Technology Assessment Disposition of Comments Report

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

Draft report available for public comment from December 11, 2019 to January 10, 2020.

Citation: Shafi T, Wilson RF, Greer R, Zhang A, Sozio S, Tan M, Bass EB. End-stage Renal Disease in the Medicare Population: Frequency and Duration of Hemodialysis and Quality of Life Assessment. Technology Assessment Program Project ID No. JHE51000. (Prepared by the Johns Hopkins University Evidence-based Practice Center under contract number HHSA 290-2015-00006I) Rockville, MD: Agency for Healthcare Research and Quality. July 2020. Available at: http://www.ahrq.gov/research/findings/ta/index.html.

Comments to Draft Report

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each draft report is posted to the EHC Program Web site or AHRQ Web site for public comment for a 3-4-week period. Comments can be submitted via the Web site, mail or E-mail. At the conclusion of the public comment period, authors use the commentators' comments to revise the draft report.

Comments on draft reports and the authors' responses to the comments are posted for public viewing on the Web site approximately 3 months after the final report is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

This document includes the responses by the authors of the report to comments that were submitted for this draft report. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

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	Peer Reviewer #1	General	Yes, the review is clinically meaningful and the questions appropriate	Thank you for your comment
2	Peer Reviewer #1	Introduction	The rationale for the frequent dialysis is not well discussed	We have added the rationale for more frequent dialysis in the last paragraph of the introduction. The effects of more frequent dialysis are discussed in greater detail in the discussion section.
3	Peer Reviewer #1	Methods	The search criteria well outlined and is logical	Thank you for your comment
4	Peer Reviewer #1	Results	The details are adequate and message clear	Thank you for your comment
5	Peer Reviewer #1	Results	They fail to mention about Lecce dialysis in which patients undergo dialysis every other day for a minimum of 3 hours to reduce the prolonged 72-hour weekend interdialytic interval	Every other day dialysis (EODD), first reported in the literature in 1978 from Lecce, Italy, may be beneficial due to increased frequency of dialysis and avoidance of long interdialytic interval. In the context of this systematic review, EODD dialysis would have qualified as "more frequent" dialysis. We did not identify any US studies of EODD. A recent opinion paper in JASN highlights the logistical difficulties with EODD (Gul A. JASN 2018. Pubmed ID: 30185467) perhaps explaining why its implementation has been difficult. As there are no studies on EODD, we did not include it in our discussion. Based on the reviewer's comment, we have added EODD as a future research recommendation.
6	Peer Reviewer #1	Results	Also, the principle of square meter hour hypothesis need to be introduced	The Babb-Scribner square meter hour hypothesis* was one of the earliest (1971) attempts to quantify solute clearance. This systematic review focused on frequency and duration of dialysis but not the delivered dose (urea clearance) of dialysis. Therefore, we limited our discussion of the solute quantitation approaches. *Babb AL, Popovich RP, Christopher TG, Scribner BH: The genesis of the square meter-hour hypothesis. Trans Am Soc Artif Intern Organs 17: 81–91, 1971
7	Peer Reviewer #1	Results	Should discuss middle molecule hypothesis and include data about protein bound protein, gut derived uremic toxin and middle molecule clearance.	We discuss uremic toxins in our discussion of the generalizability of FHN results. The discussion is relevant to a variety of uremic toxins. We recognize that lack of knowledge about uremic toxins is a major limitation to improving care of patients on dialysis. We have included this as a research recommendation.
8	Peer Reviewer #1	Discussion/ Conclusion	Limitation stated well Implications in clinical care is not explicit- they do mention about loss of residual kidney function in nocturnal HD and should discuss the benefit of reducing the long 72 hour inter-dialytic interval if any	None of the studies were explicitly designed to shorten the interdialytic interval although the shortening is an expected benefit of increased frequency. We have included evaluation of every other day dialysis as a research recommendation.

#	Commentator & Affiliation	Section	Comment	Response
9	Peer Reviewer #1	Clarity and Usability	The report well structured and major points presented well. I did not see any remarkable new points being scored.	Thank you for your comment
10	Peer Reviewer #1	Clarity and Usability	The policy decisions is weak because of the weak data	The reviewer is correct that the strength of evidence is “Low” due to the quality of available studies. We have edited several sections to highlight the meaning of low strength of evidence and emphasized the importance of clinical trials in our research recommendations.
11	TEP Reviewer #1	General	The report is clinically meaningful and addresses an important question, namely- is there an association with longer and more frequent hemodialysis and clinical outcomes by conducting a metanalysis of published literature on the topic.	Thank you for your comment
12	TEP Reviewer #1	General	Key questions are clear and of importance	Thank you for your comment
13	TEP Reviewer #1	General	Target audience could be more clear- i.e. nephrologists, patients and policy makers	The target audience includes the groups mentioned by the reviewer, and this is addressed in the section about the Decisional Dilemmas.
14	TEP Reviewer #1	Abstract	would suggest presenting more clearly how longer and more frequent were defined in the included studies	We have added this information to the abstract.
15	TEP Reviewer #1	Abstract	Would also suggest qualification regarding "strength of evidence low"	We have added this information to the abstract.
16	TEP Reviewer #1	Introduction	helpful to have historical background about evolution of treatment times	Thank you for the comment. We also felt it was important to present the historical perspective.
17	TEP Reviewer #1	Introduction	As stated above, would benefit from defining longer and more frequent (i.e. more than 3Xs a week and 4 hours a treatment?).	We added this information to the abstract and outline the exact definition of longer and frequent dialysis in the Methods.
18	TEP Reviewer #1	Introduction	Fluid removal is primary benefit listed in terms of extending dialysis treatment, any clearance benefits or disadvantages (i.e. middle molecules removal).	We expand on the effects of longer and frequent dialysis on solutes in the Discussion section. The effects of volume removal are much more clearly delineated compared to solute removal as the knowledge about uremic toxins is limited. We have also included this as a research recommendation.
19	TEP Reviewer #1	Introduction	Consider comparison to PD in the introduction i.e. longer clearance and UF time.	It is correct that longer clearance and slower ultrafiltration rate may contribute to benefits of peritoneal dialysis. However, our systematic review was focused on the effects of hemodialysis and, therefore, we did not compare and contrast hemodialysis with peritoneal dialysis.
20	TEP Reviewer #1	Methods	The inclusion and exclusion criteria are clear and justifiable (presented in table 2). Table 3 is very helpful in terms of defining what longer and more frequent HD treatment entails.	Thank you for your comment.
21	TEP Reviewer #1	Methods	Need more detail about outcomes- i.e. how are these clinical outcomes defined (i.e. blood pressure control, infectious events, etc.), what are the data sources? Are there discrepancies	We abstracted any data that fell under these outcomes, and details as available are provided in the appendices. We did not do any meta-analysis because the studies varied in many different ways, thereby limiting the utility of adding more granularity about

#	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	The racial discrepancies in the trials are important to highlight. I also would favor clearly delineating the varied definitions of duration and frequency of dialysis treatments in the included studies (would incorporate in table).	This information is included in tables for individual Key Questions and includes both the frequency of dialysis and the duration of sessions in each study. These results are also displayed in Figure 4. Racial composition of studies is highlighted in each key question's "Description of included studies" and Table 5 and 6. The frequency and duration definitions of each study can be found in summary tables 7, 8, 19, 20, 28, 29.
23	TEP Reviewer #1	Results	Similarly with outcomes would incorporate in tables how they are defined and data sources for each study included.	Data sources are identified in the appendices. We decided not to add a lot of detail about the definitions of outcomes because we didn't think it would add enough to the interpretation of the evidence, given the limited data on each of the outcomes.
24	TEP Reviewer #1	Results	Consider summary table for all clinical outcomes.	All information is available in the appendices per AHRQ guidelines. We did not want to add more summary tables because of AHRQ guidance on trying to limit the length of the main body of the report.
25	TEP Reviewer #1	Results	All relevant studies appear to have been included.	Thank you for your comment
26	TEP Reviewer #1	Results	QOL analyses appear to be exhaustive, figure 12 is excellent.	Thank you for your comment
27	TEP Reviewer #1	Discussion/ Conclusion	The implications are clearly stated, and the limitations section is well delineated.	Thank you for your comment
28	TEP Reviewer #1	Discussion/ Conclusion	The issue of type of system and location of dialysis is particularly relevant, given AAKH and transition to more home-based NxStage based therapies. The challenges of applying FHN and TiME trials are well stated, but could be framed in that context.	We have added reference to AAKH to the discussion.
29				
30	TEP Reviewer #1	Discussion/ Conclusion	The future research section could benefit from referencing AAKH and the expansion of home dialysis modalities, how is home HD dosed based on this report? How do nephrologists consider the findings when advising patients regarding home HD versus PD	We are unable to say much about home HD dose based on available data. We cannot comment on HD versus PD as that was not the focus of this report. We have added reference to AAKH in the future directions section.
31	TEP Reviewer #1	Clarity and Usability	The report is well structured and organized, there is a considerable amount of information about the varied outcomes examined in the included studies.	Thank you for your comment

#	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 utility in adding this level of granularity Data sources are identified in the appendices. We decided not to add a lot of detail about the definitions of outcomes because we didn't think it would add enough to the interpretation of the evidence, given the limited data on each of the outcomes
33	TEP Reviewer #1	Clarity and Usability	The conclusion are timely and very relevant to policy given AAKH release in July. It is a nice synthesis of existing evidence.	Thank you for your comment
34	Peer Reviewer #2	General	The report overall is very meaningful. Since most of the results consist of weak evidence or no evidence there is limited clinical impact. As such the report would benefit more as a guide for future research.	Thank you for your comment. We have a section on research recommendations based on the findings of this report.
35	Peer Reviewer #2	Introduction	No specific comments. The overall report is well laid out.	Thank you for your comment
36	Peer Reviewer #2	Methods	The methods are well defined. The analysis is primarily qualitative which is appropriate.	Thank you for your comment
37	Peer Reviewer #2	Methods	It appears the authors did not consider non-US based studies. Can these be explained and justified?	The JHU Evidence-based Practice Center was given a specific scope of work that included the US-based Medicare population. As such, we only included studies that were conducted on predominantly US populations. Non-US-based studies, in general, did not meet the above qualifications.
38	Peer Reviewer #2	Results	The results are well laid out and the tables and figures are very useful.	Thank you for your comment
39	Peer Reviewer #2	Results	Can the authors comment on the quality of the observational studies?	We have added a detailed explanation of how we assessed the study limitations of observational studies—These details can be found in the Methods section. Observational studies included in KQs1-3 were assessed using the Cochrane ROBINS tool. Judgements on the risk of bias are based on several key domains. The quality of observational studies is addressed in the Discussion, under limitations of evidence, especially regarding unmeasured and time-varying confounders. For KQ4, we did not assess the quality of the individual studies—the purpose of this question was to identify tools and specific characteristics of the tools.
40	Peer Reviewer #2	Results	Do the comparison groups show equipoise?	The comparison groups in randomized clinical trials were similar, but not identical (See Results, Key Question 1). As there is low level of evidence

#	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	Discussion/ Conclusions	The conclusions overall are well laid out	Thank you for your comment
44	Peer Reviewer #2	Discussion/ Conclusions	The report could benefit from more detailed guidance on future studies. Particularly around the potential for observational studies.	The Research Recommendations section outlines recommendations for clinical trial generalizability, econometrics, and rigorous analyses of observational data, all of which can guide future observational studies of this topic.
45	Peer Reviewer #2	Clarity and Usability	Overall the report is well structured and easy to follow. The main conclusions are well laid out.	Thank you for your comment
46	Peer Reviewer #2	Clarity and Usability	Given the lack of clear clinical guidance from the review more attention should be paid to framing future research.	We agree and have outlined detailed recommendations in the Research Recommendations section of Discussion.
47	TEP Reviewer #2	General	Overall, I understand the methodology and the concept of evidence grading as outlined. However, it is not obvious to the average consumer how one would consider any evidence within the present subject matter to be strong given that most data are derived from a few if not single study and the patient population is not often generalizable.	In all assessments, the strength of evidence (study limitations) was assessed as low, some concerns, or insufficient for the interventions and outcomes assessed. This is reflective of the fact that there were few studies, and inconsistent or imprecise data. A high (strong) level of evidence was not attained for any intervention/outcome combination for a multitude of reasons detailed in the report. We have added further details on the level of evidence classification in multiple sections of the report. Two paragraphs have been added to the end of the Discussion section: Limitations of the 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,

#	Commentator & Affiliation	Section	Comment	Response
				with judgments ranging from “some concerns” to “high” as evaluated using the Cochrane Risk of Bias-2 tool. Additionally, none of the cohort studies were judged to have low study limitations. Further, the available evidence was often imprecise or inconsistent across studies. This review was not intended to be used as a guideline or guidance document. The review was conducted to collect and present the available evidence on this specific topic. Guidelines such as those created by KDIGO, also assess the strength of the evidence of studies used to develop their recommendation. However, guidelines are meant to provide advice on clinical topics
48	TEP Reviewer #2	General	To this aim, it would help the reader to orientate better if a potential scoring algorithm could be shown to allow all readers to understand what might be considered "strong" evidence (e.g. multiple RCTs, large patient population, precise estimates etc.)	We added a scoring algorithm to help explain the assessment of study limitations as well as a more detailed description of the grading scheme.
49	TEP Reviewer #2	General	This is different than the bulk of the evidence in the field of frequent hemodialysis which may be viewed as validation, proof of concept and small demonstration RCTs.	You are correct and that’s why the level of evidence was graded as Low, meaning that future research could change our interpretation.
50	TEP Reviewer #2	Introduction	Overall, the aims and key questions were outlined.	Thank you for your comment
51	TEP Reviewer #2	Introduction	However, we did not define clearly the various modalities: 1. conventional hemodialysis 2. frequency and duration of hemodialysis are different in using conventional hemodialysis machines versus low dialysate flow machines	We categorized the modalities based on time and frequency with usual care dialysis defined as thrice weekly hemodialysis for less than or equal to 4 hours per treatment. There were no head-to-head trials comparing low dialysate flow machines (NxStage) to conventional hemodialysis machines.
52	TEP Reviewer #2	Introduction	Given that there is lack of clear definitions, the outcomes are even more heterogeneous.	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 utility in adding this level of granularity. Data sources are identified in the appendices. We decided not to add a lot of detail about the definitions of outcomes because we didn’t think it would add enough to the interpretation of the evidence, given the limited data on each of the outcomes
53	TEP Reviewer #2	Methods	I have no specific issues with the methods.	Thank you for your comment
54	TEP Reviewer #2	Results	Amongst the RCTs, the TIME study was included as a trial which aimed to test the difference of 45 minutes of hemodialysis time in clinical outcomes (namely survival). However, it must be noted that the trial failed	We did address this limitation of the TiME trial in our Results section for KQ 3 “The TiME Trial was terminated early (median followup, 1.1 years) owing to an inadequate between-group difference in session duration (goal, 45 minutes; achieved, 9 minutes)”.

#	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.	
55	TEP Reviewer #2	Results	In contrast, I feel that we omitted to include ACTIVE trial which was published in JASN 2017. This is a multinational RCT which tested the difference between 3x4.5 hours versus 24 hours or greater of hemodialysis.	This article was captured in our search, and we excluded it because there are no US patients included. The trial includes patients from Australia, Canada, China, and New Zealand (see table1 in the article).
56	TEP Reviewer #2	Results	I feel that there was a flawed inclusion and an omission of ACTIVE trial. As a result, the totality of the studies and the presented results/estimates may differ somewhat.	This article was captured in our search, and we excluded it because there are no US patients included. The trial includes patients from Australia, Canada, China, and New Zealand (see table1 in the article)
57	TEP Reviewer #2	Discussion/ Conclusions	I feel that the Discussion was very focused on the strength of evidence, however, it should be noted that the focus of many of the included studies may be viewed as mechanistic in nature.	We agree that most of the studies, particularly observational studies, are mechanistic and hypothesis generating. In grading the level of evidence, such data are classified as insufficient evidence or low level of evidence. New clinical trials are the focus of our research recommendation and will be needed to provide a higher strength of evidence.
58	TEP Reviewer #2	Discussion/ Conclusions	Overall, there is a dichotomous theme that emerged: higher dialysis dose may be correlated with several clinical benefits (blood pressure change, LVH regression, normalization of phosphate, feasibility of full term birth), however given the small patient populations that were tested and that the remain only less than a handful of RCTs in the field, the totality of evidence is graded as "low".	Yes, this interpretation is correct.
59	TEP Reviewer #2	Clarity and Usability	I think the report is written well. There were some repetitions throughout the report.	Thank you for your comment. We will review the revised version to make sure there are fewer repetitions.
60	TEP Reviewer #2	Clarity and Usability	As stated above, I feel that there are a couple of themes that emerged: 1. there is a clinical signal which favored high dose dialysis 2. However, given the small overall sample size, selected patient group, the totality of evidence is graded as "low".	Yes, this interpretation is correct.
61	TEP Reviewer #2	Clarity and Usability	I think it is important to differentiate the fact that evidence is graded "poor" because of lack of effect.	A more detailed description of how grading was conducted was added to the Methods section.

#	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.
63	TEP Reviewer #3	Results	While the TIME Trial (Dember, JASN, 2019) is by far the largest relevant trial for this review, its interpretation is debated. The trial did not achieve separation in treatment time in the study arms and was stopped early by its monitors for “futility.” Some argue that the as-treated results cannot be interpreted since the “longer” arm was not substantively longer than the shorter arm, and the trial was stopped early. Despite its size and generalizability, the study may deserve less weight in this review than it is given.	We used the AHRQ Guide to grade the level of evidence. In the absence of RCT data, the evidence is graded as insufficient. So, the grading of the evidence as insufficient for the duration of dialysis, including TiME results, would not change even if it was excluded.
64	TEP Reviewer #3	Results	While the studies were generally small and included highly selected populations, Table A (ES-3) gives the appearance of reasonably strong evidence favoring more frequent HD (simply based on the number of outcomes in which more frequent HD was favored as opposed to insufficient evidence). The text and conclusions are not entirely consistent with this table as presented.	We have revised the text and tables to make it more clear that the strength of evidence was low to support these conclusions.
65	TEP Reviewer #3	Results	The reasons for the decisions to include vs. exclude several observational studies is not clear to me. Specifically, Brunelli et al. Kid Int 2010 was included in the analysis of treatment time; yet, Flythe et al. KI 2013 was excluded, and Tentori et al. [DOPPS] Neph Dial Trans,	Flythe, 2013. This study looks at dialysis time shorter than that defined as usual care. Brunelli, 2010. The intervention was hemodialysis using a pre-post measurement design in the cohort.

#	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.
68	TEP Reviewer #3	Introduction	Related, it may be worthwhile to clarify that the review focuses on consistent more frequent or longer dialysis and outcomes rather than occasional or targeted extra dialysis treatments (shorter or longer than standard treatments). The data should not be extrapolated to the practice of ordering "extra" treatments for thrice-weekly dialysis patients who are volume-overloaded or otherwise.	Thank you for pointing this out. We have added this to the discussion. "Importantly, we defined frequent or longer hemodialysis treatments based on either a consistent prescription in observational studies or randomized intervention in clinical trials. Extra hemodialysis sessions are sometimes prescribed based on clinical needs, generally fluid overload, to thrice weekly in-center hemodialysis patients. Our results cannot be extrapolated to this practice."

#	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 (decades) and their relevance to the contemporary dialysis population is unknown. Re-exploration of content validation may be necessary. Discussion of this issue should be considered.	Although many of the QOL tools were validated quite some time ago, there have not been any major scientific breakthroughs in our understanding of the mechanisms of symptoms experienced by patients with ESRD or technological breakthroughs in hemodialysis technology. Indeed, the prevalence of many symptoms in patients with ESRD has not changed much over the past 15 years, so the validated instruments remain relevant to the contemporary dialysis population. There are certain areas where validation is needed and others where new instruments may be needed. We have included these as research recommendations.
70	TEP Reviewer #3	Discussion/ Conclusions	Overall, the review is unbiased. However, the future research recommendations focus, potentially disproportionately on the need for validated instruments/scores for uremic symptoms, a research area of interest of at least one of the authors. Perhaps this research recommendation should be broadened to include other symptoms- such as volume-related symptoms and otherwise. The authors may intend for "uremic symptoms" to be an umbrella term for all ESKD symptoms but this is not clear, and it could be interpreted more narrowly- applying only to uremic toxin-related symptoms. Valid scores and instruments for all ESKD-related symptoms are needed.	We have edited the research recommendations to clarify the distinction between symptoms due to biological effects of kidney failure --- volume overload and retention of uremic toxins ---and other factors that may contribute to QOL. We agree that valid scores for all ESRD-related symptoms are needed.
71	TEP Reviewer #3	Results	Figure 3 (page 13)- Consider adding study years to the studies listed in the figure- particularly since the USRDS year (2016) is listed.	We have added the study year to the figure
72	TEP Reviewer #3	Results	On page 18, the authors note that a notable difference between the literature review and the USRDS data is that USRDS includes institutionalized patients who were not included in reviewed studies. To avoid this difference, the authors could consider only non-institutionalized USRDS patients. I believe this designation should be apparent in the USRDS administrative claims data.	We used data from the USRDS online Render system, which does not provide this information. While the USRDS database gives percentage of patients using home dialysis, this is not cross-referenced with patient characteristics in the data and not available on the USRDS RenDER database.

#	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 from U.S. participants. This information was included in the PICOTS table (Table 2) of the draft report and we have further expanded this information so that all relevant information is in the draft report. * This amendment was made at a very early timepoint in our review process and did not lead to any major change in the studies to be abstracted for this systematic review. We do not believe that our search criteria need to be amended further or literature search needs to be updated.
75	TEP Reviewer #4	General	Please clarify the need for Medicare for inclusion in this SR. Medicare only clearly is	Medicare: U.S. Medicare population was our target population. However, Medicare enrollment was not an inclusion criterion for study selection. We

#	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 amended their inclusion criteria. I would note that clinical practice is not dictated solely by trials conducted in the United States. In the kidney space, many trials have international composition. TREAT was just more than half US based and perhaps has had more influence on US dialysis policy than any other trial in the past 20 years. IDEAL was entirely Australian and is called out repeatedly by MedPAC and CMS. CREATE was entirely European. The decision that well-conducted trials without majority US representation are not consequential to this evidence base needs to be defended, as it seems very myopic to think that knowledge from elsewhere, particularly places with similar provision of care, cannot be applicable to a US population. The entire mission of DOPPS is based on the fact that the community can learn from different practices worldwide.	Protocol Amendment: 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 from U.S. participants. This information was included in the PICOTS table (Table 2) of the draft report and we have further expanded this information so that all relevant information is in the draft report. Data poor space and non-US studies: The limitations of large randomized clinical trials of dialysis in the U.S. are widely recognized. One of the overarching goals of this systematic review commissioned by CMS and AHRQ was to summarize evidence to date from U.S. studies and make recommendations for future research. Our systematic review protocol was reviewed extensively, prior to implementation, by CMS, AHRQ, technical experts, and stakeholders. The review summarizes the current hemodialysis evidence landscape in the US and provides research recommendations. Our findings are based on a rigorous unbiased approach and we believe it would be disingenuous to now amend the protocol based on post-hoc opinion. We fully understand that many insights into disease mechanism and treatments from non-US studies may also be applicable to U.S. patients. However, U.S. dialysis care and practices are different from other

#	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 much higher than the rest of the world? J Am Soc Nephrol. 2009 Jul;20(7):1432-5. doi: 10.1681/asn.2009030282. PMID: 19443632. 169. Goodkin DA, Bragg-Gresham JL, Koenig KG, et al. Association of comorbid conditions and mortality in hemodialysis patients in Europe, Japan, and the United States: the Dialysis Outcomes and Practice Patterns Study (DOPPS). J Am Soc Nephrol. 2003 Dec;14(12):3270-7. doi: 10.1097/01.asn.0000100127.54107.57. PMID: 14638926. 170. Yoshino M, Kuhlmann MK, Kotanko P, et al. International differences in dialysis mortality reflect background general population atherosclerotic cardiovascular mortality. J Am Soc Nephrol. 2006 Dec;17(12):3510-9. doi: 10.1681/asn.2006020156. PMID: 17108318. 171. Chapter 11: International Comparisons. Minneapolis, MN: USRDS Coordinating Center. https://www.usrds.org/2018/view/v2_11.aspx. Accessed on March 17, 2020.
77	TEP Reviewer #4	General	The fact that the criteria for inclusion was amended after the fact (specifically to get the 2 FHN trials into the SR) needs to be discussed in the main manuscript and not just in supplementary materials. Again, this is a bit messy as you end up excluding the 3rd trial in this space (Culleton et al, JAMA 2007) because it has solely Canadian representation, even though very similar patients are included in the FHN trials. Given the lack of clinical trial data in this space, the exclusion of a 6 month randomized clinical trial that assessed PROs is unfortunate, particularly as the population is similar to populations included in the SR and to populations included in observational data with extensive Canadian populations	Protocol Amendment: 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 from U.S. participants. This information was included in the PICOTS table (Table 2) of the draft report and we have further expanded this information so that all relevant information is in the draft report. We disagree with the reviewer's opinion that the protocol "was amended after the fact" to include FHN trials. The protocol amendment occurred as part of the protocol development process and such amendments are not unusual after the start of abstraction. We fully recognize the pivotal nature of the Culleton Study but we cannot amend the protocol post hoc to include every reviewer's suggestion for

#	Commentator & Affiliation	Section	Comment	Response
			(Nesrallah and Hladenuwich, with Nesrallah being majority international).	including a study if it did not get included based on our inclusion and exclusion criteria. We have added a paragraph in the Discussion section that discusses the Culleton paper and other international studies that are not included in our systematic review: "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 much higher than the rest of the world? J Am Soc Nephrol. 2009 Jul;20(7):1432-5. doi: 10.1681/asn.2009030282. PMID: 19443632. 169. Goodkin DA, Bragg-Gresham JL, Koenig KG, et al. Association of comorbid conditions and mortality in hemodialysis patients in Europe, Japan, and the United States: the Dialysis Outcomes and Practice Patterns Study (DOPPS). J Am Soc Nephrol. 2003 Dec;14(12):3270-7. doi: 10.1097/01.asn.0000100127.54107.57. PMID: 14638926. 170. Yoshino M, Kuhlmann MK, Kotanko P, et al. International differences in dialysis mortality reflect background general population atherosclerotic cardiovascular mortality. J Am Soc Nephrol. 2006 Dec;17(12):3510-9. doi: 10.1681/asn.2006020156. PMID: 17108318. 171. Chapter 11: International Comparisons. Minneapolis, MN: USRDS Coordinating Center. https://www.usrds.org/2018/view/v2_11.aspx. Accessed on March 17, 2020.
78	TEP Reviewer #4	General	For observational data, you state in the amendment that you will "include studies that are conducted in the US and countries outside of the US as long as data are stratified by country." I think you missed some studies, for example, Tentori NDT 2012, that would meet these criteria.	Tentori, 2012 is a publication of the DOPPS study. This study is multi-national and the population information is not stratified by country.
79	TEP Reviewer #4	General	Using these terms for simplicity as they are used in the SR, I have some major concerns with the KQs, specifically that, in my opinion, 240 minutes is not 'extended dialysis'. The definition of "extended" and "standard" dialysis is variable throughout the SR, with the results not consistent with the proposed methodology. Table 3 in the SR should be identical to Table 1 in the protocol, but they are different, likely reflecting a post hoc change made. Table 3 in the manuscript, Table 2 in the manuscript (Intervention row, KQ3 and Comparison row,	Thank you for your careful review. The <4 hour versus >=4 hours cut-off defined in the Protocol and in the text description of key questions is correct. We have thoroughly reviewed the report to make sure the time cut-off is consistently defined when using symbols and words. There was no post-hoc change in the protocol.

#	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, 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 hours per week. This figure does not report treatment time per session. We believe that the control group of FHN is correctly assigned as usual care.
81	TEP Reviewer #4	General	The decision-making behind how a 4-hour threshold was arrived at in the document must be discussed in detail as, in the results, 4 hours appears to be considered extended. Based on US data, including data from observational studies included in the SR as well as from the background you provide in introducing this project, 4 hours is part of the standard prescribing time in the US. The DPM shows that 4 hours is a common dialysis duration, used by a substantial proportion of thrice weekly in-center hemodialysis patients.	We included the following footnote in Tables 1-3 to help explain how we sorted studies by duration of dialysis – * Usual care involves 3 treatments per week with an average of 3.5 hours per treatment and a minimum of 3 hours per treatment. We don't believe that the <4 hour versus >=4 hours cut-off needs to be changed. For studies addressing Key Question 3 or the combined Key Questions 2 and 3, duration per treatment ranged from 4 hours to 7.5 hours Furthermore, the level of evidence will still be insufficient as these are all observational studies.

#	Commentator & Affiliation	Section	Comment	Response
			4 hours or more is even more common in many other parts of the world, as mean US dialysis duration is a bit shorter than most other countries. The data from Davita in one of your included cohort studies similarly show a high proportion of 4 hour dialysis times; similarly, in ArMORR, the most common dialysis duration was 4 hours. Interestingly, the TiME trial, where one group was included as extended despite only a 9 minute difference in dialysis duration between groups and early cessation of the trial due to a failure to achieve separation (not discussed), had a mean treatment duration of only 216 minutes in the 'extended' group, introducing huge heterogeneity that absolutely needs to be discussed regarding the utility of those data. In fact, the extended dialysis duration group in TiME is slightly shorter than the mean dialysis duration in the US, potentially reflecting that these are incident patients who may still have residual kidney function.	Re Time Trial, the following in the Results section of the report clearly outlines the lack of separation between the arms: "The TiME Trial was terminated early (median followup, 1.1 years) owing to an inadequate between-group difference in session duration (goal, 45 minutes; achieved, 9 minutes)."
82	TEP Reviewer #4	General	Similarly grouping studies describing nocturnal vs 'standard and 4 hours vs 3.5 hours together is really not helpful. These studies should be reported in separate parts of tables and clearly delineated as asking a different question given the totally different interventions.	We believe that the reviewer is referring to Table 28 that has only 1 randomized trial and one non-randomized trial. The Key Question it refers to is whether more frequent and extended dialysis duration improve outcomes. The treatment construct being addressed is both dialysis frequency and dialysis time and we believe it is appropriate to keep these studies in the same Table.
83	TEP Reviewer #4	General	Overall, this is a huge deal for the face validity of this systematic review. If you discuss with the dialysis community in the US, physicians and dialysis providers will tell you that a 4 hour duration is a common and fairly standard prescription. Including 4 hours three times a week as extended devalues all the work you have done (and the inconsistency throughout the manuscript when presenting the threshold for extended diminishes confidence in the methodology).	We arrived at the 4 hour threshold in consultation with our stakeholders (including CMS, technical expert panel, key informants). Although, 4 hours is often the prescribed dialysis time, it is often not the actual dialysis time. We do not believe that the <4 hour versus >=4 hours cut-off needs to be changed. Furthermore, the level of evidence will still be insufficient as these are all observational studies. We included the following footnote in Tables 1-3 to help explain how we sorted studies by duration of dialysis – * Usual care involves 3 treatments per week with an average of 3.5 hours per treatment and a minimum of 3 hours per treatment. We don't believe that the <4 hour versus >=4 hours cut-off needs to be changed. For studies addressing Key Question 3 or the combined Key Questions 2 and 3, duration per treatment ranged from 4 hours to 7.5 hours

#	Commentator & Affiliation	Section	Comment	Response
84	TEP Reviewer #4	General	To show generalizability, you compare to the USRDS population. Key issues relate to Race/Ethnicity and Incident/Prevalent status. Comparison to USRDS is a very reasonable strategy (and actually could allow for broader inclusion criteria as discussed above because you can discuss to what element of the US population data are most relevant). Critically, there is a huge error with the interpretation of race versus ethnicity data that results in major inconsistencies and incorrect conclusions. Additionally, there is heterogeneity in included studies between incident and prevalent populations, and, when comparing to the USRDS, it is important to choose whether to compare to incident or prevalent USRDS data given the marked differences between incident and prevalent patients in their characteristics and outcomes.	The focus of our systematic review was to summarize the existing literature to date on this topic. We added some comparisons with the US dialysis population using the USRDS RenDER online tool. The analyses suggested by the reviewer are important but need a carefully designed study that uses the full USRDS data. Such an analysis is beyond the scope of this project. Thank you for pointing out this error. We have corrected the population numbers on race and not counted Hispanic as a separate group for the USRDS data. Because of the heterogeneity in the studies including incident and prevalent populations, we have included both incidence and prevalence USRDS data in the report.
85	TEP Reviewer #4	General	When comparing to the USRDS population, you make errors in the race generalizability. For example, Figures 5 and 6 and 7 are incorrect as you are counting ethnicity as a separate race, decreasing Whites correspondingly. This is not what was done in USRDS or in most of the reviewed papers, with TiME and Ayus exceptions that counted Hispanic ethnicity as a subset of race distinct from White). This greatly impacts your generalizability conclusion with regard to race, which is entirely wrong as currently written.	Thank you for pointing out this error. We have corrected the population numbers on race and not counted Hispanic as a separate group for the USRDS data.
86	TEP Reviewer #4	General	You note that "The mortality rates in RCT and observational studies was lower than the rate in the U.S. dialysis population." This is a function of matching in the observational studies and the populations that are offered more frequent and longer dialysis both in current clinical practice and in trials as well as possible differences between incident and prevalent dialysis patients and the fact that many of the patients in the FHN and elsewhere may not have been Medicare beneficiaries. Again, this means that we are	We did not use “not generalizable” in the report. The following paragraph in the Discussion is a direct quote from the FHN authors: The FHN investigators were cautious in their interpretation of the results and noted that “relative to the entire North American hemodialysis population, participants in the FHN Daily Trial were younger, had longer dialysis vintage, and by design, had low levels of residual kidney function; therefore, these results may not be generalizable to all patients.”

#	Commentator & Affiliation	Section	Comment	Response
			looking at a healthier subset of dialysis patients and, given the size of the dialysis population, it means that results are applicable to a sizable subset (rather than the blanket comment of 'not generalizable', which is highly biasing). This could be more accurately stated as results being more applicable to dialysis patients with a longer life expectancy than average (slightly younger) rather than not being generalizable.	We understand your point regarding generalizability to the target population versus applicability for the contemporary population. We have edited our report in several places to make this clear.
87	TEP Reviewer #4	General	Even if not generalizable to the entire population, these data are APPLICABLE to many thousands of US dialysis patients, and this is really important.	We understand your point regarding generalizability to the target population versus applicability for the contemporary population. We have edited our report in several places to make this clear.
88	TEP Reviewer #4	General	Additional overarching comments regarding how the results are described: a. In reading this report, it seems to word conclusions in such a way that they are more readily perceived as negative rather than the weak positive and neutral results that are presented in the review. I give examples below and mention different ways to state these results that remain factually accurate.	We comment on each of the points in subsequent sections.
89	TEP Reviewer #4	General	In much of the manuscript, it suggests targeting only Medicare ESRD patients. This is inappropriate and also inaccurate as all of the RCTs and most of the observational studies included both Medicare beneficiaries as well as non-Medicare beneficiaries. I think that targeting all US hemodialysis patients is more appropriate and would accurate, although you are also informed by non-US patients, albeit to a far lesser extent. Please edit the 'Limitations' on Page ES-3 accordingly.	We agree that the results are applicable to US hemodialysis patients, the majority of whom are covered by Medicare. We have revised the report in several places to make this clear.
90	TEP Reviewer #4	General	What comes across from reading the manuscript is that this area is insufficiently studied; that needs to be a 'Main Point'. Essentially, clinical trial data are limited for these questions. I would consider stating the Ns enrolled in the 32 FHN trials in the main points with a bullet speaking to the limitations of existing data.	We have added this information to the evidence summary and abstract.

#	Commentator & Affiliation	Section	Comment	Response
91	TEP Reviewer #4	General	You note that "The mortality rates in RCT and observational studies was lower than the rate in the U.S. dialysis population." This is a function of matching in the observational studies and the populations that are offered more frequent and longer dialysis both in current clinical practice and in trials. As noted above, TiME's inclusion here also needs to be extensively discussed and cognitively justified given that, despite the intention, it was unable to even test its question (and the question that this SR is asking).	Re Mortality Rates: We have discussed this in several places in the report. Re TiME trial: We included all published studies that met our criteria. Ability to achieve intervention fidelity was not an inclusion criterion. We have discussed the limitations of the TiME trial in the Results and Discussion sections.
92	TEP Reviewer #4	General	As you are aware, many readers, including policy-makers, will Again, this means that we are looking at a healthier subset of dialysis patients and, given the size of the dialysis population, it means that results are applicable to a sizable subset (rather than the blanket comment of 'not generalizable', which is highly biasing). This could be more accurately stated as results being more applicable to dialysis patients with a longer life expectancy than average (slightly younger) rather than not being generalizable.	We understand your point regarding generalizability to the target population versus applicability for the contemporary population. We have edited our report in several places to make this clear.
93	TEP Reviewer #4	General	Again, for the reader who is likely to only read the abstract, I would reword. Given this, please consider rewording the abstract to read that "More frequent in-center hemodialysis may improve clinical outcomes, including mortality and quality of life. The strength of evidence supporting this conclusion is low and the population studied in clinical trials able to test this question was, on average, younger and healthier than the broader US dialysis population limiting extrapolation to older dialysis patients." I think that this is a very accurate statement of these results.	We have edited the conclusion of abstract to incorporate these suggestions.
94	TEP Reviewer #4	Introduction	Page 1. Hemodialysis is effective in removing volume (not water). While it is possible to remove free water in significant excess of solute in dialysis patients, this is not what we typically do. We ultrafilter VOLUME by manipulating TMP. Free water can be	We have made this change.

#	Commentator & Affiliation	Section	Comment	Response
			removed by having high dialysate osmolality, but this is not what we do and would be limited by back diffusion of sodium anyway.	
95	TEP Reviewer #4	Introduction	Page 1. The focus did not shift to 'achieving optimal urea clearance on dialysis' as we do not know what that means. It shifted to achievement of a specific threshold of urea clearance based on low quality evidence and opinion, with a p value of 0.06 for dialysis duration in the NCDS (NEJM 1981) resulting in the primacy of time average urea concentration over dialysis time based on an association with hospitalization. It is ironic that, given the discussion of low quality evidence here, the primacy of urea clearance over volume control is presented with little hesitancy in the supporting introduction.	The KDOQI 2015 Hemodialysis Adequacy Clinical Guidelines (reviewer was a member of that panel) supports urea-based targets. The Guideline committee could have removed their emphasis on the "primacy" of urea clearance, but did not, and these targets are used as dialysis quality metrics. The guideline document states (https://www.ajkd.org/article/S0272-6386(15)01019-7/fulltext): "Small-solute clearance is currently considered the best measure of HD and its adequacy. Kt/V, the fractional urea clearance, is the most precise and tested measure of the dialyzer effect on patient survival and is the most frequently applied measure of the delivered dialysis dose." We disagree with the reviewer's comments that we are presenting "primacy" of urea clearance in the introduction. It is merely a historical introduction to why the treatment time has become shorter.
96	TEP Reviewer #4	Introduction	Page 1. You state: "First, the major benefit of more frequent or longer dialysis treatment seems to be from volume removal." Do you mean more gradual volume removal or greater total volume removal or both? Or more reliable achievement of a volume steady-state? Please be specific.	We have edited this sentence to clarify that we are referring to total volume removal.
97	TEP Reviewer #4	Introduction	Page 2. You state that 'Each dialysis treatment takes 4 to 6 hours away from a day...' This is not exactly true although is common for in-center hemodialysis. For home dialysis, sessions may be shorter, and nocturnal home or in-center dialysis may be done to avoid taking time away from a day. Please be precise here, stating in-center hemodialysis whenever that is what you are referring to.	We have changed "4 to 6" to "several hours".
98	TEP Reviewer #4	Methods	See above, as I comment extensively on Methods in the General Comments section. The inconsistency throughout the SR is troubling, and, in particular the noise around a >=4 hour versus a >4 hour threshold throughout the manuscript reflects what I think may have been poorly established definitions from the start.	We have addressed the time cut-offs and reference to US Medicare population in previous comments. We have included the Culleton paper in our Discussion: "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…"

#	Commentator & Affiliation	Section	Comment	Response
			I have issues with the limitation to US Medicare, which is really not enforced anyway (as FHN and TiME and some of the observational papers used non-Medicare patients). I would include Culleton if possible - or at least mention it in additional discussion or a sensitivity analysis given that it is the only trial like this not included.	168. Foley RN, Hakim RM. Why is the mortality of dialysis patients in the United States much higher than the rest of the world? J Am Soc Nephrol. 2009 Jul;20(7):1432-5. doi: 10.1681/asn.2009030282. PMID: 19443632. 169. Goodkin DA, Bragg-Gresham JL, Koenig KG, et al. Association of comorbid conditions and mortality in hemodialysis patients in Europe, Japan, and the United States: the Dialysis Outcomes and Practice Patterns Study (DOPPS). J Am Soc Nephrol. 2003 Dec;14(12):3270-7. doi: 10.1097/01.asn.0000100127.54107.57. PMID: 14638926. 170. Yoshino M, Kuhlmann MK, Kotanko P, et al. International differences in dialysis mortality reflect background general population atherosclerotic cardiovascular mortality. J Am Soc Nephrol. 2006 Dec;17(12):3510-9. doi: 10.1681/asn.2006020156. PMID: 17108318. 171. Chapter 11: International Comparisons. Minneapolis, MN: USRDS Coordinating Center. https://www.usrds.org/2018/view/v2_11.aspx. Accessed on March 17, 2020.
99	TEP Reviewer #4	Methods	The race numbers pulled from USRDS were done incorrectly, such that Hispanic ethnicity was counted as a distinct race, markedly decreasing the number of whites and contrasting with how the studies report race. This error impacts your generalizability conclusions.	We have revised the population numbers to be more precise about how race and ethnicity were reported.
100	TEP Reviewer #4	Results (summary)	You state in the summary that "All interventions are compared with usual care (hemodialysis 3 treatments per week, less than 4 hours per treatment)." I think thisThis is incorrect as patients could be receiving 4 hour sessions (at least in the FHN trials - see supplemental figure 2 in the FHN Daily paper for example, where there were patients getting 12 or more hours of HD per week in the control group). Be careful here as the observational data are similarly not that clean and TiME also failed to achieve separation. There should (such that both the intervention and control group were receiving less than 4 hours). There needs to be a more detailed presentation in tables of the actual duration and variability in duration of dialysis in the 'intervention' and the 'control' groups.. in the manuscript. Some of this comes out in supplemental Table 3 but these are such important data that they need to be featured in the actual manuscript in my opinion.	Regarding the FHN trial: 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 hours per week. This figure does not report treatment time per session. We believe that the control group of FHN is correctly assigned as usual care. Regarding the TiME trial: we have clarified in several places in the results section that the TiME trial failed to reach separation.

#	Commentator & Affiliation	Section	Comment	Response
101	TEP Reviewer #4	Results (summary)	Table A: I would reword the table title as follows: "Summary of outcomes in individuals receiving more frequent and or longer duration hemodialysis as compared to standard thrice weekly in-center hemodialysis" You can then make the point in the title or in the footer that all evidence is of low strength. As currently worded, I am left looking for the tables with moderate and with high strength evidence. These do not exist. In the footnote, you comment on blank entries, but there are no empty boxes.	We have revised the Table A title and added footnotes to better explain the data
102	TEP Reviewer #4	Results	Page 15. You state that "In comparison to the percentage of White patients in the USRDS population (39.6%),43 two of the RCTs (FHN Nocturnal and TiME)27, 28 reported a higher percentage (55.2% and 57.8%, respectively) while the FHN Daily trial reported a lower percentage (36.3%).26" See table 1.7 in the USRDS ADR from 2018 (prevalent patients 2016). Among HD patients (column D), the proportion of white patients is 56.7% while the proportion of blacks is 35.9%. Initially I thought that you looked at the proportion of ESRD patients, which includes PD and transplant recipients as well as HD patients, but, on further reviewing this, what I think you did for the figures is count Latino/Hispanic as neither white nor black, which is not how this was done in the trials or in the USRDS and represents a major error here. It is also important to note that FHN didn't have 0 Hispanic/Latino participants; rather they did not report ethnicity - so that part of Figure 6 is also wrong (not reported must appear different than 0%). Really, Figures 5 and 6 (and 7) need to be entirely redone.	We have revised the population numbers to be more precise about how race and ethnicity were reported, We have also corrected the problem with reporting 0 rather than not reported.
103	TEP Reviewer #4	Results	Figure 3 is interesting. Several comments: 1. The USRDS data combine both incident and prevalent dialysis patients whereas the FHN trials look only at prevalent patients with much longer vintage. This really must be noted in this figure for the USRDS vs FHN comparison. TiME interestingly has a slightly	We have added a footnote to clarify the USRDS data includes incident and prevalent populations.

#	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).	
104	TEP Reviewer #4	Results	In figure 3, please justify the inclusion of the extended follow-up data for FHN Nocturnal. If presenting this, I would also recommend presenting the data during the actual study conduct or only present the deaths during the study, reserving the extended follow-up results for the table footnote.	We have replaced the extended follow-up rate of death per patient-year with the trial period only death per patient-year.
105	TEP Reviewer #4	Results	Table 4. Wrong as Ethnicity is not distinct from race and the data in this table do differ somewhat from Table 1.7 in the US (N is a bit higher in the USRDS data table).	We have revised the population numbers to be more precise about how race and ethnicity were reported.
106	TEP Reviewer #4	Results	Table 5. Report race and ethnicity separately. If ethnicity is not reported, indicate NR rather than 0. The Ayus report states 'Ethnicity' and does break down non-Hispanic whites and non-Hispanic blacks; this study fails to report race. TiME reports race and ethnicity together, but, because USRDS does not do this, you need to stick to how USRDS does this. I would suggest asking the TiME investigators for the race breakdown. Figure 4. Incident or prevalent patients? I would consider having 2 figures - with 'prevalent' juxtaposed with the FHN and other reports and 'incident' juxtaposed with the TiME trial For USRDS data, it would be very helpful to state the specific data table that data are drawn from throughout this report. Critically, the 'RCT' bar in Fig 4 applies to the FHN trials and not to TiME. This should be made clear.	Regarding table 5, we have revised the population numbers to be more precise about how race and ethnicity were reported. Regarding the comment about figure 4: A figure was added upon revision. Figures 3, 5, 6, and 7, have all been revised to include incident and prevalent 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	Page 15. The percentage of White patients in the USRDS population was NOT 39.6%. This reflects the subtraction of the Hispanic population from the total White population and cannot be juxtaposed with data from any of the observational studies or from the FHN trials.	We have revised the population numbers on race and not counted Hispanic as a separate group for the USRDS data.
109	TEP Reviewer #4	Results	Page 18. (and appendices). I would describe the observational studies a bit differently. Essentially, rather than saying multicenter, I would be more specific and state that some of these span LDOs. (so huge swaths of the US). This is much more precise and much less confusing than the paragraph currently reads. This holds for the figures too. Rather than combining some very different observational study designs, separate out the retrospective analyses of big chain data from the other analyses. This will result in far more consistent data presentation than the current lumper approach that was taken.	We have updated the summary tables in the results to provide this information
110	TEP Reviewer #4	Results	Table 6. Again, be very cautious with conflating not reported with 0%. For example, Brunelli 2010 (ArMORR) did not have zero Hispanics; they just did not report it. Table 6 needs to state what the prescription was for the 'extended' and 'control' groups and how controls were selected for analyses.	We have revised the table to report 0 rather than not reported, and to clarify what the prescription was for the study groups.
111	TEP Reviewer #4	Results	Decision to include 4 hours as extended dialysis. This is discussed extensively above. ArMORR is an example of a report where, in usual clinical practice, 240 minutes was the most common 'extended' dialysis prescription, making this very different from some of the clinical trials, particularly TiME. The >=4 hours rather than >4 hours is a major KQ decision that I think was flawed. For example, ArMORR demonstrates that	We arrived at the 4 hour threshold in consultation with our stakeholders (including CMS, technical expert panel, and key informants). We don't believe that the <4 hour versus >=4 hours cut-off needs to be changed. Furthermore, the level of evidence will still be insufficient. We included the following footnote in Tables 1-3 to help explain how we sorted studies by duration of dialysis – * Usual care involves 3 treatments per week with an average of 3.5 hours per treatment and a minimum of 3 hours per treatment. We don't believe that the <4 hour versus >=4 hours cut-off needs to be changed. For studies

#	Commentator & Affiliation	Section	Comment	Response
			~44% of patients had a 4 hour dialysis session. This almost standard of care prescription cannot be compared to the concept of extended nocturnal or to more than thrice weekly dialysis. Even TiME had a > 4 hour rather than a >=4 hour threshold for longer dialysis. although this was not achieved (and hence the early discontinuation of the trial).	addressing Key Question 3 or the combined Key Questions 2 and 3, duration per treatment ranged from 4 hours to 7.5 hours
112	TEP Reviewer #4	Results	Page 24. The statement that "Patients included in all the RCTs and most of the observational studies were younger and more likely to be white compared with the overall U.S. dialysis population." is wrong with regard to race as discussed above. As previously noted, please be careful with comparisons with USRDS to use incident or prevalent USRDS data as appropriate.	We have revised the population numbers on race and not counted Hispanic as a separate group for the USRDS data. The revised text shows a comparable percentage of Whites.
113	TEP Reviewer #4	Results	Page 30. You state: "The primary analysis of the FHN Daily Trial26 did not have sufficient power to assess death as a primary endpoint during the 12-month study period when there were five deaths (4%) in the frequent dialysis arm and nine deaths (7.5%) in the conventional dialysis arm." More correctly, the study was not powered or designed for a mortality outcome in isolation. In my opinion, the better way to write this would be to simply state the number of deaths and then launch your next sentence. So, please consider editing this to read, "During the 12-month study period when there were five deaths (4%) in the frequent dialysis arm and nine deaths (7.5%) in the conventional dialysis arm." More correctly and precisely, the study was notneither powered ornor designed for a mortality outcome in isolation. In my opinion, the better way to write this would be to simply state the number of deaths and then launch your next sentence. So, please consider editing this to read, "During the 12-month study period when, there were five deaths (4%) in the frequent dialysis arm and nine deaths (7.5%) in the conventional dialysis arm. Frequent dialysis (6 times per week) was associated with statistically significant	We have revised the statement accordingly. "During the 12-month study period. there were five deaths (4%) in the frequent dialysis arm and nine deaths (7.5%) in the conventional dialysis arm."

#	Commentator & Affiliation	Section	Comment	Response
			beneficial effects on the two primary composite outcomes: of death or increase (from baseline to 12-months) in LV mass (Hazard Ratio (HR) 0.61; 95% confidence interval (CI), 0.46 to 0.82); and death or decrease (from baseline to 12-months) in physical-health composite performance (HR 0.70; 95% CI, 0.53 to 0.92), compared with conventional thrice-weekly, in-center hemodialysis." This is a much better stating of results.	
114	TEP Reviewer #4	Results	Page 41. Not to beat a dead horse, but… Is a 4 hour dialysis session 'extended dialysis'? I do not think that this is the common definition. This does not really impact the FHN data and would not impact Culleton but has a huge impact on some of the observational data and, given that TiME had a mean duration of less than 4 hours achieved in the extended dialysis allocation, it is important there as well. This holds for the other KQs too.	We included the following footnote in Tables 1-3 to help explain how we sorted studies by duration of dialysis – * Usual care involves 3 treatments per week with an average of 3.5 hours per treatment and a minimum of 3 hours per treatment. We don't believe that the <4 hour versus >=4 hours cut-off needs to be changed. For studies addressing Key Question 3 or the combined Key Questions 2 and 3, duration per treatment ranged from 4 hours to 7.5 hours
115	TEP Reviewer #4	Results	Page 41, when you say 'longer', that is an accurate statement. Extended would be inaccurate. Key finding regarding race is incorrect. The RCT, TiME, was similar in age to USRDS incident, making that generalizability statement overly generalized. Critically, TiME, due to the failure to gain separation, ended up having an intervention group consistent with what we would call usual care. (some longer, some shorter, and a mean close to the US average). This is really important to stress.	Regarding longer versus extended—the question clearly states "extended" hemodialysis duration. Regarding race: We have corrected the section regarding race. Regarding generalizability: we have updated our conclusions. Regarding TiME trial—we have discussed failure to gain separation in the report.
116	TEP Reviewer #4	Results	Page 41-43. The KQ here is "Does extended hemodialysis duration (daytime, 4 or more hours per session, or nocturnal, overnight) improve objective outcomes over the long term (more than 6 months) compared with usual length hemodialysis duration (less than 4 hours)?" The Miller study does not really meet the KQ criteria as written, unless you clearly explain that you are comparing only the 4+ hour group.	Miller et al was included here because we were including studies for KQ 3 that were looking at duration only, and included at least one arm with an extended duration time, per our protocol.

#	Commentator & Affiliation	Section	Comment	Response
117	TEP Reviewer #4	Results	You can simplify tables 19 and 20 by removing the three times per week from each row and adding to the title. In Table 20, maintain consistency - Rivara you report minutes while for the others you report criteria for inclusion. Ideally you report criteria for inclusion in the specific time group and then report actual duration (mean or median) for that specific group. State the 'control' for Troidle (same patients prior to transition to nocturnal). I would be consistent throughout the manuscript on using either minutes or hours for duration.	For consistency with the other sections, we think frequency and duration should be reported in the comparison column. We agree that the duration units should be recorded consistently. This has been corrected in all tables
118	TEP Reviewer #4	Results	Figure 9, Miller is reported in a manner inconsistent with how the study is described in the table. Miller is reporting in their Figure 3, the HR for 240+ as compared to a reference of 210-<240 minutes. If comparing to the other duration groups, results are different. Please show this clearly. In Figure 9, Brunelli is also reported incorrectly. The HR of 1.42 is for shorter duration, so the result does NOT favor control.	Regarding Brunelli—we have recalculated the results to reflect extended versus shorter hemodialysis duration and have revised the figure. Regarding Miller: the data has been re-reviewed and it is correct.
119	TEP Reviewer #4	Results	Page 46. You state: "In the TiME Trial's full analysis population, those receiving longer dialysis gained an average (SD) of 1.93 (0.98) kg, while those receiving conventional hemodialysis gained an average (SD) of 1.88 (1.00) kg; these results were not significantly different (p=0.28)." This is not an accurate statement. Please rephrase something like: "Those incident patients receiving hemodialysis in facilities randomized to prescribe longer dialysis..." Again, TiME got 9 minutes of separation between groups, so you need to acknowledge the ITT design whenever you discuss this trial. and be cautious when describing what they actually received in TiME.	We have revised the text as suggested.
120	TEP Reviewer #4	Results	Table 28 and Figure 10 (and similar area). For Rocco, you show the results of mortality for the extended follow-up when patients were off of their intervention. This likely is not what you should do for obvious reasons. If you elect to show the post-trial outcomes, show FHN	We have added a footnote stating that the FHN Nocturnal Trial results are post-trial off-intervention outcomes.

#	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
			different from what is being used in contemporary practice (NxStage)." This is overall true, although people are using Fresenius machines as well in the US, so I would insert a statement to indicate 'what is currently being used most often in contemporary US practice." This is more correct.	
124	TEP Reviewer #4	Discussion/ Conclusion	Table 46. You comment that the control groups in the observational studies had lower mortality. This is only true for a small subset of the observational studies.	In most observational studies where patients receiving more frequent dialysis or longer dialysis are matched to control patients, the mortality in control patients is lower than the general population. We have added "most" before "observational studies".
125	TEP Reviewer #4	Clarity and Usability	See comments above. The fact that 240 minutes (4 hours) was considered extended dialysis is, in my opinion, a very poor choice, given how common this prescription is both in the US and worldwide. This introduces major heterogeneity into the 'intervention' group, precluding much of a conclusion.	We arrived at the 4 hour threshold in consultation with our stakeholders (including CMS, technical expert panel, and key informants). We don't believe that the <4 hour versus >=4 hours cut-off needs to be changed. Furthermore, the level of evidence will still be insufficient. We included the following footnote in Tables 1-3 to help explain how we sorted studies by duration of dialysis – * Usual care involves 3 treatments per week with an average of 3.5 hours per treatment and a minimum of 3 hours per treatment. We don't believe that the <4 hour versus >=4 hours cut-off needs to be changed. For studies addressing Key Question 3 or the combined Key Questions 2 and 3, duration per treatment ranged from 4 hours to 7.5 hours
126	TEP Reviewer #4	Clarity and Usability	I would organize the observational data differently, as discussed above.	We have organized the results to be consistent with our research methods.
127	TEP Reviewer #4	Clarity and Usability	Please include an early table defining key terms used in later data tables, such as 'consistency', 'directness', 'precision' and 'strength of evidence', including how this was arrived at.	A glossary is included at the end of the report, following AHRQ guidance.
128	TEP Reviewer #4	Clarity and Usability	Please state in a table somewhere why you gave various studies various designations (i.e., state precisely why TiME had a high risk of bias).	To address this concern, we have added very detailed information in the Methods section about how risk of bias was ascertained. We also include (and cite) the risk of bias judgement for each study in the appendix.
129	TEP Reviewer #5	General	The report is clinically meaningful. Results are clear to me. Key question 4 is less clinically meaningful as the other key questions.	Thank you for your comment
130	TEP Reviewer #5	Introduction	Clear and concise	Thank you for your comment
131	TEP Reviewer #5	Methods	Yes, yes, yes, yes	Thank you for your comment: We are assuming that these are affirmative answers to subquestions on Methods
132	TEP Reviewer #5	Results	The results section is very clear with the tables succinctly describing the trials and studies and the results.	Thank you for your comment

#	Commentator & Affiliation	Section	Comment	Response
133	TEP Reviewer #5	Discussion/ Conclusions	The major findings are clearly stated. I think the future research questions were very good.	Thank you for your comment
134	TEP Reviewer #5	Clarity and Usability	The report is well organized. I do think key question is not as relevant. The summary of the studies is very important for nephrologists and for policy.	Thank you for your comment. We addressed all questions in the scope of work in the hopes that our information would impact the care of individuals on hemodialysis in the future.
135	Peer Reviewer #3	General	This report provides a very important summary of the available evidence for hemodialysis duration and frequency. The approach is rigorous and systematic. The key questions are clear and appropriate.	Thank you for your comment.
136	Peer Reviewer #3	General	The addition of assessment of QOL measurements at first seems to be tacked-on, but make sense on further reflection, as these are the metrics that are essential to know when mortality and morbidity benefits cannot be found.	Thank you for your comment.
137	Peer Reviewer #3	General	This reviewer has a specific concern regarding the absence of diabetes as patient characteristic for Key Question 1, which seems to be a major omission.	We did not abstract comorbidity data as there are major issues with ascertainment of comorbidities particularly in retrospective cohort studies using registry data. Based on data from one study, the prevalence of diabetes may be underestimated by as much as 13% (https://www.ncbi.nlm.nih.gov/pubmed/28151871).
138	Peer Reviewer #3	Introduction	The Introduction is well-written and provides a nice, brief historical perspective how hemodialysis care in the US got to its current state. It also makes plain the need for this type of assessment.	Thank you for your comment.
139	Peer Reviewer #3	Methods	Inclusion and exclusion criteria are logical and justified. The search strategies are explicitly stated and robust.	Thank you for your comment.
140	Peer Reviewer #3	Methods	There appears to be a typo on Page 6, Table 2, Row "Population", Column "Exclusion criteria". says "All KQs: non-U.S participants constituting < 50% of study population". Either the "<" should be changed to ">" or the "non-US" should change to "US".	We have reviewed the document and corrected this typo.
141	Peer Reviewer #3	Results	Results are presented and discussed in a clear and systemic fashion. There was often extreme level of detail in discussion of the secondary outcomes of the trials. If this exhaustive detail is not necessary these sections could be condensed without sacrificing the important points of the assessment.	Thank you for the comment. Since we did not perform a meta-analysis, due to the heterogeneity of the studies, we felt it was important to synthesize details from individual studies as they do inform our thinking about future directions.

#	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 most important additional cardiovascular risk factor. It could replace smoking and education (currently listed in the tables). Diabetes prevalence will be reported in most if not all of the literature cited (as opposed to smoking and education). Diabetes prevalence should then be mentioned in the demographic comparisons of the studies to the USRDS population.	We did not abstract comorbidity data as there are major issues with ascertainment of comorbidities particularly in retrospective cohort studies using registry data. Based on data from one study, the prevalence of diabetes may be underestimated by as much as 13% (https://www.ncbi.nlm.nih.gov/pubmed/28151871).
143	Peer Reviewer #3	Results	Figure 5 is redundant. The data is already given in Table 4, and the relevant comparisons are made in Figure 6.	Figure 5 presents the overall racial distributions in the US dialysis population and Figure 6 provide race distribution by individual studies. They provide different information.
144	Peer Reviewer #3	Discussion/ Conclusion	The limitations and heterogeneity of the current literature is well described.	Thank you for your comment
145	Peer Reviewer #3	Discussion/ Conclusion	As a personal bias, this reviewer feels that the discussion for future research should have more emphasis on the last two bullets (lines 23 and 33) on page 84 (and that they should be listed first).	We have rearranged the research recommendations as suggested.
146	Peer Reviewer #3	Discussion/ Conclusion	On pg. 80, line 9-11, the authors make the important point that urea kinetics do not fully quantify the removal of other uremic toxins. Following this logic, determining optimal dialysis may require studies which measure the clearance of other surrogate biomarkers which might better correlate with mortality and cardiovascular outcomes. This would seem to be an important addition to the future research section.	We have added this to our research recommendations.
147	Peer Reviewer #3	Discussion/ Conclusion	Page 81. Line 18-38. The authors correctly point out that a major difference between the FHN and the TiME trials are residual renal function (and by extrapolation, dialysis vintage). This has implications for the comparison of patient characteristics to USRDS in the results sections. For TiME, a better comparison for this assessment would be USRDS patients who are within 120 days of dialysis initiation rather than the entire population.	We obtained information on the US dialysis population using the online reporting tool from USRDS (RenDER). We have updated figures to identify incident and prevalent populations.

#	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-dialysis blood pressure, ultrafiltration intensity, and post-dialysis recovery time.	The grading is based on specific guidance as explained in the methods section. The low level of evidence is due to lack of RCT data which is a major limitation in many dialysis studies.
152	Fresenius	General	The Discussion in the draft generally fails to synthesize the breadth of evidence about hemodialysis frequency and duration. There are multiple nuances in randomized clinical trials and observational studies that together suggest clearly efficacious roles of increased hemodialysis frequency and duration in the treatment of end stage kidney disease. We encourage the authors of the draft to review the November 2016 supplemental issue of the American Journal of Kidney Diseases, entitled "Intensive Hemodialysis: Potential for Improving Patient Outcomes." This issue includes six narrative reviews that focus on	We have added a section to the Discussion to better address your concern: "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 much higher than the rest of the world? J Am Soc Nephrol. 2009 Jul;20(7):1432-5. doi: 10.1681/asn.2009030282. PMID: 19443632. 169. Goodkin DA, Bragg-Gresham JL, Koenig KG, et al. Association of comorbid conditions and mortality in hemodialysis patients in Europe, Japan, and the United States: the Dialysis Outcomes and Practice Patterns Study (DOPPS). J

#	Commentator & Affiliation	Section	Comment	Response
			then-published randomized clinical trials of hemodialysis and large observational studies of hemodialysis frequency.	Am Soc Nephrol. 2003 Dec;14(12):3270-7. doi: 10.1097/01.asn.0000100127.54107.57. PMID: 14638926. 170. Yoshino M, Kuhlmann MK, Kotanko P, et al. International differences in dialysis mortality reflect background general population atherosclerotic cardiovascular mortality. J Am Soc Nephrol. 2006 Dec;17(12):3510-9. doi: 10.1681/asn.2006020156. PMID: 17108318. 171. Chapter 11: International Comparisons. Minneapolis, MN: USRDS Coordinating Center. https://www.usrds.org/2018/view/v2_11.aspx. Accessed on March 17, 2020.
153	Fresenius	Evidence Summary	In the first item of the Main Points section, the phrase "a few other clinical outcomes" lacks specificity. The Evidence Summary is likely to be quoted in future publications, so it is important that language is precise and explicitly delineates both benefits and harms of increasing hemodialysis and/or duration.	We revised the first bullet point: "More frequent hemodialysis, compared with usual care, was associated with improvement in total mortality, LV mass, blood pressure, and a few other clinical outcomes (low strength of evidence from 1 randomized controlled trial (RCT) and 3 observational studies)."
154	Fresenius	Evidence Summary	The importance of the observation that "[t]he mortality rate in RCT and observational studies was lower than the rate in the U.S. dialysis population" is unclear. It is well-known that participants in randomized clinical trials are generally healthier than non-participants. It is also well-known that the home dialysis patient population in the United States is, on average, younger and healthier than the in-center hemodialysis population. However, both randomized clinical trials and observational studies of hemodialysis frequency and duration do include elderly patients and patients with substantial comorbidity, especially cardiovascular comorbidity. We contend that the draft implies an exclusion of such patients from relevant studies, despite that most studies specified no such exclusion criteria. The applicability of individual studies should be evaluated based on more than mean values of baseline patient characteristics. To that point, the draft, both in the Evidence Summary and elsewhere, should better assess whether individual studies are likely or unlikely to be biased.	We have made several revisions throughout the report emphasizing the applicability of the studies. We believe that one of the biggest challenges with evidence in the dialysis space is lack of large RCTs. We have put emphasis on the small size of RCTs in our revised report.
155	Fresenius	Evidence Summary	"Usual care" is parenthetically defined by 3 treatments per week, with less than 4 hours per treatment. However, this is a problematic convention.	We arrived at the 4 hour threshold in consultation with our stakeholders (including CMS, technical expert panel, and key informants). We don't believe that the <4 hour versus >=4 hours cut-off needs to be changed. Furthermore, the level of evidence will still be insufficient.

#	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 Network (FHN) Nocturnal Trial, mean treatment duration in the thrice-weekly arm was 256 minutes (i.e., greater than four hours). This is an important observation, for it positions thrice-weekly treatment in the FHN Nocturnal Trial as an outlier, relative to most prescriptions of conventional hemodialysis in the United States.	We included the following footnote in Tables 1-3 to help explain how we sorted studies by duration of dialysis – * Usual care involves 3 treatments per week with an average of 3.5 hours per treatment and a minimum of 3 hours per treatment. We don't believe that the <4 hour versus >=4 hours cut-off needs to be changed. For studies addressing Key Question 3 or the combined Key Questions 2 and 3, duration per treatment ranged from 4 hours to 7.5 hours
156	Fresenius	Introduction	Regarding the so-called "Decision Dilemma," it is important to note, as Morfin et al (PMID: 27772642) did, that each hemodialysis treatment comprises not only treatment, but also time to travel between home and dialysis facility and post-dialysis recovery time. From this perspective, the idea that a conventional hemodialysis treatment consumes 4 to 6 hours is likely very conservative. A large reduction in post-dialysis recovery time may overwhelm an increase in dialysis treatment hours, due to the prescription of frequent and/or longer treatments.	We have changed 4-6 hours to "several hours".
157	Fresenius	Methods	Our primary concern with the methodology of the assessment is that many relevant studies of hemodialysis frequency and duration were excluded from consideration. The nephrology literature is well-known to include a paucity of randomized clinical trials. Furthermore, literature about hemodialysis frequency and duration remains relatively limited. A trustworthy review of relevant evidence should include all the highest-quality studies.	The Methods clearly describe our inclusion and exclusion criteria—we are looking at a US hemodialysis population and determined though discussions with technical experts that populations outside of the US may receive different treatments, or may not be similar enough to the US population to include. We have detailed the reason for exclusion for specific articles you point out in subsequent comments.
158	Fresenius	Methods	We contend that the restriction of evidence to United States patients is arbitrary and ultimately limits the utility of the assessment. There is no bona fide evidence that physiologic effects of frequent and/or longer hemodialysis vary among countries and patient genotypes, given a fixed comparator.	We have added a paragraph to the discussion section: "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…"

#	Commentator & Affiliation	Section	Comment	Response
			The effect of the restriction to United States patients is the elimination of two randomized clinical trials: the trial of frequent nocturnal hemodialysis by Culleton et al (PMID: 17878421) and the trial of thrice-weekly nocturnal hemodialysis by Jardine et al (PMID: 28151412). Given the paucity of randomized clinical trial evidence in all aspects of dialysis care, it is certainly arguable that exclusion of two randomized clinical trials from this assessment amounts to good science. In fact, we contend that the Centers for Medicare & Medicaid Services is well-served by synthesis of dialysis-related data from other countries, considering that other high-income English-speaking countries have historically rendered frequent and/or longer hemodialysis to greater shares of patients than the United States has.	168. Foley RN, Hakim RM. Why is the mortality of dialysis patients in the United States much higher than the rest of the world? J Am Soc Nephrol. 2009 Jul;20(7):1432-5. doi: 10.1681/asn.2009030282. PMID: 19443632. 169. Goodkin DA, Bragg-Gresham JL, Koenig KG, et al. Association of comorbid conditions and mortality in hemodialysis patients in Europe, Japan, and the United States: the Dialysis Outcomes and Practice Patterns Study (DOPPS). J Am Soc Nephrol. 2003 Dec;14(12):3270-7. doi: 10.1097/01.asn.0000100127.54107.57. PMID: 14638926. 170. Yoshino M, Kuhlmann MK, Kotanko P, et al. International differences in dialysis mortality reflect background general population atherosclerotic cardiovascular mortality. J Am Soc Nephrol. 2006 Dec;17(12):3510-9. doi: 10.1681/asn.2006020156. PMID: 17108318. 171. Chapter 11: International Comparisons. Minneapolis, MN: USRDS Coordinating Center. https://www.usrds.org/2018/view/v2_11.aspx. Accessed on March 17, 2020.
.. 159	Fresenius	Methods	The exclusion of studies without a comparison group is logical, as single-arm studies are uniquely subject to biased inference due to regression to the mean. However, in the context of hemodialysis frequency, the effect of the restriction is the elimination of one of the largest prospective cohort studies of frequent hemodialysis, the Following Rehabilitation, Economics and Everyday-Dialysis Outcome Measurements (FREEDOM) study. This study resulted in three peer-reviewed publications pertaining to patient-centered outcomes: a study of depressive symptoms and post-dialysis recovery time (PMID: 20673601), a study of restless legs symptoms and sleep disturbances (PMID: 21415315), and a study of health-related quality of life (PMID: 22622497). The authors of the report may choose to grade these studies as low-quality, but essentially disregarding these studies does not clearly strengthen the assessment. As an aside, the study by Troidle et al (PMID: 17603977) included sixteen patients who initiated in-center nocturnal hemodialysis for three sessions per week. Most outcomes in this	The FREEDOM study was included in KQ 4. It was not included in KQ 1-3 for lack of comparison group. See the last paragraph of the inclusion/exclusion section of the Methods. Troidle is a pre-post study, and therefore is included. The grading of evidence is based on specific guidelines provided in the AHRQ Methods Guide. The low level of evidence is due to lack of RCT data which is a major limitation in many dialysis studies. The low grading is not a pejorative or “degrading” of the studies as implied in this comment.

#	Commentator & Affiliation	Section	Comment	Response
			study are simply reported as mean values at baseline and again after six months of nocturnal treatment. There is no comparison group. If this study is included in the assessment, then by logical extension, all the FREEDOM study publications should be included.	
160	Fresenius	Methods	Another important issue is the apparent decision to separately evaluate Key Question (KQ) 2, Key Question 3, and the composite of Key Questions 2 and 3. The separate analysis of the composite of Key Questions 2 and 3 (i.e., effects of both frequent and longer hemodialysis) is logical, but inadvertently weakens the synthesis of data about frequency. The FHN Daily Trial, FHN Nocturnal Trial, and the excluded randomized clinical trial by Culleton et al (PMID: 17878421) are all tests of hemodialysis frequency. Furthermore, the authors should observe that even the FHN Daily Trial is a test of both frequency and duration, as mean session length in the intensive hemodialysis arm of the trial was not only shorter than with conventional hemodialysis in the United States, but also slightly shorter than what is typical of home hemodialysis in the United States. None of these trials can be regarded as tests of the effects of treatment frequency, ceteris paribus. The clinical reality is that increasing frequency of hemodialysis typically results in increasing cumulative duration of hemodialysis per week. We encourage the authors to revisit the aggregation of relevant evidence to answer Key Question 2 and Key Question 3.	The study report follows the format outlined in the study methods which were developed with extensive input from CMS, technical experts, and key informants, including representatives of Fresenius. While we could rearrange the grouping, the level of evidence will not change based on the grouping. The nocturnal trials (FHN and Culleton) were trials of frequency AND duration in a specific patient population, i.e. those willing to undergo nocturnal dialysis. The results from the nocturnal trials are therefore applicable to a similar patient population.
161	Fresenius	Methods	Although the publication date restriction is reasonable to preserve the applicability of the review to contemporary dialysis patients in the United States, it remains true that older literature includes observational evidence about associations of longer hemodialysis with outcomes. Importantly, intensive hemodialysis did not become “feasible” when the NxStage	Yes, the use of home hemodialysis was very low prior to the approval of NxStage dialysis machine in 2005, which is the reason for this restriction. We have replaced “became feasible” with “increased” in the sentence referred to by the reviewer.

#	Commentator & Affiliation	Section	Comment	Response
			System One was commercialized. Instead, utilization of frequent home hemodialysis increased after the commercialization of NxStage equipment.	
162	Fresenius	Methods	The inclusion criteria refer to the United States end stage renal disease “Medicare” population, but most studies do not attempt to quantify the distribution of payers. It is unclear why any reference to Medicare coverage exists in the methodology of this assessment.	The 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. 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.
163	Fresenius	Results	Regarding KQ 1, even in the context of the described methodology, there are multiple studies that are inexplicably omitted from the list of studies, including:	We will address each individual comment below.
164	Fresenius	Results	A key observational study by Weinhandl et al (PMID: 25085647). This is a very large retrospective study of all-cause and cause-specific hospitalization risks in daily home hemodialysis versus matched in-center hemodialysis patients with Medicare fee-for-service coverage. Importantly, this study reports strong associations of daily home hemodialysis with lower risk of hospital admissions principally attributable to cardiovascular disease, especially heart failure and hypertensive disease, and higher risk of admissions principally attributable to infection	We have added this study to our review. The population of this observational study overlaps the study by Miller, 2010 which was included in the draft report.
165	Fresenius	Results	A smaller observational study by Johansen et al (PMID: 19692997).	We have added this study to our review. The population of this observational study overlaps the study by Lockridge, 2011 which was included in the draft report.
166	Fresenius	Results	The randomized crossover trial by Laskin et al (PMID: 28389745). The reasons for its omission are unclear. This is the only randomized trial of hemodialysis frequency in the pediatric patient population.	This pilot trial of 6 children was excluded for Key Questions 1-3 as it did not meet the minimum six month follow-up period specified in the protocol. Data were abstracted for KQ 4.
167	Fresenius	Results	The prospective cohort study by Achinger et al (PMID: 23016876) is not included in the draft. This study assessed vascular access outcomes in patients undergoing short daily hemodialysis versus conventional hemodialysis and reached conclusions that	We have added this manuscript to the report.

#	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.	Lockridge is included. Kjellstrand is not. Our inclusion criteria specified that for multinational studies, the US population should be >=50% or the results should be stratified so that the US results can be abstracted. In the Kjellstrand paper, 40% of the population is from the U.S. and the results are not stratified by country. This led to the exclusion of the Kjellstrand paper.
169	Fresenius	Results	Again to the point of inclusion or exclusion of studies with patients not in the United States, it is admittedly unclear why the study by Hladunewich et al (PMID: 24525032) was retained, considering that most patients in the higher tertiles of hemodialysis hours per week resided in Canada, not the United States. We contend that inclusion is appropriate, as pregnant dialysis patients in Toronto are unlikely to respond to intensive hemodialysis in a manner that is dramatically different than similar patients in the United States.	Hladunewich: The US and Canadian data is stratified, so this study is eligible
170	Fresenius	Results	There are several minor, yet substantive issues in the draft that should be addressed:	We will address each individual comment below.
171	Fresenius	Results	Dialysate flow rates are available in most publications about randomized clinical trials of hemodialysis frequency and/or duration, but those results are not described in this report.	The focus of our study was the effects for duration and frequency of dialysis on outcomes. The dialysis flow rates contribute to solute clearance targets which were not the focus of this review.
172	Fresenius	Results	The authors of the report are incorrect in asserting that the number of screen patients in the FHN Daily Trial was not reported. Those	Thank you for pointing out this paper which was not cited in any of the FHN publications. We have added this information to the Discussion.

#	Commentator & Affiliation	Section	Comment	Response
			data were reported by Sergeyeva et al (PMID: 22505248).	
173	Fresenius	Results	Figure 4 does not depict a clear comparison. The histogram shows the distribution of age in the US hemodialysis patient population, but the superimposed bars merely show the range of mean age values across studies. The report should more completely describe the ranges of age in studies of frequent and/or longer hemodialysis.	Figure 4 (now Figure 5 in the revised report): we have revised the figure title: "Mean age distribution of hemodialysis patients in the U.S. in 2016, from the USRDS"
174	Fresenius	Results	Many tables of results list sample sizes. However, obscured in some tables of this draft is the number of patients undergoing frequent hemodialysis in each study. Table 8 is an excellent example. The study by Mathew et al (PMID: 27528548) included 50,756 patients, but only 160 underwent frequent hemodialysis. Meanwhile, the study by Weinhandl et al (PMID: 22362906) included 273,487 patients before a matching algorithm was executed and the analysis itself included 1,873 patients who underwent frequent hemodialysis and 9,365 matched patients who underwent conventional hemodialysis. We encourage the authors of the draft to revise tables to accurately depict which studies included the highest numbers of patients exposed to frequent hemodialysis, so that readers may understand which studies provide the highest precision.	We have added a figure (Figure 4) to the review to better detail the study populations. The specific data requested is available in the Appendices; specific appendices are called out in the report.
175	Fresenius	Results	Regarding KQ2 and "Mortality and Related Composite Endpoints," all composite endpoints that included death in the FHN Daily Trial were dominated by the non-death component. This is not adequately described in the draft.	The presentation of the results in this section clearly states that "The primary analysis of the FHN Daily Trial[26] did not have sufficient power to assess death as a primary endpoint during the 12-month study period when there were five deaths (4%) in the frequent dialysis arm and nine deaths (7.5%) in the conventional dialysis arm."
176	Fresenius	Results	Any discussion of extended follow-up in the FHN Daily Trial should be accompanied by a note that treatment frequency was not monitored after the conclusion of the protocol-specified 12-month follow-up interval.	We have revised the draft to make it clearer.
177	Fresenius	Results	In the description of Weinhandl et al (PMID: 22362906), cause-specific mortality hazard ratios for causes other than cardiovascular	We reported cardiovascular deaths and infection related deaths from Weinhandl and have now also described the findings for other cause-

#	Commentator & Affiliation	Section	Comment	Response
			death and infection are not described. Moreover, although the infection-related mortality hazard ratio is nonsignificant, it foreshadows the association of frequent hemodialysis with infection-related hospitalization risk, which is reported by Weinhandl et al (PMID: 25085647). Nevertheless, it is important to remark that the studies by Weinhandl et al conflated frequency and setting of hemodialysis. Infection-related mortality and morbidity may reflect the influence of dialytic setting and even cannulation technique (e.g., buttonhole cannulation) more than the influence of frequency.	specific deaths such as cachexia/dialysis withdrawal, other specified cause, and unknown (page 27). In our description of Weinhandl findings, we specifically note it is a study evaluating the effect of daily home hemodialysis. "More frequent daily home hemodialysis was associated with a lower risk of death due to cachexia or dialysis withdrawal (HR 0.63; 95% CI, 0.41 to 0.95) or an unknown cause (HR 0.59; 95% CI, 0.44 to 0.79). However, the effect of more frequent hemodialysis on cardiovascular disease mortality (HR 0.92; 95% CI, 0.78 to 1.09), infectious disease mortality (HR 1.13; 95% CI, 0.84 to 1.53), or other specified cause (HR 1.06; 95% CI, 0.81 to 1.37) was non-significant."
178	Fresenius	Results	The study by Brunelli et al (PMID: 26692402) suggests that the choice of hemodialysis machine and, correspondingly, dialysate flow rate may have little impact on clinical outcomes. This is not discussed in the report.	The objective of our report was to assess the effect of more frequent dialysis on outcomes. The comparative effectiveness of HHD systems (NxStage System One vs. 2008K@home) for more frequent dialysis or varying dialysate flow rates is beyond the scope of this project.
179	Fresenius	Results	The FHN Daily Trial (PMID: 23393319) found no significant effect of treatment frequency on arteriovenous fistula/graft loss (P = 0.58). This finding should be noted.	We did not include this as one of our primary outcomes.
180	Fresenius	Results	The report included several quotations of patients in Troidle et al (PMID: 17603977), which are negative in nature. The report fails to include other patient quotations in Troidle et al, which are positive in nature. Specifically, patients stated that "I have never felt this great on [hemodialysis]" and that "I get off the machine and resume activity right away."	We have removed the quotations from the results section.
181	Fresenius	Results	The TiME trial (PMID: 31000566) achieved a 9-minute difference in treatment time per session. The results of this study should be placed in proper context with observational studies that examined much larger differences in treatment time. Of course, it is unlikely that nine minutes per treatment will substantially alter clinical outcomes and quality of life. Thus, it is arguably disingenuous to cast the TiME trial as an assessment of treatment duration. It does, however, illustrate the	We have now included a description of the prescribed and actual treatment time per session for the TIME trial in the Description of Included Studies to provide this important context. "The TiME Trial randomized 266 facilities to assign their patients to usual care or hemodialysis sessions lasting 4.25 hours (255 minutes) or more, and then assigned patients at these facilities to their intervention strategy.[28] However, the implementation of dialysis sessions lasting 4.25 hours or more varied greatly by facility and only a small proportion of participants received this duration for the majority of dialysis sessions. For example, for the primary analysis group, the difference in the mean prescribed dialysis session duration and the mean delivered dialysis session duration between

#	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 magnitude of the reduction in LV mass was greater among patients with elevated LV mass at baseline (132 g or greater) (mean difference: - 22.7g; 95% CI, -36.7 to -8.7) compared with less than 132 g (mean difference: -3.6g; 95% CI, -12.4 to 5.2; p for interaction less than 0.0001)." However, this difference was not noted in the Nocturnal trial. We have now included the subgroup analysis for baseline LVM in the figure.
183	Fresenius	Discussion	In both the Results and Discussion sections, we harbor major concern about the lack of clear justification for grading the strength of evidence. The authors of the draft have apparently regarded all effects of hemodialysis frequency and duration as having low strength of supporting evidence. We agree that some domains, especially mortality and	The FHN Trials were inherently different enough for the trial investigators to treat them as separate trials and not combine the data. We believe that lumping the data together in a meta-analysis is an incorrect approach knowing that the study selection criteria and interventions were very different.

#	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.	
184	Fresenius	Discussion	We argue that the authors should re-assess their grade about the effect of increased hemodialysis frequency on pre-dialysis blood pressure. A reduction in pre-dialysis systolic blood pressure after initiation of frequent hemodialysis is one of the most robust effects observed in both the literature and clinical practice, yet this outcome is accompanied by low strength of evidence. To be certain, the authors could remark that the effects of hemodialysis frequency and/or duration on ambulatory blood pressure—which, unlike pre-dialysis blood pressure, is linearly associated	We followed AHRQ guidance for the grading of the evidence, and added more information about those methods in the methods section. We followed the GRADE criteria while making our assessments. These criteria were devised to ensure that evidence grading is unbiased.

#	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.	We have included a more detailed description of how the evidence was graded (see the revised Methods). To date, there is either insufficient evidence or low strength of evidence to support any interventions.
186	Fresenius	Discussion	In addition, the assertion of low strength of evidence that frequent and longer hemodialysis treatments lower ultrafiltration rate is difficult to accept at face value. The ultrafiltration rate is mathematically lower when treatment frequency and duration are both increased.	The low strength of evidence is driven by study characteristics and not the mathematical property of the ultrafiltration rate formula.
187	Fresenius	Discussion	Finally, the authors of the report state, "The FHN results can… only be generalized to prevalent hemodialysis patients with anuria and may not be applicable to incident hemodialysis patients with significant residual kidney function." That is a reasonable interpretation of the FHN Daily Trial, but of course, it is also worthwhile to note that increased hemodialysis frequency is a feature of home hemodialysis in the US, and home hemodialysis is typically prescribed to patients who have accumulated multiple years of in-center hemodialysis treatment. In that sense, the FHN Daily Trial and the practice of home hemodialysis in the US are closely aligned.	Thank you for your comment
188	Kidney Care Counsel	General	In our experience, patients are prescribed more frequent or longer duration dialysis treatments based on a clinical assessment of their acute and chronic conditions. Peer-reviewed studies find that numerous conditions, such as cardiovascular complications, are best managed with the prescription of additional or longer duration hemodialysis sessions.5 The National Kidney Foundation's Kidney Disease Outcomes Quality Initiative (KDOQI) Guidelines support	We have reviewed the KDOQI guidelines and the systematic review that accompanied those guidelines. The Clinical Practice Guidelines, developed by a panel of subject matter experts, includes both a synthesis of evidence and expert opinion. The quoted guideline (4.1.1) is an expert opinion and the review evidence for the opinion is considered "Not Graded".

#	Commentator & Affiliation	Section	Comment	Response
			the use of more frequent dialysis for both chronic and acute conditions. Specifically, the KDOQI Clinical Practice Guideline for Hemodialysis Adequacy, 2015 Update, Guideline 4.1.1 instructs physicians to "[c]onsider additional hemodialysis sessions or longer hemodialysis treatment times for patients with large weight gains, high ultrafiltration rates, poorly controlled blood pressure, difficulty achieving dry weight, or poor metabolic control (such as hyperphosphatemia, metabolic acidosis, and/or hyperkalemia)." We recommend AHRQ include a review the KDOQI Guidelines and underlying research and analysis in the Final Report.	
189	Kidney Care Counsel	General	The duration of the dialysis treatment is driven by clinical factors, as providers and the treating clinicians monitor patient outcomes and modify the duration to adequately and optimally dialyze the patient in the most efficient time frame possible. However, the duration of dialysis may also be affected by patient inputs, including some patients who decline to dialyze for the full time recommended by their treating clinicians. Unfortunately, the AHRQ-defined concept of "usual care" does not comport with clinical guidelines or practice experience of dialysis providers.	We fully agree that a multitude of factors contribute to the dialysis duration observed on dialysis and some of them may not be captured in the studies. However, the majority of the patients in the U.S. are dialyzed in-center thrice daily for less than 4 hours per treatment, which comprises the usual care received by hemodialysis patients in the U.S.
190	Kidney Care Counsel	General	The Draft Report indicates that "[a]ll interventions are compared with usual care (hemodialysis 3 treatments per week, less than 4 hours per treatment)."6 Unfortunately, AHRQ's defined concept of "usual care" does not represent the standard of care or dialysis provider experience. For example, a one-hour dialysis treatment would meet the definition of "usual care" as it is less than four hours per treatment. Similarly, the AHRQ definition of "usual care" implies that it would be "unusual" for a patient to dialyze for more than four hours, yet we find that patients with a higher Body Surface Area may regularly require four	We used the classification system to categorize "published" studies in different groups. The intent is not to recommend "usual care" as the "optimal care" strategy. We do not imply that anywhere in our report and have carefully reviewed the report to verify that.

#	Commentator & Affiliation	Section	Comment	Response
			to five hours of dialysis treatment. As a result, we believe the AHRQ definition of "usual care" likely includes treatments that are much shorter than the standard of care and excludes treatments that are slightly longer, but still within the standard of care.	
191	Kidney Care Counsel	General	The AHRQ definition of usual care is the basis through which all of the studies are reviewed for the Draft Report and is therefore of paramount importance. We are concerned that this definition may have distorted the conclusions drawn from AHRQ's literature review. We ask that AHRQ review, for example, the KDOQI Clinical Practice Guideline for Hemodialysis Adequacy, 2015 Update, Guideline 4.1, which recommends that "patients… undergoing thrice weekly hemodialysis be prescribed a bare minimum of 3 hours per session." We recommend that AHRQ revise the definition of "usual care" to better represent the standard of dialysis care and update the analysis of the literature review accordingly in a Final Report.	The KDOQI guideline recommended thrice weekly hemodialysis for a minimum of 3 hours per treatment, which is consistent with how we classified study interventions.
192	Kidney Care Counsel	Methods	AHRQ has employed a rigorous set of criteria for study inclusion and analysis that is likely built upon deep experience in meta-analysis across the healthcare system. The methodology excluded studies for which AHRQ determined there was no comparison group for outcomes, studies where less than 50 percent of the participants were from the United States, and studies that included non-hemodialysis patients. Included studies for KQs 1, 2, and 3 were limited to those published between January 1, 2005 and April 1, 2019. In addition, AHRQ employed screening mechanisms that focused on article titles and abstracts and required consensus among reviewers prior to inclusion. While we appreciate the academic rigor of AHRQ's analysis, several features of the ESRD population make these academic criteria difficult to achieve and may result in overlooking valuable studies that may provide	The rigorous set of study eligibility criteria were defined in the review protocol, and we followed those criteria in conducting the review. We disagree with changing the eligibility criteria now. In the discussion section, we acknowledge some other studies that did not meet the strict criteria for inclusion in the review.

#	Commentator & Affiliation	Section	Comment	Response
			insight regarding frequency and duration of hemodialysis treatment across modalities.	
193	Kidney Care Counsel	Methods	Randomized clinical trials comparing modalities, specifically those comparing outcomes for home hemodialysis and in-center dialysis, are difficult to accomplish in the ESRD population because patients cannot be randomly assigned to home hemodialysis. Home dialysis patients need to be physically stable enough to effectively dialyze at home and must have the physical and mental capacity to utilize home dialysis equipment and supplies. They must have a suitable home environment with enough space for their machine and supplies and their home must meet sanitation standards. Home dialysis patients need to have a support network, such as friends and family, to help if a problem arises. The choice of modality is ultimately a personal one for patients and not one that can be assigned at random without raising ethical and clinical concerns.	We recognize the factors influencing choice of home modalities including the ones listed here.
194	Kidney Care Counsel	Methods	Given these unique circumstances affecting the ESRD population, AHRQ, should consider the inclusion of a broader range of studies. For example, we recommend that AHRQ consider studies that may not have a "comparison group" because, as discussed above, such study models may be more limited in the ESRD context.	We specify Medicare population for a number of reasons. The scope of this review was intended to inform Medicare policy. Through numerous discussions with technical experts, it was determined that studies on populations that were not predominantly conducted in a US population would not help inform Medicare policy. The inclusion of a comparison group is a key consideration for rigorous epidemiological design of any study, including studies of patients on dialysis. We recognize the unique nature of the factors contributing to the choice of home dialysis and the difficulty in finding matching patients using registry data. Nevertheless, many studies have attempted this comparison, making the best use of available data. Furthermore, there are advanced statistical methods to account for selection bias and confounding which are widely used in epidemiological studies but have not yet been applied to studies of dialysis patients. We have included them as research recommendations.
195	Kidney Care Counsel	Methods	AHRQ should also reconsider assertions in the Draft Report of "low confidence" and "bias" when evaluating available studies and data. As discussed above, it is difficult to study certain aspects of the dialysis population	We have included a more detailed description of how the evidence was graded (see the revised Methods). To date, there is either insufficient evidence or low strength of evidence to support any interventions.

#	Commentator & Affiliation	Section	Comment	Response
			without some degree of "bias," for example a selection bias in the case of home hemodialysis modalities as discussed above. We support the acknowledgement of this inherent selection bias in the studies and literature. However, given the challenges in overcoming this selection bias, these studies likely represent the best available data and research for the ESRD population.	There are advanced statistical methods to account for selection bias and confounding which are widely used in epidemiological studies but have not yet been applied to studies of dialysis patients. We have included them as research recommendations.
196	Kidney Care Counsel	Methods	We are also concerned that AHRQ created a time limitation for KQs 1, 2, and 3 that is related to the release of a specific home dialysis machine and does not look more broadly at frequency and duration of dialysis treatments. Specifically, we note that AHRQ imposed a time limitation for the inclusion of studies for KQs 1, 2, and 3 "to those published between January 1, 2005 and April 1, 2019," under the theory that "more frequent dialysis is generally prescribed at home, and became feasible after the availability of the NxStage home hemodialysis machine in 2005."7 This statement suggests that in advance of their analysis, AHRQ conflated the examination of frequency and duration with not only the home hemodialysis modality, but with a specific dialysis machine. We recommend that AHRQ reconsider this time limitation and rationale and ensure that the Final Report is based upon a review of the literature regarding frequency and duration that includes available studies on both in-center and home hemodialysis modalities.	The time limitation is important to include studies relevant to contemporary dialysis population. The home hemodialysis population was less than 2000 patients in the U.S. prior to the approval of NxStage machine in 2005. The population had increased 4-fold to approximately 9,000 patients by 2014. So it is important to include studies which present the contemporary perspective on use of home hemodialysis. In the U.S., home hemodialysis remains the principal way of providing more frequent hemodialysis. However, our review includes *all* studies of dialysis frequency and duration, irrespective of whether the care was delivered at home or in-center. So our review includes "a review of the literature regarding frequency and duration that includes available studies on both in-center and home hemodialysis modalities" published since 2005.
197	Kidney Care Counsel	Methods	Applicability to Medicare ESRD Patient Population: The Draft Report indicates that the "reported studies had limited generalizability to the U.S. Medicare hemodialysis population."8 KCC is interested in how AHRQ was able to extract payer information from the studies analyzed. Our experience with many of these studies indicates that most do not report outcomes based on payers, such as Medicare. We therefore appreciate that there is a statement in the Draft Report that the	Thank you for this important comment. <u>Medicare population:</u> The 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. Based on our inclusion criteria, our results are generalizable to the U.S. hemodialysis population. We have carefully reviewed the report to ensure that this comes across as intended.

#	Commentator & Affiliation	Section	Comment	Response
			findings of the study have "limited generalizability to the U.S. Medicare hemodialysis population." We are concerned, however, that policymakers may attempt to make such generalizations and extrapolations to the Medicare population, nevertheless. We therefore ask that AHRQ make clearer in the Final Report where they are and are not able to identify Medicare patients in the reviewed studies and to underscore potential concerns with generalizing the findings of the Final Report to the Medicare hemodialysis population.	
198	Kidney Care Counsel	Reviewer conclusion	KCC appreciates the opportunity to provide comments to the Technology Assessment for Public Comment, Project Title: End Stage Renal Disease in the Medicare Population: Frequency and Duration of Hemodialysis and Quality of Life Assessment. We offer these comments with the goal of ensuring that AHRQ's Final Report provides an accurate and useful tool for policymakers as it relates to the delivery of complex, live-saving dialysis care.	Thank you for your comments
199	Kidney Care Counsel	Reviewer conclusion	Every day KCC members treat patients who have clinical needs for more frequent or longer duration hemodialysis. Unfortunately, the chronically underfunded ESRD PPS does not have the capacity to deliver unreimbursed, more frequent dialysis to Medicare beneficiaries. We have seen as recently as 2017, ill-advised proposed policies that would have presented serious barriers to the delivery of medically necessary care.	Thank you for your comment
200	Kidney Care Counsel	Reviewer conclusion	Every day KCC members treat patients who have clinical needs for more frequent or longer duration hemodialysis. Unfortunately, the chronically underfunded ESRD PPS does not have the capacity to deliver unreimbursed, more frequent dialysis to Medicare beneficiaries. We have seen as recently as 2017, ill-advised proposed policies that would have presented serious barriers to the delivery of medically necessary care.	Thank you for your comment

#	Commentator & Affiliation	Section	Comment	Response
201	Kidney Care Counsel	Reviewer conclusion	We appreciate the academic rigor with which AHRQ has sought to design and pursue this project related to ESRD. However, we believe that the Draft Report should be amended to afford policymakers a clearer picture of the relevant clinical data regarding the frequency and duration of dialysis for in-center and home hemodialysis patients. We ask AHRQ to consider our comments as well as comments submitted by others in the kidney community who are involved in the delivery of care on a daily basis. The stakes are high for AHRQ to get this right for Medicare beneficiaries living with ESRD.	Thank you for your comment
202	Renal Physicians Association	General	RPA appreciates the scope of work undertaken by AHRQ, particularly highlighting the disconnect between payment policies and treatment, but we have concerns with AHRQ's assessment of insufficient or low evidence to support more frequent dialysis or dialysis of increased duration. Furthermore, RPA is concerned that the negative effects of increased access issues were presented without grading the evidence.	Thank you for your comment, we will address the specific input and recommendations below. The grading of evidence is based on criteria set forth in the AHRQ evidence guide and applies to all systematic reviews. We recognize that dialysis evidence is limited, however, the criteria for grading the evidence remain the same. We have attempted to synthesize the literature in the best way possible and outline future directions to address unanswered questions. Vascular access: We have revised the abstract to align with the results which show that the strength of evidence is low for vascular access complications
203	Renal Physicians Association	Recommendations	Individualized therapy is key to providing patient-centered care. The provision of more frequent and/or intensive dialysis exemplifies this approach for patients with dialysis-dependent ESRD. Changes in delivery of care structures, reimbursement models and payment policies of CMS and other third-party payors to support therapies other than the dominant in-center, thrice-weekly, 3-4 hours treatment paradigm are needed. RPA's specific recommendations are provided below: 1. Longer and more frequent dialysis should be an option available to all patients for whom there is potential for clinical and quality of life benefits. Nephrologists should assess their	Thank you for your comments.

#	Commentator & Affiliation	Section	Comment	Response
			patient population for those patients who might benefit from more intensive hemodialysis. 2. Medicare, Medicaid, and commercial health insurers should adopt payment policies that increase the availability of more intensive dialysis therapies (either SDHD or NHD) to patients as prescribed by the patient's nephrologist. 3. Funding both in the form of Medicare reimbursement for dialysis through the ESRD Prospective Payment System (PPS) and research funding to NIH should be provided to advance technologies that promote the practice of more intensive hemodialysis, whether at home or in-center, for its convenience and cost effectiveness, but mostly for the clinical benefits it provides. 4. CMS and NIH should support well designed large RCTs to further evaluate this matter, but these studies should not preempt the provision of the best care possible based on current evidence.	