Technology Assessment Disposition of Comments Report

Title: *End-stage Renal Disease in the Medicare Population: Frequency and Duration of*

Summary of Peer Reviewer, TEP, and Public Comments and Author Response

Below is a list of common themes brought up by the commenters.

1. Missing articles
2. Critique of limiting the review to the US Medicare population
3. Comments on clarity of the inclusion and exclusion criteria
4. Requests for clarification on risk of bias assessment in individual articles and grading of the body of evidence for specific outcomes.

Changes made to the draft report to address these comments.

1. Missing articles: We reviewed all articles the commenters identified as "missing." After our evaluation we identified two articles that were missed and added these to the final report. The remaining articles were assessed as not applicable for the following reasons: not primarily a US Medicare population; no comparison group; interventions were not consistent with our inclusion criteria.
2. US Medicare population: Our scope of work was to assess the US Medicare population, therefore we limited our review to studies that included a US population, included more than 50% US participants, or stratified data by country.
3. Inclusion/exclusion criteria: We reviewed the inclusion and exclusion criteria and made minor adjustments. These adjustments corrected typos we believe occurred during copy editing. The report was consistent in how studies were included and classified.
4. Risk of bias and Grading: We added extensive explanations to the Methods section on strength of evidence and grading. We additionally added paragraphs to the Discussion section noting how the evidence in this review can be used.

Peer Reviewer, Technical Expert, and Public Comments and Author Response

#	Commentator & Affiliation	Section	Comment	Response
1	P			

#	Commentator & Affiliation	Section	Comment	Response
9	Peer Reviewer #1	Clarity and Usability		Tc -0.004lTw 99.3ejoy

#	Commentator & Affiliation	Section	Comment	Response
			between studies as to how outcomes are defined- and if so, how was this reconciled.	differences in the reporting of outcomes. However, we have included basic information about the outcomes that were reported in Appendix A
22	TEP Reviewer #1	Results		

#	Commentator & Affiliation	Section	Comment	Response
32	TEP Reviewer #1	Clarity and Usability	There is an opportunity to more clearly present the outcomes examined, definitions and data sources in aggregate.	We abstracted any data that fell under these outcomes, and details as available were provided in the appendices. Since we did not do any meta-analysis, we do not believe there is much

#	Commentator & Affiliation	Section	Comment	Response
				or insufficient evidence from these studies, there remains clinical equipoise on this topic.
41	Peer Reviewer #2	Results	Does there seem to be the potential to learn about this question through more observational work?	Yes, rigorously designed observational studies can answer some of the relevant questions. Our recommendations are included in the Research Recommendations section of Discussion.
42	Peer Reviewer #2	Results	Is there any evidence of treatment outcome heterogeneity. Are their some populations that may benefit more than others?	There may be treatment heterogeneity due to the selective nature of clinical trial populations. Descriptions of included studies for KQs 2 and 3 make note of this. Within the trial populations, heterogeneity was not reported but may be limited by small sample size.
43	Peer Reviewer #2			

#	Commentator & Affiliation	Section	Comment	Response
			to achieve the needed separation and thus despite its study design as a RCT, the actual inclusion of this trial creates a different problem and was not evident within the body of the report until Discussion.	

#	Commentator & Affiliation	Section	Comment	Response
				We do not have a “poor” grade of evidence, but we do have low strength of evidence. The algorithm in the methods make clear how the grades were assigned. Based on the AHRQ Guide, low level of evidence could be due to one or no RCTs, multiple study limitations, and inconsistent or imprecise estimates of effect size.
62	TEP Reviewer #3	Results	More information on comorbid health conditions of patients in the trials, observational studies, and the U.S. 2016 hemodialysis population are needed. Table 4 (page 14) displays the characteristics of age, race, education, and smoking. Comorbidity status from USRDS and the listed studies should be included. This critique applies throughout the review.	It is difficult to compare study populations based on the reported comorbidity characteristics as they are ascertained in different ways in different studies. It is well recognized that the comorbidities obtained from CMS form 2728 significantly underestimates the true comorbidity burden. This is further reflected in low mortality rates observed in the control arm of the clinical trials and the matched cohorts in observational studies. Further abstraction of comorbidity data is unlikely to add much to our understanding of differences between studies. Participant characteristics for these studies is available in Appendix E table 4.

#	Commentator & Affiliation	Section	Comment	Response
			2012 was not even listed. The reason given for exclusion for the Flythe article was lack of an intervention. While Brunelli et al used a prospective cohort, there was no intervention. Treatment times were prescribed clinically, just as the treatment times studied in the Flythe and Tentori articles. This is important as the Brunelli study is discussed at length in the review.	Tentori, 2012: A publication of the DOPPS study. A multinational study not stratified by country.
66	TEP Reviewer #3	Results	Related, inclusion of observational studies at all gives some pause due to the inherent limitations of observational studies and concern for substantial residual confounding and other biases.	We agree but the scope of our project was to review all available evidence, summarize findings to date, and make recommendations for future research. Therefore, inclusion of observational studies is relevant to this report.
67	TEP Reviewer #3	Results	One of the challenges of combining these data is the heterogeneity of interventions (e.g. "longer" and "shorter" dialysis as well as "more frequent" dialysis were defined differently across many of the studies). Given the general paucity of evidence to begin with, this is a substantial limitation and may deserve greater acknowledgement.	We agree and this substantial limitation is reflected in the Low and Insufficient levels of evidence grading for most studies. This is highlighted in the Discussion under "Limitations of the Systematic Review Process," and forms the basis for our Research Recommendations. A paragraph has been added to the Discussion section: Limitations of the Systematic Review process--Systematic Review process--Systematic Review process— Across all outcomes addressed in key questions 2, 3, and the combined 2 and 3, the strength of evidence was assessed as either low or insufficient. As described in the methods section of this report, we followed AHRQ guidance when we assessed the strength of evidence.[181] Following these guidelines reduces bias in assessing overall strength of evidence. A number of factors impacted these strength of evidence assessments. A primary contributing factor to lower strength of evidence assessments was important study limitations. None of the RCTs had low study limitations, with judgments ranging from "some concerns" to "high" as evaluated using the Cochrane Risk of Bias-2 tool.[182] Additionally, none of the cohort studies were judged to have low study limitations. Further, the available evidence was often imprecise or inconsistent across studies.

#	Commentator & Affiliation	Section	Comment	Response
			Similarly, most of the included more frequent dialysis studies considered dialysis frequencies of more than 4x/week. Suggest more explicitly stating the populations (and practices) to which this review can reasonably apply.	
69	TEP Reviewer #3	Discussion/ Conclusions	Many of the patient-reported outcome measures (PROMs) considered (e.g. CHEQ) were developed quite some time ago [redacted] contemporary dialysis population is unknown. Re-exploration of content validation may be necessary. [(nec)-2.7 (e)13.4 scussic3.3 (of40.4f40.4f40.4-2.c3.3 (o(scu)16.u)10.7 (b12i)2.3 hounc)-2.7 d bece	

#	Commentator & Affiliation	Section	Comment	Response
73	TEP Reviewer #4	General	In this report, the evidence review team compiles data on longer or more frequent dialysis as compared to usual/standard dialysis. There are several errors and inconsistencies in this report. I call out as many as I noticed, some major. Given these major issues and the importance of the question being asked, I feel strongly that this document should undergo a second round of peer review following responses to reviews and comments.	We thank the reviewer for their detailed review of the report. We have revised the report, as outlined in subsequent sections, to clarify potential inconsistencies in writing.
74	TEP Reviewer #4	General	US Medicare Population. The systematic review is somewhat inconsistent in the approach to the overall population. Specifically, the overarching criteria specifies US Medicare ESRD patients. This is not what is done. There needs to be a clearer explanation of inclusion and exclusion criteria within the actual manuscript including a mention within the actual manuscript of the amended search criteria and updating of how you are referring to eligibility based upon this amendment, the detailed rationale for these criteria included in the actual manuscript rather than the appendix, a review by the ERT of how included studies meet or do not meet these criteria, and a revisiting of the literature for missed studies based upon the criteria as written. I understand that the broad topic is dictated by the title; however, there are latitude in how the inclusion and exclusion criteria are conceptualized.	We used a consistent approach to identify our population. We have revised the methods section to clarify this information. Medicare population: U.S. Medicare population was our target population. However, Medicare enrollment was not an inclusion criterion for study selection. We have clarified that we included all U.S. hemodialysis studies of adults and children as over 90% of all U.S. ESRD patients are eligible for Medicare. To maintain generalizability to our target population, we included multinational studies if the U.S. participants constituted more than or equal to 50% of the study population or if the results were stratified by country to allow abstraction of results from U.S. participants. Amended search criterion: As outlined in Section VII of the publicly available Protocol(https://www.ahrq.gov/sites/default/files/wysiwyg/research/findings/ta/topicrefinement/esrd-protocol-2019-amended.pdf), the only criterion that was amended clarified that we will abstract multinational studies if the U.S. participants constituted more than or equal to 50% of the study population or if the results were stratified by country to allow abstraction of results

#	Commentator & Affiliation	Section	Comment	Response
			not the case for any of the studies that are included in the SR, either trials or observational studies, with the exception of those solely based on USRDS. FHN for example does not report insurance status but likely has a substantial non-Medicare population. TiME similarly has a substantial non-Medicare population, consistent with incident dialysis patients. Given that most dialysis patients in the US will be Medicare beneficiaries at some point in their treatment, this restriction, which is not enforced clearly or consistently anyway, should be removed from the text. I would also note that there were undoubtedly 'institutionalized' patients within included studies. The same comments apply.	have clarified that we included all U.S. hemodialysis studies of adults and children as over 90% of all U.S. ESRD patients are eligible for Medicare. To maintain generalizability to our target population, we included multinational studies if the U.S. participants constituted more than or equal to 50% of the study population or if the results were stratified by country to allow abstraction of results from U.S. participants. Based on our inclusion criteria, our results are generalizable to the US hemodialysis population. We have carefully reviewed the report to ensure that this comes across as intended. Institutionalized: We included studies where the dialysis was performed in-center or at home. If there are studies where institutionalized patients were included but not reported in methods of the paper, then institutionalized patients might have been included. We are not aware of any such studies based on our review and cannot make this assumption based on conjecture.
76	TEP Reviewer #4	General	In a data poor space, this seems to be a suboptimal limitation and, in fact, this limitation was recognized by the ERT when they	

#	Commentator & Affiliation	Section	Comment	Response
				developed countries such as Canada, Japan, and many European countries. The setting in which the trial was conducted can have implications on the findings and therefore maintaining focus on the U.S. studies is prudent. We have added a section in the Discussion outlining these differences and their implications for dialysis care: "This systematic review was designed to synthesize information of relevance to the U.S. hemodialysis population. The U.S. dialysis population is significantly different from the dialysis population in the rest of the developed countries.168-171…" 168. Foley RN, Hakim RM. Why is the mortality of dialysis patients in the United States

#	Commentator & Affiliation	Section	Comment	Response
			KQ1) and Table 1 in the manuscript (Intervention row, KQ3 and Comparator Row, KQ 1 and 4) all include 4 hours as standard dialysis while appendix table 1 (identical to manuscript table 3 in all other regards) and the comparator KQ 3 rows in Tables 1 and 2 in the manuscript define extended dialysis as 4 hours or more. This inconsistency is troublesome for a lack of transparency in methodology and possible post-hoc defining of this critical aspect of the systematic review. The first paragraph of the discussion messes this up as well, when it states: "We defined usual care as thrice weekly hemodialysis with a total treatment time LESS THAN OR EQUAL TO 12 hours per week. We defined longer hemodialysis as thrice weekly hemodialysis with treatment time GREATER THAN OR EQUAL TO 12 hours per week	
80	TEP Reviewer #4	General	Additionally, because 4 hours is a common duration, the control group in the FHN trials not infrequently were receiving 4 hour prescriptions. Please see supplemental figure 2 in the FHN Daily paper for example. Overall, the inconsistency within the SR methods and text on the threshold is highly troubling, and, at ultimately, the wrong decision appears to have been made.	The average time per dialysis session in the FHN Daily trial was 213 minutes or approximately 3.5 hours (Table 2; Chertow 2010). We reviewed Supplemental Figure 2 of the FHN Daily Paper. In the 3/week group: a) There was a small group of patients (~2.5% estimated from the figure) that received treatments more than 3 times per week. b) The majority (78%) of the patients had a weekly treatment time <12 t-d [(b))3.7 (T)15 (he m)-3 (a)13.3 (j)-2.7 (.)15.3 3

#	Commentator & Affiliation	Section	Comment	Response
			4 hours or more is even more common in many other par 0 9 1tmm • u^ìÀÀî™tEEu^ìªÌ W§á3é—¤SE| –"vNù 0 Pb'd$À) ': !0S}ïºÊS –ûï·á3Þ–$U~ûì¥7á3®™³< ™ < 2	

#	Commentator & Affiliation	Section	Comment	Response
84	TEP Reviewer #4			

#	Commentator & Affiliation	Section	Comment	Response
			looking at a healthier subset of dialysis patients and, given the size of the dialysis	

#	Commentator & Affiliation	Section	Comment	Response
			different problem as it is more likely to have more patients in the 1st 90 days, when mortality is highest (and therefore appears to do worse than USRDS). Options could be to tease out age and vintage-adjusted USRDS comparators to juxtapose with the trials or really explain this well in the footer. This gets at applicability of these data to patients (which I really do not want to call generalizability as saying not generalizable minimizes large swaths of dialysis patients).	

#	Commentator & Affiliation	Section	Comment	Response
107	TEP Reviewer #4	Results	Page 15 - incident or prevalent USRDS population. Be specific. Given that TiME recruited incident while the other trials (and most of the observational studies) are prevalent, you may want to report these comparisons separately.	Because of the heterogeneity in the studies including incident and prevalent populations, we have included both incidence and prevalence USRDS data in the report.
108	TEP Reviewer #4	Results		

#	Commentator &

#	Commentator & Affiliation	Section


Agency for Healthcare
Research and Quality

#	Commentator & Affiliation	Section	Comment	Response
			Nocturnal in 2 lines, one for on study and one for post-trial. Clearly note in the footnotes that you are using the long term follow-up and not the during trial data if this is your decision (again though, you should be focusing on the underpowered on-intervention data). In fig 11, for Rocco, you report on Trial. You cannot have this both ways, at least not without making it totally, completely clear.	
121	TEP Reviewer #4	Results	In the section on instruments, be specific re: dialysis population vs ESRD population, recalling that ESRD includes kidney transplant. Many of these items were developed for dialysis and not for ESRD.	We have revised the KQ 4 section to be more clear that we are including studies assessing patients with ESRD treated by dialysis
122	TEP Reviewer #4	Discussion/ Conclusion	The generalizability comment, specifically that these studies 'have limited generalizability' appears overstated, particularly for the observational studies. Specifically, although the population doing longer or more frequent dialysis is different than the overall dialysis population, this population is still consistent with a large swath of the US dialysis population. Clinically this is important as practicing nephrologists are not prescribing frequent or extended dialysis for all patients, but rather specifically for a subset. I worry that this aspect of the abstract is stated as a negative, while it could be restated as: "The studies of more frequent or longer hemodialysis regimens are more generalizable to younger and higher functioning dialysis patients." This is the same conclusion as is currently written but the different wording avoids the pejorative. With regard to race, FHN Daily, which is the best of the trials on this topic, had more than 40% black participants (so, be careful on the white conclusion in the generalizability comment in the abstract where you state that 'ALL study populations were younger, more likely to be white, and had lower mortality rates.').	We have revised the conclusions to focus on applicability rather than generalizability. The comment that the population doing longer or more frequent dialysis “is more consistent with a large swath of US dialysis population” is based on conjecture rather than facts.
123	TEP Reviewer #4	Discussion/ Conclusion	Table 46. You state: "The longer more frequent and longer hemodialysis treatments were provided hemodialysis systems that are	We have made this change.

#	Commentator & Affiliation	Section	Comment	Response

#	Commentator & Affiliation	Section	Comment	Response
133	TEP Reviewer #5	Discussion/ Conclusions	The major findings are clearly stated. I think the future resear	

#	Commentator & Affiliation	Section	Comment	Response
142	Peer Reviewer #3	Results	This reviewer has a significant comment regarding Page 14, Tables 4 and 5. The comparison of patient characteristics should include % diabetic, the single.4 (14,)85339 re W n Bactiihoue (ac)-po.7 (t)4 545002 Tw 9 -7 0 T0.6 (ude)13.4 (79r)3.7 ( )]TJ ET3	

#	Commentator & Affiliation	Section	Comment	Response
148	Peer Reviewer #3	Clarity and Usability	The report is very well organized and structured, with clearly defined points. It summarizes the state of the literature in a way that is easily digestible and stimulates thought.	Thank you for your comment
149	Fresenius	General	We wish to commend the authors of the draft on identifying many important studies of hemodialysis frequency and duration, including both Frequent Hemodialysis Network (FHN) trials, and summarizing study results about intermediate outcomes, clinical outcomes, and quality of life.	Thank you for your comment
150	Fresenius	General	The collection of relevant studies is flawed. Several randomized clinical trials of hemodialysis frequency and large observational studies of frequent home hemodialysis were excluded. Other observational studies were excluded because they ostensibly lacked a comparator, but other studies with similar designs were included.	We appreciate your concern about missing articles and have reviewed all of the studies you mention in subsequent comments. We do not believe we missed or erroneously excluded articles from this review. The protocol for this study and the inclusion/exclusion criteria were developed with extensive input from CMS, technical experts, and key informants, including Fresenius representatives.
151	Fresenius	General	The grading of strength of evidence appears to lack justification. In particular, the homogeneity of grading is suspect. Although we agree that there remains low strength of evidence that hemodialysis frequency and duration definitively modulate risks of death and hospitalization, we disagree that there is low strength of evidence that hemodialysis frequency modulates pre-	


Agency for Healthcare
Research and Quality

#	

#	Commentator & Affiliation	Section	Comment	Response
			First, although in most United States cohorts of patients receiving usual care, mean treatment duration is less than four hours, a sizable minority of patients in these cohorts are prescribed at least 4 hours per treatment. Second, in the Frequent Hemodialysis	


Agency for Healthcare
Research and Quality

#	Commentator & Affiliation	Section	Comment	Response

#	Commentator & Affiliation	Section	

#	Commentator & Affiliation	Section	Comment	Response
			are in direct conflict to those in the FHN Daily Trial.	
168	Fresenius	Results	Furthermore, the inclusion of the study by Lockridge et al (PMID: 21435157) implicitly raises an important question about the requirement of a comparator group. Technically, Lockridge et al studied a cohort of patients undergoing nocturnal home hemodialysis and compared the survival of that cohort with United States Renal Data System (USRDS) estimates of survival on conventional hemodialysis. This use of aggregated data from an external source to inform a comparison is qualitatively identical to the methodology employed by Kjellstrand et al (PMID: 18458034), who studied a cohort of patients undergoing short daily hemodialysis and likewise compared the survival of that cohort with USRDS estimates of survival on conventional hemodialysis. The report should consistently include or exclude studies of this nature.	

#	Commentator & Affiliation	Section	Comment	Response
			data were reported by Sergeyeva et al (PMID: 22505248).	
173	Fresenius	Results		

#	Commentator & Affiliation	Section	Comment	Response
			practical challenge of changing treatment duration in the facility setting.	the intervention and the usual care group was small (219; 95%CI, 217 to 222 minutes and 216; 95%CI, 214 to 219 minutes in the intervention group vs. 210; 95%CI, 209 to 213 minutes and 207, 95% CI, 206 to 211 minutes in the usual care group, respectively)" We have previously described the practical challenges to changing treatment duration identified in the TIME trial. We have revised this statement to more explicitly state that due to limited adoption of the intervention, the study was unable to determine whether extended hemodialysis sessions improves clinical outcomes. "No significant differences were seen in adherence to dialysis sessions in the TiME Trial, where 83.3 percent of patients in the usual care and 82.3 percent in the intervention group experienced a missed dialysis session. However, session duration did decrease over time, impacting the intervention group more than the control group. Due to insufficient uptake of the intervention, the study was unable to determine whether extended hemodialysis improves clinical outcomes. The authors indicated that both facility and patient factors were responsible for not achieving the desired 4.25 hours per session in the intervention group. Facility factors included perceptions by nephrologists and staff of lack of need for longer dialysis or potential burden. Patient factors included unwillingness to have longer dialysis sessions.
182	Fresenius	Results	A nuanced interpretation of the left ventricular mass in the FHN Nocturnal Trial requires synthesis of the data in Chan et al (PMID: 22360996). The Nocturnal Trial included a large share of patients without left ventricular hypertrophy at baseline, in whom regression would not be expected. Notably, among patients with left ventricular mass > 132 g at baseline, intensive versus conventional hemodialysis significantly lowered left ventricular mass and left ventricular mass index. This observation is not apparent in Figure 11.	We agree with the reviewer that it is important to note that "34% (Daily Trial) and 28% (Nocturnal Trial) of subjects had LVH at baseline. For the Daily trial, we did describe the difference in reduction in LVM by baseline LVM in the text Under results for KQ2 (LV Mass and Ventricular Volumes): "The (s)-2.7 (i)1 (w)19.3 (i)-0..6 (ar)19.3 (i)-0..02 Tw [(he 398.3 (j)-)2 (U)6 (n

#	Commentator & Affiliation	Section	Comment	Response
			hospitalization risks, are characterized by low strength of evidence, as that evidence is almost exclusively observational in its design. However, with respect to some physiologic and quality of life outcomes, there is much stronger and consistent evidence of effects of intensive hemodialysis, including evidence derived from randomized clinical trials. To this point, we strongly encourage the authors to conduct meta-analyses of the effect of hemodialysis frequency on clinical outcomes that were assessed in the FHN Daily Trial, the FHN Nocturnal Trial, and the trial of frequent nocturnal hemodialysis by Culleton et al. In a random effects model of the these trials, we have found that the frequent versus conventional hemodialysis engenders summary effects of -13.4 g on left ventricular mass, -9.6 mm Hg on pre-dialysis systolic blood pressure, -4.9 mm Hg on pre-dialysis diastolic blood pressure, -1.0 mg/dL on serum phosphorus, +2.4 points on the physical component score of the SF-36 quality of life survey, and +3.4 points on the mental component of the SF-36 quality of life survey. All these effects are evidently statistically significant ($P < 0.01$), and only the effect of intensive hemodialysis on serum phosphorus exhibited evidence of heterogeneity, with a predictably larger effect associated with nocturnal hemodialysis.	

#	Commentator & Affiliation	Section	Comment	Response
			with risk of major adverse cardiovascular events—remains unknown.	
185	Fresenius	Discussion	The authors should also re-assess their grade about the effect of increased hemodialysis frequency on post-dialysis recovery time. The findings of both FHN trials (PMID: 28094031), as well as the reported change in recovery time in the FREEDOM study (PMID: 20673601), consistently point toward a large reduction in recovery time after initiation of frequent hemodialysis.	

#	Commentator & Affiliation	Section	Comment	Response
			without some degree of “bias,” for example a selection bias or [illegible]	

#	Commentator & Affiliation	Section

#	Commentator & Affiliation	Section	Comment	Response
201				

[redacted]ction	Comment	Response