pre-bundle service categories when developing the adjustment factors. We understand it may not be feasible to develop factors based on a single-equation model for CY 2016, but expect to see such a change in a future revision.⁵

Also, like KCP, MedPAC requested that CMS provide greater transparency in terms of the data provided, asking “for the control variables, from both the facility- and patient-based regressions,” and requesting that CMS “explain the calculation of the weights used to combine factors from each regression.”⁶ This information would allow MedPAC and stakeholders to “assess the validity and understand the relative importance of each aspect of the model.”⁷

Once CMS revises the adjusters, it will also be necessary to update the budget neutrality standardization factor to ensure its accuracy. Given The Moran Company’s analyses (*see* Technical Appendix for the current analysis provided by The Moran Company) of the adjusters and the fact that they are not paid out at expected levels, KCP remains deeply concerned that the budget neutrality adjustments preserve an artificially and inaccurately high standardization factor that reduces the overall program budget below that prescribed by statute as the basis for budget neutrality for the entire system.

B. The documentation requirements for the co-morbid case-mix adjusters are unnecessarily burdensome and not clinically appropriate.

While KCP was pleased that for CY 2016, CMS eliminated two of the original six discretionary⁸ comorbid case-mix adjusters, we remain concerned that the remaining four continue to inflict an unnecessary burden on dialysis facilities, clinicians, and patients, if the facility chooses to claim the adjuster. Given that there

⁴*Id.*

⁵*Id.*

⁶*Id.*

⁷*Id.* at 4.

⁸SSA § 1881(b)(14)(D)(i).

is also no practical clinical need for the adjusters as defined and that the documentation requirements are unnecessarily burdensome, we encourage CMS in the final rule for CY 2018 to eliminate the remaining comorbid case-mix adjusters. Further, facilities report in many cases that the cost and burden of documentation is not worth the benefit in the payment of the adjuster, and so these adjusters are reported at a lower frequency than estimated in CMS's standardization assumptions. CMS takes money out of the base rate based on its estimation of the frequency with which adjusters are expected to be claimed. So it is not paying out the adjuster dollars expected, resulting in underpayment relative to the required level of payment projected by CMS.

In the CY 2016 rulemaking cycle, CMS correctly concluded that the bacterial pneumonia and monoclonal gammopathy case-mix adjusters were unduly burdensome on facilities and patients. However, it failed to recognize that the remaining four comorbid patient-level adjusters also were unnecessarily burdensome because they are not relevant to how patient care is delivered or the cost of that care. In fact, these four comorbid case-mix adjusters do not serve a policy purpose and no patient with these conditions faces an access to care problem.

The documentation requirements for facilities to claim the pericarditis, gastrointestinal tract bleeding with hemorrhage, hereditary hemolytic or sickle cell anemia, and myelodysplastic syndrome create the same overly burdensome requirements that CMS has recognized for bacterial pneumonia and monoclonal gammopathy. The burdens on facilities and patients make the cost of documenting the requirements greater than any potential benefit a facility might receive from claiming the adjuster. The bottom-line is that Medicare should not require and pay for additional tests that impose additional physician visits and cost sharing obligations on patients who are already required to receive dialysis three-times a week.

Pericarditis. Pericarditis is the inflammation or swelling of the thin sac-like membrane surrounding the heart. This condition is one that can occur suddenly and rarely lasts long. Only a small number of patients receiving dialysis experience this condition. While patients may report atypical chest pain and nephrologists may listen to the patient's heart for the pericardial friction rub, nephrologists rarely require patients to incur the cost of having an electrocardiogram before simply treating the condition based upon the patient's report and listening to the patient's heart.

Yet, to document this condition, the Benefits Policy Manual requires "suggestive electrocardiogram changes (*e.g.*, widespread ST segment elevation with reciprocal ST segment depressions and PR depressions) not previously reported" in addition to listening for the pericardial friction rub. Thus, as with the bacterial pneumonia, the documentation requirements are inconsistent with current

diagnostic practices. Because of this fact and the small number of patients with the condition and its short duration, KCP recommends that CMS eliminate the pericarditis comorbid case-mix adjuster and allow facilities to rely upon the outlier policy for those patients who may incur higher costs because of this condition.

Gastrointestinal (GI) Tract Bleeding with Hemorrhage. While GI tract bleeding may account for the greater use of some drugs or biologicals, the cost of meeting the current documentation requirements exceeds any potential benefit that this adjuster might provide. In speaking with nephrologists, it is clear that while many patients may experience GI bleeds, the treatment protocol is to treat to the condition rather than require patients to receive one of the expensive tests the Benefits Policy Manual sets forth. Few dialysis patients obtain an endoscopy, colonoscopy, adionuclide scan, or radionuclide imaging, and/or angiography to confirm the condition. Even if a patient does undergo one of these procedures, it can be difficult to identify the actual clumping of the arteries that cause the bleed. Additionally, once a patient has had one of these procedures, it is unlikely that a nephrologist would order a second or third one simply to confirm what he/she already knows has likely occurred again. Thus, for the same reasons that CMS removed the monoclonal gammopathy comorbid case-mix adjuster (documenting it requires patients to undergo procedures they otherwise would not), CMS should eliminate the GI bleeding comorbid case-mix adjuster. Facilities that experience higher costs related to patients with this condition can instead rely upon the outlier policy.

Hereditary Hemolytic or Sickle Cell Anemia. While recent studies have shown that hereditary hemolytic or sickle cell anemia result in higher ESA utilization, this condition is present in a small percentage of the ESRD population,⁹ and the current documentation requirements do not align with clinical practice. Nephrologists monitor patient hemoglobin levels to determine the dosing of ESAs. While it may be of interest to know definitively whether a patient has hereditary hemolytic or sickle cell anemia, the fact that a patient requires more ESA to maintain target hemoglobin levels is independent of the specific diagnosis. Thus, rather than require patients who do not already know their status to undergo one of the tests outlined in the Benefits Policy Manual, nephrologists focus on managing the patient's anemia. Thus, it becomes extremely difficult, if not impossible, for facilities to meet the documentation requirements for this condition. An approach that is more consistent with clinical practice would be to rely upon the outlier policy to address the higher costs.

⁹Vimal K. Derebail, Eduardo K. Lacson, Jr., Abhijit V. Kshirsagar *et al.*, "Sickle Trait in African-American Hemodialysis Patients and Higher Erythropoiesis-Stimulating Agent Dose" *J. Am. Soc. Neph.* (Jan. 23, 2014)(published online).

Myelodysplastic Syndrome. The extremely few dialysis patients with myelodysplastic syndrome (MDS) are battling not only kidney failure, but also a potentially fatal blood cancer. In caring for these patients, nephrologists focus on keeping them from being uremic. Patients will indicate that they have MDS, but very few oncologists or hospitals are willing to provide dialysis facilities with the documentation of the biopsy necessary to meet the requirements of the Benefits Policy Manual and it would be wasteful to order a second biopsy merely to establish this case-mix adjuster. Given the small number of dialysis patients with MDS and the difficulties in documenting the disease, we recommend that CMS eliminate the comorbid case-mix adjuster and instead rely upon the outlier policy.

It is likely that the flawed methodology identified by MedPAC and The Moran Company resulted in the contractor recommending and CMS adopting these comorbid case-mix adjusters. We are concerned that the regression analysis methodology relies upon variables that are not in fact independent of one another, but are treated as if they are. Without addressing these short-comings, it is not surprising that the contractor found a higher cost associated with conditions that those who work in the community do not see. If facilities do not perceive cost to be higher for these conditions, there is no reason for them to avoid treating patients with these conditions, no threat to access, and no policy purpose served by having a payment adjuster.

Thus, KCP continues to recommend the elimination of these four comorbid case-mix adjusters because they serve no clinical value in the real-world setting of dialysis treatments and impose unnecessary time and cost burdens on patients.

C. Recent modifications to the Age Adjuster should be rescinded and the previous Age Adjuster reinstated.

While the Age Adjuster may statistically predict variation in facility costs, the current adjuster and payment multipliers do not align with clinical experience, making their application overly complicated, inappropriate, inefficient, and unnecessarily burdensome. MedPAC's letter regarding the CY 2016 ESRD PPS raised concerns that the methodology, rather than the cost of caring for patients, drove the establishment of the current adjuster and called for CMS to modify the methodology. "We would expect the relative cost of dialysis treatment across age categories to remain relatively stable over time, and are concerned that such shifts indicate that the estimated factors are highly sensitive to the model's specification and that the model lacks robustness."¹⁰

¹⁰MedPAC, *supra* note 3, at 3.

KCP members corroborated MedPAC's concerns. In discussing the patient characteristic of age with the physician, nurse, and other health care professional organizations within KCP, there is a general belief that older patients have more complications and may require the use of additional labor and resources. However, the methodology currently employed to establish the age adjuster does not direct the dollars to the individual patients who require the most care.

The current Age Adjuster sets the reference group at ages 70-79. The contractor determined, contrary to actual clinical experience, that the 70-79 age group does not incur significantly higher costs than other age groups. Under the current Age Adjuster, facilities receive 7 percent more for patients who are 60-69 years old and 11 percent more for patients 80 years and older. Yet, there is no adjustment for those patients who are 70-79 years old. This result implies that patients 70-79 years old require fewer services and items than those in the other age groups, which is inconsistent with the experience of clinicians caring for dialysis patients. In fact, the choice of the 70-79 group as a reference with no adjustment appears to be a requirement of the statistical method and is not based on actual cost variation associated with age.

Additionally, the Age Adjuster as finalized for CY 2016 and in effect today substantially increased its value. Yet, The Moran Company found no data that validate this 159 percent increase in the average adjuster value. CMS has not provided, and KCP has not been able to identify, any reason for this substantial difference. MedPAC's comments also question why there was such a substantial change when it finds the "dialysis treatment across age categories to remain relatively stable over time."¹¹ CMS's own data show that the age distribution of the dialysis population has changed very little when comparing the 2009 age mix (which was used to finalize the 2011 base rate) to that in 2013. Similarly, there is very little difference in separately billed services by age category in 2013 data, reflecting the decrease in use of ESAs and some other drugs since 2011. There have been no concerns expressed about limited access to dialysis services for any of the age groupings. Thus, it is unclear and in fact appears inappropriate to modify the age adjusters as proposed.

MedPAC and The Moran Company have stated that the problems with the Age Adjuster strongly suggest that there is something wrong with the regression analysis. The Moran Company has raised concerns about using a regression analysis to determine the relationship between age and facility costs, as well as with using cost report data to establish an association of age and composite rate costs. It has found only a very limited association of age with separately billable costs. The result of the use of the Age Adjuster essentially randomizes payment, rather than

¹¹*Id.*

targeting payments to patients with specific characteristics associated with higher costs.

Because the current Age Adjuster inappropriately shifts dollars away from more costly patients to those that are not as costly, this adjuster is unduly burdensome. It also distorts the distribution of funds within a facility that could have negative consequences for the quality of care being provided. Therefore, rather than an adjustment that is based upon inadequate data and a flawed methodology, KCP recommends that CMS work with the kidney care community and suspend the use of the Age Adjuster in the ESRD PPS.

D. The BSA and BMI Adjusters Need To Be Modified.

As CMS looks at simplifying the system and reducing burden, it is also appropriate to address the problems of the discretionary¹² BSI and BMI Adjusters used in the ESRD PPS. The current adjusters represent two different measures for the same patient characteristic and for underweight patients undermining any policy goal. While KCP supports including an adjuster(s) to account for patient underweight and overweight for the ESRD PPS, we remain concerned that the current BMI and BSA adjusters are poorly constructed and do not appropriately direct funds to patients based on the cost of providing services. These adjusters as currently designed are methodologically inappropriate. In discussing the patient characteristic of weight with the physician, nurse, and other health care professional organizations within KCP, there is a general sense that physicians rely more often on the BMI to adjust patient treatments. BSA is important for evaluating overweight patients. It may be possible to restructure a weight related adjuster using a single scale so that underweight and overweight patients are mutually exclusive. However, using cost report data is not an appropriate basis for such adjusters. Thus, the use of these adjusters is unnecessarily burdensome and shifts dollars in a way that is not appropriate clinically or analytically.

BMI and BSA are both different measures for the same patient characteristic. As such, they are highly correlated and should not function as independent variables in a regression analysis because they essentially measure the same thing. Patients who are underweight and qualify for a positive adjuster for low BMI are also subject to a BSA adjuster, which applies to all patients, including those with a low BMI. The BSA adjuster for low BMI patients is negative and offsets almost all of the benefit of the positive low BMI adjuster. Low BSA also correlates with advanced age, so this adjuster also cancels out the positive adjuster for the >80 age group.

¹²SSA § 1881(b)(14)(D)(i).

Clinicians have indicated that underweight and overweight at certain thresholds do require more staff time and, for overweight patients, different supplies or equipment may also be necessary. There is absolutely no evidence that cost varies with a continuous variable like BSA, and there is no reason to have an adjuster for normal weight. Because the BSA adjuster as currently designed applies to all patients, it does not meet the policy goal of recognizing the point at which body size results in higher staffing costs or specialized equipment.

Thus, rather than continue adjusters that cancel each other out, KCP recommends that CMS suspend using the BSA and BMI adjusters. KCP is committed to working with CMS and the contractor to identify a better approach that would allow the weight of patients to be appropriately incorporated into the patient level adjusters.

E. The Low-Volume and Rural Adjusters unnecessarily overlap.

In its review and simplification of the payment system, CMS should also modify the ESRD PPS Low-Volume Adjuster and eliminate the Rural Adjuster because it correlates with the Low-Volume Adjuster. Therefore, KCP continues to propose that CMS instead rely upon a two-tiered Low-Volume Adjuster policy, with the Low-Volume adjuster being the first tier and the second tier applying to facilities with 4,001-6,000 treatments per year.

During the last few year's MedPAC has made a series of recommendations about how adjusters in rural areas could be better targeted. To align with the MedPAC's general recommendations of a more targeted approach, KCP recommends that CMS rely upon a two-tiered Low-Volume Adjuster policy. The current Low-Volume Adjuster would constitute tier one. As an interim step CMS should suspend the use of the Rural Adjuster.

Then, CMS could create a second Low-Volume Adjuster, rather than maintain the less-specific Rural Adjuster that overlaps the Low-Volume Adjuster. As an interim step, CMS should use the dollars allocated to the Rural Adjuster for this second tier Low-Volume Adjuster. This second tier adjuster would apply to facilities with 4,001-6,000 treatments per year. The other Low-Volume adjuster requirements would also apply.

We believe that the 4,001-6,000 range is appropriate based upon an analysis prepared by The Moran Company. While it is not possible to replicate the geographic isolation criteria, The Moran Company was able to perform an analysis of the 2013 cost report data to show the distribution of low treatment volumes in relation to facilities' margins. Based upon this analysis, it is clear that facilities with 4,001-6,000 treatments per year also experience significantly negative margins per

treatment and rural facilities with more than 6,000 treatments generally exhibited normal Medicare margins, making it inappropriate to provide them with a low-volume adjustment. This second tier low-volume adjuster would allow CMS to target the dollars directly to those facilities that struggle because of a smaller patient base.

II. KCP asks that CMS eliminate unnecessarily burdensome policies in the ESRD quality programs.

A. CMS should adopt a limited set of measures that matter and refrain from an unrestrained proliferation of measures.

While KCP strongly advocated for the implementation of the first pay-for-performance/value-based purchasing program in Medicare and still supports it, the program is on a trajectory that has made it overly burdensome and threatens its ability to serve as a catalyst for improving quality of care and lowering costs. The primary problem with its current path as set by the previous administration is that more and more measures are being added to the ESRD QIP in a manner that does not promote or align with a kidney disease-specific strategic approach to improving the life of patients with the disease.

As the previous administration developed the its National Quality Strategy, the KCP took the critical step of developing a quality blueprint for improving quality in the delivery of care to individuals with kidney failure. In “A Strategic Blueprint for Advancing Kidney Care Quality,” KCP identifies the key areas central to making an impact on these goals and recommends focusing on 32 strategic opportunities to do so. KCP’s vision is that the identification of a comprehensive, yet parsimonious, core set of strategic recommendations will help patients with kidney disease live “Life to the Fullest.” KCP believes care can and should be improved to:

- Improve survival;
- Reduce hospitalizations;
- Improve health-related quality of life; and
- Improve patient experience with care.

As the first set of Medicare providers to embrace pay-for-performance/value-based purchasing programs that linked payment to the quality of care being provided, dialysis facilities have been at the forefront of efforts to improve quality and lower overall health care costs. However, as the ESRD QIP has evolved, we have been deeply concerned that the approach taken to incorporate new measures into the program has weakened its goals and threatens its ability to allow patients to live life to the fullest.

The current path of the ESRD QIP seems to be focused on adding more and more measures to the program with no measures being retired, even as measures “top out.” MedPAC and other thought leaders also have raised concerns about diluting the impact of quality programs, especially value-based purchasing programs, by incorporating too many measures. Critically important clinical measures, such as reducing catheters, are competing for percentage points with other measures that have less clinical significance to patients. We appreciate that in previous rulemakings CMS has agreed in the past that it would prioritize measures that track important patient outcomes. Yet, the sheer number of measures and continued proposals to add more measures to the ESRD QIP dilute these important measures inappropriately. We also recognize that CMS has statutory constraints and must address certain measurement areas, even as applicable and available measures may be topped out. Nevertheless, CMS has failed to adopt reasonable suggestions to address these burdens (*e.g.*, adopting other measures, reducing the burden of data collection, or by reducing the weight of contribution of such measures to a facility’s Total Performance Score.) CMS also continues to spend significant funds on developing new measures, in some cases where a National Quality Forum-endorsed measure already exists.

With respect to prioritizing measures, MedPAC has recommended that “[t]he set of measures should be small to minimize the administrative burden on providers and CMS.”¹³ MedPAC also noted that the current trajectory of value-based purchasing programs:

creates an incentive for providers to focus resources on the exact care processes being measured, whether or not those processes address the most pressing quality concerns for that provider. As a result, providers have fewer resources available for crafting their own ways to improve the outcomes of care, such as reducing avoidable hospital admissions, emergency department visits, and readmissions and improving patients’ experience of care.¹⁴

KCP encourages CMS to pause its current measure development efforts and instead meaningfully engage with the entire kidney care community, not simply a small group of TEP members, to identify a small set of core measures that matter. This work could identify measures that could be retired as well.

The ESRD QIP has been successful in many ways in its efforts to achieve its original goals. Yet, it runs the risk of no longer being able to achieve them if we do

¹³MedPAC, *Report to the Congress*, “Chapter 3: Measuring Quality of Care in Medicare” 41 (June 2014).

¹⁴*Id.*

not take the time to heed MedPAC's warnings and work together to create system where the measures included in the ESRD QIP are measures that matter.

While KCP greatly appreciates that CMS continues to apply a standard set of criteria for purposes of evaluating measures for retirement or removal, this approach achieves a different purpose than the one articulated by MedPAC and the kidney care community. To address MedPAC's concern, as well as concerns expressed by the kidney care community, we recommend that as a starting point, CMS: (1) adopt criteria based on that articulated by the National Quality Form (NQF) in its work to assist sectors of the health care industry in prioritizing measures; (2) engage in a public notice-and-comment process for prioritizing measures that uses these criteria and is data-driven; and (3) evaluate candidate measures using a similar process before they are added to any ESRD quality program, especially the ESRD QIP.

- 1. Adopting criteria based on that articulated by the NQF in its work to assist sectors of the health care industry in prioritizing measures.**

The NQF, which the Congress, CMS, and the health care community has relied upon to endorse measures and to make recommendations for measures for federal programs, has recognized that while the National Quality Strategy provides a general framework to assess gaps in quality measures, "a deliberate prioritization of measures is necessary to close critical gaps in measurement." The NQF has focused on five domains in which to prioritize measures for the health care system generally. These criteria also would be useful for prioritizing measures within an area of health care, such as managing and treating kidney failure.

To that end, KCP recommends that CMS adopt the following criteria for prioritizing ESRD quality measures, especially those used within the ESRD QIP.

- The measure is **actionable and improvable** by dialysis facilities.
- There is a significant **performance gap** that if addressed would improve patient outcomes in terms of reduced hospitalizations, improved mortality, improved quality of life, and/or improved patient experience.
- The measure will have a **high impact** on improving patient outcomes by reducing hospitalizations, improving mortality, improving quality of life, and/or improving the patient experience.
- The measure **reflects integrated care (across settings and providers, including transitions).**

- The measure as reported is **easy for patients to understand and interpret.**
- **Measurement of the particular aspect of care has lack of, or low unintended/adverse consequences**, in terms of improving patient outcomes by reducing hospitalizations, improving mortality, improving quality of life, and/or improving the patient experience.
- The measured aspect of care is **meaningful to patient/caregiver and patient-centered.**
- The measure focuses on **outcomes.**
- The **burden of measurement** is low.
- The measure will **drive system-level improvement.**
- The **benefit of measure outweighs burden of collection.**

2. Engaging in a public notice-and-comment process for prioritizing current and future measures that uses these criteria and is data-driven.

Establishing the criteria is the first step of a public process that would evaluate the measures currently in ESRD quality programs, particularly the ESRD QIP, to ensure that the programs provide sufficient flexibility to allow dialysis facilities to prioritize improving the quality of care their patients receive. Because measures for the QIP must be adopted through notice-and-comment rulemaking, we recommend incorporating the criteria evaluation into that process.

For the first iteration of prioritization, it may be useful for CMS to publish an Advance Notice of Proposed Rulemaking and publish the results of that inquiry. Specifically, CMS could set forth the criteria (perhaps as adopted in the final rule published in November of this year) and ask interested stakeholders to rank each existing QIP measure using these criteria, as well as to explain the rationale for their rankings. CMS could then publish these results to promote transparency and allow the community to engage in a meaningful dialogue about which measures, if any, should be removed from the current set. Then in the 2018 rulemaking cycle, CMS could propose such changes and finalize them based on stakeholder input.

For future measure decisions, CMS in the relevant rulemaking cycle could suggest candidate measures to be added to the ESRD QIP and seek comments on their inclusion both in terms of how they should be prioritized using the criteria, as well as how the prioritization of the existing set of measures would change if

candidate measure were added. For example, if a patient-reported outcomes measure (PROM) related to pain management were added to the ESRD QIP, the existing pain management measure collected through facility reports might no longer be a prioritized measure because it may be captured in the PROM.

In addition to including the stakeholder comments, it is important to make sure that any measure meets the criteria established by the NQF for endorsing measures. In particular, any measure that does not meet the scientifically accepted definitions of validity and/or reliability should not be incorporated into any quality program. As it evaluates the current measures and proposes new candidate measures for the ESRD QIP, it should include the data supporting each measure in terms of the NQF endorsement criteria. This step would increase transparency and increase the opportunity for finding common ground.

KCP favors an open and transparent process that includes clear prioritization criteria, which aligns with the NQF's process. There are many ways CMS could undertake this important, yet missing, step, and KCP is open to working with CMS on the process, so long as CMS is inclusive, open, and transparent. We would not, for example, support the sole use of technical expert panels, which limit meaningful discussion to a few hand-selected members and whose discussions are tightly controlled by contractors and CMS and for which summary reports do not capture the full range of discussion and dissent.

The current path of the ESRD QIP can be redirected to ensure that it achieves the goals that the Congress, CMS, and KCP envisioned when it was initially established and implemented. As CMS did in those early days, we ask that it work with KCP directly to solidify the successes of the program and allow it to evolve in a manner that will build on those achievements.

B. Minimum Number of Cases

A second aspect of the current QIP that threatens its ability to achieve its goals is the use of the minimum number of cases established under guidance to implement the HIPAA Privacy Rule for the very different context of the ESRD QIP. By using a number meant to ensure the privacy of individuals, rather than ensure the integrity of the data being reported, CMS has introduced unnecessary randomness into the process of scoring the quality measures. This randomness means that facilities cannot predict how their actions will impact the outcomes and so makes the measures meaningless in terms of improving quality. It also makes the outcome of these measures meaningless to patients because the small number of patients drives the outcome, not the actual care being provided.

While we are pleased that CMS has attempted to address KCP's concerns through the use of the Small Facility Adjuster (SFA) for the various QIP measures,

CMS's own measure testing data demonstrate that results for many facilities with small patient numbers will be driven more by luck than by actual performance, in particular for the standardized ratio measures. CMS uses values that are too low and will result in random volatility that the Small Facility Adjuster cannot offset.

For example, consider the Standardized Transfusion Ratio (STrR) measure. When the STrR measure was considered for NQF endorsement, it was found to have very low reliability, especially for small facilities. The inter-unit reliability¹⁵ (IUR) for facilities with sample sizes below 46 patients was about 0.4, suggesting that 60 percent of inter-facility difference was due to random noise and not underlying performance. IURs increase as a function of sample size. Therefore, smaller samples would be associated with lower IURs. Based on the NQF documentation, one would expect the vast majority of STrR variation to be due to random variation across the 10-21 patient-years at risk that CMS has proposed for the small facility adjustment for STrR. All the small facility adjustment does is raise the scores for small facilities, but it does not offset the substantial effect of random variation for small sample sizes.

To reduce the burden on providers and improve quality of care, we encourage this Administration to adopt consistent criteria for establishment of the minimum data requirements and reliability for the standardized ratio measures. We recommend that CMS set the minimum data requirement for each measure at the sample size at which the IUR reaches 0.70, the value commonly used at NQF, as further noted in Adams.¹⁶ (We further note that this value is lower than what is often cited in the published literature—i.e., a value of 0.80).¹⁷ That is, the minimum sample size would be set at the point where at least 70 percent of the observed result would be driven by actual performance. Anything below that means that too high a proportion of the observed result is simply chance. If CMS opts not to adopt this convention, we recommend that the top end of the SFA range be set at a sample adequate to achieve an IUR of 0.7. At that point, enough of the observed result is likely due to actual performance.

For discussion purposes **only**, we illustrate how adopting an IUR of 0.5 and then a subsequent SFA reliability threshold that yields an IUR of 0.7 would at least

¹⁵ From the NQF Measure Worksheet for STrR: A small IUR (near 0) reveals that most of the variation of the measures between facilities is driven by random noise, indicating the measure would not be a good characterization of the differences among facilities, whereas a large IUR (near 1) indicates that most of the variation between facilities is due to the real difference between facilities.

¹⁶ J.L. Adams, "The reliability of provider profiling: A tutorial." *RAND Health* (2009).

¹⁷ Nunnally J, Bernstein I, *Psychometric Theory*, 3rd ed., New York: McGraw-Hill (1994); Garson GD. Scales and standard measures. *Statnotes: Topics in Multivariate Analysis*. (September 2008); Landis J, Koch G, The measurement of observer agreement for categorical data, 33 *Biometrics* 159-74 (1977).

result in a more fair and meaningful representation of quality for small facilities. Based on the NQF documentation for the STrR measure, for example, we estimate that the minimum sample size required to achieve an IUR of 0.5 is about 50 patient-years.¹⁸ We would further estimate that a sample size of about 75 patient-years would be required to achieve an SFA of 0.7. Under these parameters, the Minimum Data Requirement for the STrR measure would be 50 patient-years at risk and the SFA range would be 50-75 patient-years at risk. Again, we emphasize this is offered only as an example: We are not advocating that an initial IUR of 0.5 be used based on facility size.

CMS should apply similar logic to the SRR and SHR measures to determine the Minimum Data Requirement and SFA range. For illustrative purposes, the NQF documentation for SHR suggests a sample of about 200 patients to achieve an IUR of 0.5 and a sample of about 300 to achieve an IUR of 0.7. However, we again note that in its submission to NQF, CMS expresses these results as a function of the number of patients in a facility while CMS now proposes to set the values for the QIP using patient-years at risk. Finally, for SRR, CMS presented reliability data to NQF for which even for large facilities with >121 patients, the IUR was only 0.61. For SRR implementation, CMA proposes an adjuster of 11-41 index discharges, but this structure is even less transparent and makes it impossible to estimate an appropriate SFA range.

In summary, we urge CMS to adopt clear and transparent criteria for measure reliability to set the range for the Minimum Data Requirements and the SFA, and to update the SFA ranges for the standardized ratio measures accordingly. As we have in that past, KCP recommends that CMS adopt a minimum sample size of 26 patients scoring measures and eliminate the small facility adjustment process. Using sample sizes as small as 11 is inconsistent with measurement best practice and exposes the QIP results to random results that are not fully compensated by the small sample size adjustment. As KCP documents in our comment letter on the ESRD QIP on specific measures, there are ample technical ways in which small facilities can be included, yet random results avoided.

We note that one of CMS's goals for the ESRD QIP is to align with the National Quality Strategy. Principle 6 of the NQS states that, "National standards for health care quality and consistent approaches to measuring quality are essential components of the National Quality Strategy." We do not believe that including small sample sizes and trying to then compensate for their inevitable random effects – a process which is, to our knowledge, unique to the QIP – is consistent with the goals of the NQS, and we urge CMS to raise the minimum sample size.

¹⁸ The NQF documentation only provides IUR figures for ranges of sample sizes, so we cannot calculate a precise sample size threshold for a .5 IUR.

C. Reducing Burdens on Patients

KCP agrees that it is critically important to evaluate patients' experiences when receiving dialysis and continues to support including the ICH CAHPS measure in the ESRD QIP. However, as noted during the recent Patient-Report Outcomes TEP by several members, the ICH CAHPS measure response rate is extremely low, likely due in large part to patient survey fatigue. If CMS remains serious about understanding patient experience and having facilities work to improve it, then the ICH CAHPS measures needs to be adjusted immediately. Until it is, it should not be used as a clinical measure.

First, CMS should address the fatigue problem by modifying the measure to address concerns about the burden on patients and to align the specifications with those that AHRQ relied on when it tested the measure, as well as to ensure the accuracy of its fielding.

We would like to work with CMS to identify ways to address the burden and cost issues associated with administering the survey. In previous letters, we have raised concerns about patients being unable to finish the complete survey because of its length and recommended that CMS divide it into the three sections that were independently tested. Given that the Agency has not yet made this modification, we ask that CMS work with us and the patient organizations to find another alternative that promotes the completion of the survey by patients. Similarly, we have raised concerns about the requirement to administer the survey twice each year. We would like to better understand why administering the survey once each year is inadequate. In fact, the American Institutes for Research/RAND *et al.* have described in detail the difficulties in translating the results from ICH CAHPS into interventions resulting in meaningful improvement when administered more frequently than once a year.¹⁹ We also recommend that CMS coordinate with the Networks to reduce duplication in its administration.

D. Misalignment between Five Star and QIP

Within the Medicare program alone, there are multiple quality initiatives (*e.g.*, ESRD QIP, Dialysis Facility Compare, Fistula First/Catheter Last, CROWNWeb, ICH-CAHPS reporting, and the new ESRD Five Star Rating System), each with its own goals, procedures, data collection systems, and data definitions. In addition, the ESRD Networks also operate their own series of quality initiatives separate and apart from the federal programs. Not only is it confusing and burdensome to facilities, physicians, nurses, and other health care professionals who must provide

¹⁹ See, American Institutes for Research, RAND, Harvard Medical School, Westat, Network 15. Using the CAHPS® In-center Hemodialysis Survey to Improve Quality: Lessons Learned from a Demonstration Project. Rockville, MD: Agency for Healthcare Research and Quality (Dec. 2006).

the data through different IT systems with different business rules, but the multiple programs with different measures and data definitions result in inconsistencies in reporting outcomes and quality performance. The most egregious overlap in these programs is between the ESRD QIP and the ESRD Star Rating program, which MedPAC has recognized as well.

KCP recommends that CMS eliminate the unnecessarily burdensome ESRD Star Rating program because it is duplicating the public reporting requirements of the ESRD QIP is so inconsistent with the ESRD QIP that it leads to confusing publicly reported outcomes for facilities. The ESRD Star Rating program includes quality-related metrics that often look the same, but have different measure micospecifications (*e.g.*, reporting periods) than those used for the ERSD QIP. In addition, the methodology for awarding stars in the ESRD Star Rating program is also inconsistent with the ESRD QIP methodology (which is established by statute) for establishing Total Performance Score (TPS) used to assign the ESRD QIP five tier penalties. These differences create unnecessary burdens on facilities and creates inconsistent publicly reported quality results that patients have indicated are extremely confusing.

Given that the Congress mandated the ESRD QIP, KCP recommends that CMS eliminate the ESRD Star Rating program and instead apply stars using the Total Performance Scores (TPS) required by the ESRD QIP. For example, based on the proposed TPS scoring methodology for PY 2019, stars could be awarded in the ESRD QIP as follows:

Total Performance Score	Reduction	Star
100-61	0%	☆☆☆☆☆
60-51	0.5%	☆☆☆☆
50-41	1.0%	☆☆☆
40-31	1.5%	☆☆
30-0	2.0%	☆

Eliminating this unnecessarily burdensome program would accomplish” the goals of reducing unnecessary burdens, while increasing quality of care, lowering costs, and “making the health care system more effective, simple, and accessible.”

Streamlining these programs would allow for the creation of a standardized data dictionary; alignment of measure specifications; elimination of redundancies among the various programs; consolidation of a single IT system; and, most importantly, consistency in the outcomes being reported.

The alignment of ESRD QIP and ESRD Star Rating by the elimination of the different measures and methodology in the ESRD Star Rating program would likely

improve patient care as well. Good data drives good decisions. The current fragmentation produces inconsistent results upon which physicians and others are unlikely to rely. If they do rely upon them, the decisions they make will be suboptimal because the flaws within the various programs. If facilities and physicians can trust the information being reported because the focus is on a single set of valid and reliable measures that are actionable, rather than on competing public reporting programs, the Medicare quality programs would be more likely to accomplish their goals of improving patient outcomes and reducing unnecessary hospitalizations.

Patients who want good quality information also find the various programs confusing and difficult to use as they make treatment decisions. As we have commented in previous letters, the launch of ESRD Star Rating has created significant confusion for patients who are confronted with two very different sets of information. First, they will see their facilities' QIP scores, which may show a zero reduction. Then, if they look at ESRD Star Rating, they will see entirely different results.

We understand that there may be concerns that the ESRD QIP does not provide sufficient differentiation among facilities; however, that should not be the goal of a quality program. First, it is important to set quality goals/benchmarks and acknowledge when providers achieve them, as the ESRD QIP does. Second, it is important to work with the community to establish the clinically metrics that when reported provide an assessment of facility performance and can lead to changes in facility actions to improve those outcomes. Simply creating another program that relies upon a forced distribution to create differentiation that does not exist in reality, which is the methodology adopted by the ESRD Star Rating program, is misleading.

In sum, KCP strongly encourages CMS to take a step back and re-evaluate its approach to the ESRD Star Rating program and eliminate this unnecessarily burdensome requirement. KCP and others in the kidney care community, as well as other interested stakeholders, would welcome the opportunity to have a constructive and meaningful dialogue to establish a system that achieves the intent of those who sought to create it.

Administrator Verma

August 11, 2017

Page 21 of 33

III. Conclusion

KCP appreciates the opportunity to provide comments on the RFI. We look forward to working with the Secretary and Administrator on addressing the concerns in this letter. We would welcome the opportunity to meet to discuss some of how we can help the Administration achieve the critically important goals outlined in the RFI. Please do not hesitate to contact Kathy Lester at (202) 534-1773 or klester@lesterhealthlaw.com if you have any questions.

Sincerely,

A handwritten signature in black ink that reads "Frank Maddux, M.D." with a stylized flourish at the end.

Frank Maddux, M.D.
Chairman
Kidney Care Partners

Appendix A: KCP Members

AbbVie
Akebia Therapeutics, Inc
American Kidney Fund
American Nephrology Nurses' Association
American Renal Associates, Inc.
American Society of Nephrology
American Society of Pediatric Nephrology
Amgen
AstraZeneca
Baxter
Board of Nephrology Examiners and Technology
Centers for Dialysis Care
DaVita Healthcare Partners Inc.
Dialysis Clinic, Inc.
Dialysis Patient Citizens
Fresenius Medical Care North America
Fresenius Medicare Care Renal Therapies Group
Greenfield Health Systems
Keryx Biopharmaceuticals, Inc.
Kidney Care Council
National Kidney Foundation
National Renal Administrators Association
Northwest Kidney Centers
Nephrology Nursing Certification Commission
NxStage Medical, Inc.
Renal Physicians Association
Renal Support Network
Rogosi Institute
Sanofi
Satellite Health Care
U.S. Renal Care

**Technical Appendix Prepared by
Rachel Feldman, Partner, The Moran Company**

- **Recommendation:** Eliminate Comorbid Case Mix Adjusters, particularly given burden on patients and providers—burden offsets value of adjusters—no policy objective achieved with continuation of these adjusters.
- **Recommendation:** Suspend Age and Body Case Mix Adjusters—analysis shows there are incorrectly constructed, distort payments, and do not allocate payment in a manner that improves accuracy of payment or protects access to care—no policy objective achieved with these adjusters in current form—Kidney community recognizes association of age and body size with cost and agrees to work with CMS to develop valid basis for consideration of these adjusters.
- **Recommendation:** Revise Low Volume & Rural Adjusters to Two Tiered Low Volume Adjuster with Elimination of Rural Adjuster.
- **Recommendation:** CMS should seek an objective second opinion on contractor work used to make major revisions to payment, when serious criticism of methodology is documented by responsible stakeholders and MedPAC.

Overview: Statistical methods and data must fit to produce valid results. Research underlying payment adjuster values in the ESRD PPS is seriously flawed. The critique that follows provides an independent objective analysis of the methods used to produce adjuster values in the 2016 ESRD proposed and final rules commissioned by the Kidney Care Community and developed by researchers at The Moran Company. Many of these points were also made by MedPAC in its comments on the 2016 proposed rule. CMS relied upon contractor review of technical criticism and dismissed any serious concerns about its methods, claiming in general that “nothing is perfect”, and relying on a variety of “we continue to believe” statements without any serious responses to serious and responsible criticism of statistical methods and underlying assumptions.

Payment Adjusters Should Serve a Policy Function

The purpose for development of payment adjusters in a Medicare PPS is to target additional payment to providers incurring higher cost due to either facility conditions or patient characteristics. When adjusters target higher payment to patient characteristics, this needs to be done based on reliable data documenting that higher costs are consistently incurred by providers treating patients with the target characteristic. When adjusters target higher payment to facilities, this needs to be done based on reliable data demonstrating that facility characteristics are the source of higher costs that disadvantage those facilities in the payment system, and that those characteristics are outside the control of the facility. The over-arching purpose of payment adjusters is to ensure patient access to care.

Critique of Methods: Adjuster Values Cannot be Accurate

In the development of payment adjusters for the End Stage Renal Disease Prospective Payment System (ESRD-PPS), CMS and its contractor used statistical methodologies and identified adjuster variables in a manner that cannot produce valid or reliable adjuster values, and does not serve the above policy purposes. The discussion that follows explains why the chosen methodology and its implementation cannot produce valid adjusters. Statistical methods are only valid if the data to which they are applied are a fit to the methods. Statistical methods applied to data that do not meet the requirements for reliability and validity will produce results that are not accurate, may not be meaningful, and can be volatile from year to year: the resulting adjusters will target funds in ways that have no relationship to the policy that the methodology is designed to implement. Fundamental requirements of a regression model were not met in the analyses used to design the ESRD-PPS payment adjusters.

To produce valid and reliable results, a regression analysis must be based on a sound research design and must adequately address the assumptions made by the mathematical properties of the regression analysis. The major assumptions we will discuss are noted below and are commonly documented in many standard texts on regression methods.

1. Regression assumes that the model is correctly specified. If a regression model is not correctly specified, the results will be biased, and will not reflect an accurate impact of the predictor variables (also called independent variables) on the dependent variable. Correct specification requires:
 - a. All variables that could predict change in the dependent variable (cost per treatment first equation, cost of separately billed items in second equation) were included in the regression model. The process for selecting variables and evaluating them for inclusion in the regression model was not comprehensive and there is considerable reason to believe that the variables that were selected were not those that drive cost variation.
 - b. All variables must be statistically or theoretically related to the dependent variable in the regression model.
 - c. The coefficients of the predictor variables (the value assigned to the adjuster as a result of the regression) are assumed to not change during the period of analysis.
2. The observations, in this case treatments, are uncorrelated with each other. This means that all treatments are assumed to be independent of each other. In this context, treatments occur in a sequence linked to an individual patient such that treatment cost may for one treatment may be related to prior treatment, the duration between treatments, events that interrupt

treatments such as hospitalization, and the patient's health status at the time of treatment. Therefore, treatments are not independent of each other.

- a. If ordinary least squares (OLS) is used, the result will be that it is no longer possible to trust significance tests.
 - b. If observations are, in fact, correlated as is the case with dialysis treatments, then this correlation between observations should be modeled in the regression using generalized least squares (GLS). This would require modeling the variation within the series of treatments for each patient. GLS could not be done using cost report data as these data have no link to the patient. GLS could be done in the separately billed data. We found no documentation to suggest that this method was used.
3. Regression assumes that there is not random error built into the predictor or independent variables.
- a. Because regression assumes that the data contains no erroneous values, the data is modeled as if all values are correct. If data has a great deal of error contained it, the result is that the coefficients of the predictor variables will be biased and will not reflect the effect of the predictor variable on the dependent variable. There is considerable error in the cost report data used. The separately billed data may meet the conditions for this assumption.
4. The predictor variables are not correlated with each other.
- a. Correlation of predictor variables will increase the variability associated with a coefficient, and will reduce the accuracy of the adjuster. This will increase the degree to which small changes in the data can create large changes in the results. It will also make the interpretation of correlated variables unclear. We find that there is considerable correlation among the predictor variables.

In the discussion that follows, we demonstrate that the regression methodology used, to the extent that it was last explained in the original 2008 Report to Congress, or in the 2016 Proposed Rule, includes multiple violations of these core assumptions.

In addition to violation of the core requirements of the regression methodology, we assert the following problems with the application of the methodology:

- Average facility cost per treatment based on cost report data (excluding drug cost) has no relationship to patient characteristics. It is not possible to distinguish variation in cost per treatment that is associated with patient

- characteristics from facility treatment volume and a number of corporate financial, clinical, purchasing and labor policies that dominate the ultimate cost per treatment at the facility level. Reliance on cost report data to construct adjusters for patient characteristics is inappropriate and results in an invalid model.
- Because the adjuster variables explain less than 10% of the variation in cost, the model should have been re-evaluated before being used.
 - The selection and omission of variables in the regression model was not transparent. Variables known to affect cost were omitted and variables of questionable importance were included.
 - Exaggerated statistical significance of variables based on a universe, not a sample, has resulted in adjusters with questionable statistical or clinical significance.
 - The combination of the coefficients from the two regressions into a single adjuster is problematic. The weighting is not described, but it would not be correct to assume that the distributions for the two regressions are the same. If the distributions are not the same (which we believe is certainly the case), then the accuracy of the resulting adjuster will be compromised. Also there are costs that are duplicated in both regressions, particularly for laboratory tests, further muddying the analysis and eroding accuracy.
 - Because of the poor fit of the model to appropriate data, and the high level of correlation among the adjuster variables, we do not believe that this regression model can be fixed.

Variation in average facility cost per treatment derived from cost reports does not represent variation in the cost associated with treating patients with different characteristics, such as those CMS identified for adjusters.

An analysis must use data that measures the concept under analysis. The relevance and construction of the facility level variables used by CMS have not been explained, nor has CMS explained whether other variables were considered. The only facility level variable that we agree may be linked to variation in average facility cost per treatment is treatment volume.

We disagree, and have data to demonstrate, that size of chain and non-profit/for-profit status have no bearing on variation in treatment cost. In our analysis of cost report data, we find that the following characteristics of owner entities drive most of the variation in treatment cost:

- Overall financial policies
- Labor policies which may vary geographically
- Allocation of overhead to facilities

- Purchasing policies which may vary for sub-sets of facilities
- Contracting practices

Even at the local facility level, our understanding from company leaders is that variations in staffing intensity, the source for cost variation in cost report data, are not related to patient characteristics, but that more time intensive patient care is, for the most part, integrated into the standard scheduling and operation of the clinics. In any case, the aggregate data in cost reports is not in any way directly associated with variation in patient characteristics. Because of this, the variable concepts for adjusters cannot be measured by these data.

The specific regression models used have negligible power to explain variation in cost, and therefore should not have been used to estimate adjusters.

The results from the regressions reported in the 2008 Report to Congress and referred to in the 2016 proposed rules indicate that the adjuster variables explain little of the variation in cost – less than 2% of the total variation in composite rate costs, and less than 8% of separately billable costs. This level of explanatory power of the adjuster variables is very low, only marginally greater than zero. Variables with such low impact on cost may not have been included if the volume of claims and of facilities had not been so large as to result in statistically significant results. Because a population is used rather than a sample, we believe that more restrictive requirements should be used for determination of statistical significance.

Model Specification and Variable Selection

Inclusion of relevant variables and exclusion of irrelevant variables is vital to producing a valid model. In each of the proposed and final rules published since the publication of the 2011 proposed rule, CMS has stated that a stepwise regression was used to arrive at the adjusters. Stepwise regression is a method of adding variables to a regression equation in turn, based on statistical significance or in some predetermined order of entry. The criteria used for entry and exclusion of variables into the regression models used to create the adjusters was not explained, and it is unclear why some variables that may affect cost were excluded from consideration, while other variables were included. The issue of variable selection is magnified by the difficulties of interpreting statistical significance in data sets as large as those used. Because statistical significance alone is not a dependable method of variable selection, the judgment of the researchers becomes paramount in the selection of variables.²⁰

²⁰ Kennedy, P. *A Guide to Economics 1st Edition*, MIT press.

In discussions with ESRD facility operators and clinicians, alternative variables are viewed as clinically significant and potentially relevant to the variation in the costs of delivering dialysis services. In CMS's discussion, age and body size are referred to as useful variables, largely because they are easy to record reliably in claims. It is true that clinicians would agree that body size, specifically severe underweight and significant overweight may be related to the level of effort in caring for patients, and therefore costs. However, clinicians are not suggesting that any continuous variable for body size that provides a payment adjustment for all patients is appropriate. Similarly, body size is actually a proxy for the reason for expecting more intensive care needs, and, as a proxy for underweight and overweight, the referents are different. In the case of underweight patients, the needs of the patients are associated with malnourishment and frailty and the care needed to support stability in the patient's health. Underweight patients will not use more separately billed services than overweight patients, nor will they use more dialysis time. Overweight patients may require more dialysis time, use more drugs due to dosing related to body size, and, for the morbidly obese, may require the facility to purchase special equipment such as lifts and larger sized dialysis chairs.

In the case of age, clinicians also view age as a proxy for factors that increase the cost of care and do not suggest that a continuous or tiered age variable is appropriate. Their concerns for patients with advanced age is not based strictly on age, but the correlation between age and certain characteristics that limit the patient's ability to participate in his or her own care (e.g., nursing home residency, non-ambulatory, paralysis, dementia). They note that it is common for dialysis patients, due to their generally compromised health, to acquire these characteristics before the age threshold of 80.

It does not appear that the development of the model was based on input from dialysis providers who are most aware of those factors that influence the cost of dialysis treatment. In the 2008 analysis, the research team appeared to generate a list of conditions that could potentially be related to dialysis costs, and using many years of historical claims, performed statistical significance tests (which we note were not likely meaningful due to their application to a large population). In response to criticism of the list of comorbidity adjusters in the proposed 2011 rule, CMS reduced the list considerably, and agreed with the critics that the conditions could not be determined to meet the criteria for independence of variables. The response in the final rule was to allow for only one—with the highest value--adjuster to be paid if the patient qualified for multiple comorbidity adjusters. This action in no way corrected the fundamental flaw in the methodology that made the adjuster values invalid.

Large number of ESRD claims and facilities exaggerate statistical significance.

The large number of facilities (most of the universe) and treatments used in the two regressions have resulted in exaggerated statistical significance of coefficients. This

is because coefficients become more statistically significant as the size of a “sample” increases. Statistical significance is most useful to evaluate selection of variables when **actual samples** are being used. In these regressions, CMS uses as much of the universe as it can, rather than having statistically sampled the universe. As a result, statistical significance as used by CMS no longer has the meaning it does with actual samples. In the 2008 report to congress, the authors note this problem, stating:

...given the very large number of ESRD patients with Medicare claims, statistical significance is a necessary but not a sufficient condition for including a potential patient characteristic as a case-mix adjuster. Even variables with very small relationships to costs or payments are likely to be statistically significant in patient-level analyses.²¹

UM-KECC, despite correctly identifying the problem of large sample size and statistical significance in its 2008 report, and repeated in the 2016 revision of adjustres, employs variables with indeterminate statistical significance. A categorized age variable is consistently used as an independent variable in these regressions. In the report to Congress, CMS provides detail on statistical significance of the regression coefficients associated with these variables. In the 2008 report, the age categories 45 to 59 and 70 to 79 were not significant at the .05 level, and in the case of separately billable services, the category 18 to 44 was not statistically significant at the .05 level. Only pediatric patients and patients over 80 were consistently statistically significant in these analyses. Given the large sample size, if age were an independent driver of cost, we would expect a greater level of significance. Note that none of these specifications were disclosed for the updated regressions used to estimate the proposed 2016 payment adjusters.

There is no stated clinical or economic rationale for the selection of control variables

Given the lack of information that statistical significance provides in large populations such as these, the model specification must be made on the basis of clinical or economic theoretical criteria. Of the seven control variables, no link in the dialysis literature, and no analyses confirm a link between the variables and the actual costs of dialysis. There is no theoretical argument for the use of these control variables that might provide a case for assuming a link to cost.

It is not acceptable to omit variables from the regression model that are related to costs.

While a decision to not create an adjuster from a regression coefficient can be defended on policy grounds, the exclusion of an important variable from the

²¹ A Design for a Bundled End Stage Renal Disease Prospective Payment System: Report to Congress 2008, p 27.

regression analysis cannot. The first assumption of regression modeling is that all relevant predictor variables are included in the model and irrelevant ones excluded²². If a variable that has an effect on the cost of treatment is not included in the equation, and this variable is not perfectly uncorrelated with the remaining variables, the coefficients for the remaining variables will reflect the variation that they have in common with the omitted variable. Race of the patient, while not an adjuster, has been demonstrated to effect cost, for example. Other examples of variables that may affect cost, as identified by clinicians, include patient characteristics that interfere with the patient's ability to participate in his/her own treatment (e.g., nursing facility residents, paralysis, dementia). The category of age over 80, is essentially a proxy for these elements as discussed above. Because CMS has not reported the full results of the regression equations used to derive the adjusters for the 2016 proposed ESRD rule, it is impossible to assess the degree of misspecification of the regression model.

Non-independence of predictor variables results in imprecise parameter estimates^{23,24}.

In the regressions on composite rate costs and separately billable services on age, body mass index (BMI), body surface area (BSA), comorbidities, and new patient status, that were conducted by UM-KECC and reported to Congress in 2008, the level of correlation between the independent variables is severe. Body surface area and BMI are both estimated using height and weight, and are by definition, highly correlated, and therefore, not independent of each other. Older dialysis patients tend to have smaller BMI/BSA and lower rates of obesity than younger dialysis patients²⁵. "New to dialysis" correlates with all other variables. There are three consequences of these correlations that we believe make these adjusters unreliable and unsuited for use in a reimbursement system.

Adjusters derived from correlated "independent" variables are imprecise.

Significant correlation between predictor variables will force the regression to only use variation unique to a variable, ignoring the variation that is correlated with another predictor variable. This results in variables with less precise adjuster values. Because the value of the adjuster is based on only the variation of the

²² Kennedy, *A Guide to Economics 1st Edition*, MIT press, p. 57.

²³Kennedy, *A Guide to Economics 1st Edition*, MIT press, p 129.

²⁴ Neter, J. and Wasserman W. *Applied Statistical Models Regression, Analysis of Variance, and Experimental Designs* 1974 Richard D. Irwin Press Homewood, Illinois, p. 344

²⁵ Kramer, H., Saranathan, A., Luke, A., et al. Increasing Body Mass Index and Obesity in the Incident ESRD Population *J Am Soc Nephrol* 17: 1453-1459

“independent” variable unique to that variable, confidence intervals around that adjuster, the range of possible values, is larger than it would be if the variable were truly independent of the other variables.

The adjusters derived from correlated variables no longer have meaning.

A regression coefficient, and the derived adjuster, represent the change required in an independent variable (e.g. age) to create a change in the dependent variable (e.g., composite rate services), holding all other independent variables constant. However, when two predictor variables are highly correlated, as are BSA and BMI, it makes no sense to state that a decrease in BSA, holding BMI constant, creates a change in cost. Because older patients tend to have smaller body size and younger patients larger body size, it makes little sense to model age and body size as if they are independent of each other. Therefore it is quite possible that the significant results obtained for age may be related to body size or some other un-specified variable.

Small changes in the data can result in large changes in the adjusters.

When predictor variables are correlated, even small changes in the data can result in large changes in the adjusters. There has been no demonstrated shift in either the distribution of the age of dialysis patients, nor would the effect that age has on cost over the past ten years have changed. The shift in the age category associated with the lowest cost from 40 to 59 in the proposed 2011 rule, to 60 to 69 in the final 2011 rule, and now to 70 to 79 in the proposed 2016 rule, is an indication that the age variable, highly correlated with other variables included in the regression, is unreliable, and is sensitive to small changes in the data.

The problem of correlated independent variables cannot be solved by omission of a variable

Because it is unacceptable to omit variables that are related to the dependent variable from the regression equation, any attempt to repair this regression by omitting some of the correlated variables will create new problems. The remaining coefficients and resultant adjusters will be biased, and will no longer reflect an accurate effect that the remaining variable has on cost.

Conclusion

The regression analyses used by CMS for the proposed and final 2016 ESRD PPS Rule violates the core assumptions for a valid analysis. We have demonstrated that:

- The regression analyses used to produce the adjuster values included in the ESRD PPS are not correctly specified: there is inadequate rationale provided to explain the selection of the variables and no discussion of other variables that clinicians would suggest better explain variation. Based on our research and expert knowledge of those delivering services, the selection of variables is certainly incomplete.
- The assumption regarding the independence of observations is not anywhere discussed by CMS or its contractor. The unit of observation (the dialysis treatment) is not independent: treatments are linked to individual patients, and one treatment and its costs may very much depend on the results of the last treatment and the interval between treatments. We see no evidence that the required statistical modeling was used to account for this lack of independence among observations.
- Our analysis of cost reporting documents that there is a large amount of error in the data that is unrelated to actual variation in cost. There are large amounts of missing data in the fields that are rolled up into the total cost field used by the researchers. CMS has not disclosed how it handled trimming data for unbelievable values and other types of error. We know that hospital cost reports are frequently highly inconsistent with independent facility cost reports, and are often missing, or have large amounts of missing data. Without addressing the known significant level of error in the data source, the assumption that the data are error free is violated.
- Virtually all of the patient level predictor variables are correlated, with different levels of correlation between pairs of variables:
 - BSA and BMI measure exactly the same characteristics, are entirely correlated and their adjusters cancel each other out.
 - Age is correlated with both BSA and BMI, particularly the lower body size and advanced age. The BSA adjuster cancels out part of the age adjuster for patients over 70.
 - Age is correlated with new patient status, as new patients are more likely to be younger, though patients start dialysis at all ages.
 - New patient coefficients are based on exactly the same data as the other patient characteristics, thereby duplicating cost variation and muddying the cost variation for other coefficients.
- Low volume and rural status are correlated, as well, as acknowledged in the proposed rule by CMS.

Therefore, the adjuster values that are generated by this model are not valid or accurate.

Furthermore, the cost report data upon which the analyses rely represents a poor fit to the task of explaining variation in cost associated with patient level

characteristics. When the meaning of “higher cost care” is reflected in more intense staff time, a more expert mix of staff time, longer duration of dialysis, or special equipment, then these are the sources of data that should be used to develop adjusters, not total facility average cost without drugs. The costs of facility operations for capital, overhead, contracting, labor and benefit policies, and purchasing, are driven by corporate policies that cannot be accounted for by any of the variables included in the regression analysis. There may well be a better way to design a model to produce adjuster values that are valid, once an acceptable model can be specified. The Kidney Care Partners (KCP) and The Kidney Care Council (KCC) are available to assist CMS in formulating a statistically valid approach to developing payment adjusters.

Finally, the payment adjusters are applied to each year’s update to the base rate using original and out-of-date research as the basis for calculating that dollars that are removed from the unadjusted original base rate, then updated by ESRD update factors. This is referred to as “standardization”. CMS has consistently overstated the “standardization” factor, removing more dollars from the unadjusted base rate than are ever paid out for adjusters, because providers do not claim adjusters at the rates originally assumed by CMS. CMS should be adjusting the standardization factor each year to the actual rate at which providers claim adjusters. Not doing so is reducing payments below the level mandated by the original statute.