KIDNEY CARE PARTNERS

June 8, 2020

National Quality Forum
1099 14th Street NW
Suite 500
Washington DC 20005

RE: NQF All-Cause Admissions and Readmissions Project, Spring 2020 Cycle

Kidney Care Partners (KCP) appreciates the opportunity to comment on the measures under consideration for endorsement in the National Quality Forum's (NQF) All-Cause Admissions and Readmissions Project, Spring 2020 Cycle. KCP is a coalition of more than 30 organizations, comprised of patient advocates, dialysis professionals, care providers, researchers, and manufacturers, dedicated to working together to improve quality of care for individuals with Chronic Kidney Disease (CKD) and End-Stage Renal Disease (ESRD). This letter addresses the two new measures submitted for review within the project, the *Standardized Emergency Department Encounter Ratio (SEDR) for Dialysis Facilities* (NQF 3565) and the *Standardized Ratio for Emergency Department Encounters Occurring within 30 Days of Hospital Discharge (ED30) for Dialysis Facilities* (NQF 3566).

I. Overarching Concerns

KCP recognizes the importance of assessing emergency department (ED) utilization by individuals with ESRD. Nevertheless, we have numerous concerns about the proposed *Standardized Ratio for ED Encounters Occurring within 30 Days of Hospital Discharge* (ED30) and *Standardized ED Encounter Ratio for Dialysis Facilities* (SEDR) metrics. We believe the measures as currently specified will not improve the quality of care or outcomes for dialysis patients—and may in fact exacerbate existing sociodemographic status (SDS) and geographic disparities. Below we detail several overarching concerns and make several recommendations applicable to both metrics; concerns specific to the individual measures are then addressed.

i. **Medicare Advantage (MA) Patients.** Unlike CMS's other standardized measures for dialysis facilities, the SEDR and ED30 (and Standardized Transfusion Ratio) exclude MA patients because their numerator case identification relies on outpatient claims, which are largely unavailable for these patients. We appreciate the difficulty CMS faces adapting its measures to the changing Medicare environment, but have substantial concerns with this approach. Specifically, we believe the exclusion of MA patients will create an untenable scenario in which these ED measures will effectively address a population that diverges considerably from that of the other QIP measures. This may be of particular importance with the ED30 measure, as CMS promotes it as the complement to the *Standardized Readmission Ratio for Dialysis Facilities* (NQF 2496), wherein the two measures together provide a full picture of patients who require emergent care following hospital discharge. But as the SRR includes MA patients and the ED30 does not, the denominator populations are inherently

different, and the picture provided by these complementary measures would be misleading.

Additionally, CMS notes in its measure submission materials 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. We believe that such variability in coverage patterns compromises the validity of the measures, putting states, regions, and individual facilities with a low proportion of MA patients at a substantial disadvantage with the ED measures.

ii. **All-Cause Construct.** As proposed, ED30 and SEDR capture all ED visits by ESRD patients, regardless of cause. KCP strongly objects to this construction, believing that it is too expansive in scope and will unfairly penalize dialysis facilities for random ED visits that are beyond their control and sphere of influence. Our analysis of ED encounters during 2015 (prior to implementation of ICD-10 diagnosis coding), showed that approximately 30 percent of encounters among dialysis patients were accompanied by principal discharge diagnoses in the range from 780.*x* to 799.*x* (*Symptoms, Signs, And Ill-Defined Conditions*). This lack of specificity about the nature of morbidity in the ED demonstrates that ED encounters cannot be readily attributed to any one health care provider, let alone an outpatient dialysis provider.

iii. **Ratio Construct.** As we have done with CMS's other standardized ratio measures (the SMR, SHR, SRR, and STrR), KCP again strongly recommends that ratio measures be avoided and that risk-adjusted rates or year-over-year normalized rates be used. For the ED30 and SEDR measures in particular, we note that there is precedent for this approach; specifically, CMS has developed and actively maintains stewardship of two NQF-endorsed home health ED utilization measures (NQF 0173 and 2505) that use the type of risk-adjusted rate to which we're referring.

iv. **Exclusions.** KCP recommends incorporating two additional exclusions into the ED30 and SEDR measure specifications: 1) ESRD patients who seek care in an ED for *any* reason (*including* those related to ESRD and dialysis care) after missing their most recent scheduled dialysis session; and 2) ESRD patients who reside in/are discharged to a Long-Term Care or Skilled Nursing Facility. We make the former recommendation on the basis that it is unreasonable to penalize a facility for medical issues for which it has not had the opportunity to intervene or arising from lack of adherence to prescribed care, and the latter because a dialysis facility should not be held accountable for medical decisions made by another provider (i.e., the LTC or SNF) and are beyond its realm of control.

v. **Urgent Care Centers.** KCP recommends that urgent care center revenue codes be included in the ED30 and SEDR numerators. The ED measures are inextricably tied to geographic locale, including but not limited to availability of EDs vs. urgent care centers. Because urgent care is not encompassed by the two measures (with the exception of centers located within an existing emergency

room), facilities where an ED option is more readily available geographically than urgent care will be inordinately penalized by these measures as compared to facilities with the same patient mix where urgent care is available. We believe this will exacerbate existing SDS and geographic disparities of the type documented by the December 2016 report issued by the Office of the Assistant Secretary for Planning and Evaluation.[1]

vi. **Risk Models.** We note that risk model testing yielded an overall C-statistic of 0.665 for the ED30 and 0.61 for the SEDR, raising concerns that the models will not adequately discriminate performance. Smaller units, in particular, might look worse than their actual performance. We reiterate our long-held position that a minimum C-statistic of 0.8 is a more appropriate indicator of a model's goodness of fit, predictive ability, and validity to represent meaningful differences among facilities.

II. Standardized Ratio for ED Encounters Occurring within 30 Days of Hospital Discharge (ED30)

KCP has identified a number of concerns and makes recommendations specific to the ED30, as follows:

i. **Reliability.** KCP posits the ED30 is not reliable as specified. Reliability testing for measure yielded an overall IUR of 0.451 across all facilities, indicating that only 45 percent of the variation in a score can be attributed to between-facility differences (signal) and 55 percent to within-facility differences (noise)—by statistical convention, a "poor" degree of measure reliability.[2,3] KCP believes it is incumbent on CMS to address the measure's empirically demonstrated lack of reliability and use an adjuster or otherwise account for the poor reliability before the measure receives further consideration.

Moreover, we fear the reliability for small facilities in particular might be substantially lower than the overall IURs, as has been the case with other CMS standardized ratio measures. To illustrate our concern, the *Standardized Hospitalization Ratio for Dialysis Facilities* (NQF 1463) was reported in 2013 (the most recent stratified data provided by CMS) to have an overall IUR of 0.70. However, the IUR was only 0.46 ("poor" reliability) for the nearly 35 percent of facilities (n = 2,028) meeting CMS's definition of "small" (<=50 patients, for the SHR). Without evidence to the contrary, KCP is concerned that the ED30 reliability is similarly lower for small facilities, effectively rendering the metric meaningless for use in performance measurement in this sizeable group of providers. Consistent with our previous stance on this matter, we believe it is incumbent on CMS to demonstrate reliability for all facilities by providing data by facility size and use its testing data to assess the impact of a "small numbers"

[1] U.S. Department of Health and Human Services, Office of Assistant Secretary for Planning and Evaluation *Report to Congress: Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs*, December 2016. https://aspe.hhs.gov/system/files/pdf/253971/ASPESESRTCfull.pdf. Last accessed May 19, 2020.
[2] A reliability statistic of 0.70 is generally considered as "acceptable" reliability.
[3] Adams, JL. *The Reliability of Provider Profiling: A Tutorial.* Santa Monica, California:RAND Corporation. TR-653-NCQA, 2009.

effect on reliability and to empirically determine appropriate facility-level exclusion parameters and adjust the specifications accordingly.

Finally, we note that CMS has incorporated a new reliability statistic into its testing protocol, the "Profile IUR", or "PIUR". The PIUR, which itself is quite low for this measure at 0.570, was developed by CMS's measure developer contractor UM-KECC to address the unacceptably low measure reliability "that can result when many facilities have outcomes similar to the national norm, even though the measure is still very useful to identify facilities with extreme outcomes." However, NQF's Scientific Methods Panel (SMP) noted in its April 1, 2020 conference call that the QIP measures are not *intended* to identify facility outliers, but rather to distinguish performance between providers. The Panel disagreed with the developer's assertion that the PIUR is an appropriate measure of reliability for the QIP measures, maintaining that the applicable statistic is the IUR. We concur with this assessment and further propose that a measure incapable of discerning performance between providers approximating the norm is not a meaningful or valid measure.

ii. **Stratification of Reliability Results by Facility Size.** KCP notes that unlike testing results provided for its other standardized ratio measures, CMS has provided no stratification of ED30 reliability scores by facility size; we are thus unable to discern how widely reliability varies across the spectrum of facility sizes. In particular, we are concerned that the reliability for small facilities is substantially lower than the overall IUR of 0.45 (already poor), as has been the case with other standardized ratio measures. For instance, the *Standardized Transfusion Ratio for Dialysis Facilities* (STrR) measure (NQF 2979) was found to have an overall IUR of 0.60—a "moderate" degree of reliability—however, the IUR for the STrR was only 0.3 for small facilities ("poor" reliability), which were defined by CMS for this measure as <=46 patients. KCP is thus concerned that the already-unacceptably low overall ED30 reliability (IUR = 0.45) is likely even lower for small facilities, effectively rendering the metric meaningless for use in performance measurement in this group of providers. We believe it highly likely that small facilities with as few as one or two patients who utilize ED services will be unfairly characterized as poor performers. KCP believes it is incumbent on CMS to demonstrate reliability for all facilities by providing data by facility size.

iii. **Meaningful Differences in Performance.** KCP posits that validity of the ED30 is low. An essential component of NQF's evaluation of validity is a demonstration of meaningful differences in performance. Testing results indicate that the ED30 can only distinguish differences in performance in less than 6 percent of facilities—specifically, 2.85 percent of facilities were classified as "better than expected" and 3.05 percent as "worse than expected." Simply put, the measure is unable to assess meaningful variations in performance in the overwhelming majority (94.10 percent) of facilities. This inability to discriminate between facilities illustrates the futility of using this measure, as specified, in a public reporting or value-based purchasing program—end-users will ultimately be unable to effectively compare or make informed decisions about the quality of

care provided in various facilities. Again, KCP recognizes the importance of assessing ED utilization by individuals with ESRD; however, testing results do not support the premise that the proposed ED30 metric will provide a valid (or reliable, as just noted) representation of quality.

III. Standardized ED Encounter Ratio (SEDR) for Dialysis Facilities

KCP had identified a number of concerns and makes recommendations specific to the SEDR, as below.

i. **Reliability.** Reliability testing for the SEDR yielded an overall IUR ranging from 0.62 to 0.63—a decrease from a previous version of the measure we reviewed 2017, then 0.64 to 0.72. We have significant concerns with a measure for which reliability has demonstrably decreased. And as with the ED30, reliability statistics were not stratified by facility size, again raising concerns about inadequate measure performance in small facilities, as has been the case with other CMS standardized ratio measures. With no evidence to the contrary, we cannot simply assume that the SEDR will provide reliable, meaningful information in this group of providers and urge CMS to supply reliability data by facility size.

Finally, as with the ED30, KCP concurs with the SMP's conclusion that the developer's proposal to use the PIUR in lieu of a poor or declining IUR is wholly inappropriate. We again posit that a measure incapable of discerning performance between providers approximating the norm is not a meaningful or valid measure.

ii. **Stratification of Reliability Results by Facility Size.** As with the ED30, CMS has not provided stratification of SEDR reliability scores by facility size, making it impossible to discern how widely reliability varies across the spectrum of facility sizes. Again, we are concerned that the reliability for small facilities may be substantially lower than the overall IUR, as has been the case with other standardized ratio measures and that small facilities with even one or two patients who utilize ED services might be unfairly characterized as poor performers. KCP believes it is incumbent on CMS to demonstrate reliability for all facilities by providing data by facility size.

iii. **Meaningful Differences in Performance.** KCP posits that the validity of the SEDR is low. Again, an essential component of the NQF's evaluation of validity is a demonstration of meaningful differences in performance. Empirical testing indicates that the SEDR can only distinguish differences in performance in approximately 5.65 percent of facilities (0.60 percent were characterized as "better than expected" and 5.05 percent as "worse than expected"); the measure was unable to assess meaningful variations in performance in the overwhelming majority (94.35 percent) of facilities. This inability to discriminate between facilities illustrates the futility of using this measure, as specified, in a public reporting or value-based purchasing program—end-users will ultimately be unable to effectively compare or make informed decisions about the quality of care provided in various facilities. We also note that the SEDR discrimination is

substantially more skewed towards poor performers than the ED30, providing additional evidence that the model is not performing well. We reiterate our recognition of the importance of assessing ED utilization by individuals with ESRD. Testing results, however, do not support the validity (or reliability, as noted above) of the SEDR; it will not provide an accurate and meaningful representation of quality as currently specified.

KCP again thanks you for the opportunity to comment on this important work. If you have any questions, please do not hesitate to contact Lisa McGonigal, MD, MPH (lmcgon@msn.com or 203.530.9524).

Sincerely,

Kidney Care Partners

Akebia
American Kidney Fund, Inc.
American Nephrology Nurses Association
American Renal Associates
American Society of Nephrology
American Society of Pediatric Nephrology
Amgen, Inc.
Ardelyx
AstraZeneca
Atlantic Dialysis Management Services, LLC
Baxter International, Inc.
Board of Nephrology Examiners Nursing Technology
B. Braun Medical, Inc.
Cara Therapeutics, Inc.
Centers for Dialysis Care
DaVita, Inc.
Dialysis Patient Citizens, Inc.
DialyzeDirect
Fresenius Medical Care North America
Fresenius Medical Care Renal Therapies Group
Greenfield Health Systems
Kidney Care Council
National Kidney Foundation, Inc.
National Renal Administrators Association
Nephrology Nursing Certification Commission
Renal Physicians Association
Renal Support Network
Rockwell Medical
Rogosin Institute

Satellite Healthcare, Inc.
US Renal Care
Vertex
Vifor Pharma