Comparative Effectiveness Research Review Disposition of Comments Report

Research Review Title: Serum Free Light Chain Analysis for the Diagnosis, Management, and Prognosis of Plasma-Cell Dyscrasias

Draft review available for public comment from December 16, 2011 to January 18, 2012.

Research Review Citation: Rao M, Yu WW, Chan J, Patel K, Comenzo R, Lamont JL, Ip S, Lau J. Serum Free Light Chain Analysis for the Diagnosis, Management, and Prognosis of Plasma Cell Dyscrasias. Comparative Effectiveness Review No. 73. (Prepared by the Tufts Evidence-based Practice Center under Contract No. 290-2007-10055-I.) AHRQ Publication No. 12-EHC102-EF. Rockville, MD: Agency for Healthcare Research and Quality. August 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

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

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review 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.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. 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.

Source: http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1229
Published Online: August 23, 2012

Commentator & Affiliation	Section	Comment	Response
Rex Astles (Public Reviewer)	Executive Summary	The acronym "KC" is used before it is defined.	Thank you. We have checked that all abbreviations (including KQ) are defined at first mention.
Peer Reviewer #1	Introduction	The review is clear and concise. The introduction flows well into the body of the paper and provides good connection to the importance of the key questions. The inclusion of the recommendations of the International Myeloma Workshop and other professional societies and identification of where evidence based medicine principles may be lacking to support those recommendations is particularly useful.	Thank you.
Peer Reviewer #3	Introduction	The introduction is well-written and sets the stage nicely for the study, providing the necessary background for the reader.	Thank you.
Peer Reviewer #5	Introduction	There are no issues with the Introduction.	Thank you.
Peer Reviewer #8	Introduction	This does lay the groundwork for the report	Thank you.
Peer Reviewer #9	Introduction	Okay	Thank you.
Peer Reviewer #1	Methods	Inclusion and exclusion criteria are justified. The search strategy is explicitly stated and logical. Definition criteria for outcome measures are appropriated as are statistical methods. One of the members of the AACC Evidence based Medicine Subcommittee pointed out that reference numbers in the text did not agree with citations.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Methods	There is an entity that has been defined as light chain MGUS [LC-MGUS] (Ref 63 in this document). These cases appear to progress to MM and associated diseases at the same rate as MGUS and therefore can be grouped into the PCDs along with what we have traditionally called MGUS. Since this is an entity that has no clinical symptoms and is defined only by an abnormal rFLC, it becomes difficult to think about false positives. If that is the case, then it should not be necessary to include only studies with all samples sent for testing and exclude studies that start with diagnostic samples obtained from patients with a diagnosis of PCD. Our own study (Ref 58 in this document) is the largest study to date and documents the increased diagnostic sensitivity of adding FLC to a number of diagnostic panels and indicates by diagnosis where there is increased sensitivity. By defining KQ1 as improving diagnostic sensitivity only in undiagnosed patients, the study unnecessarily restricts the data pool.	We thank the reviewer for mentioning the entity of LC-MGUS and agree that if the FLC ratio is the only abnormal test that signals the condition, it is difficult to think about false positives. This conundrum exemplifies the challenges surrounding evaluation of a test in monoclonal disorders. The existence of different disease groups implies that different tests will perform differently by context and have to be evaluated with this in mind. While the effectiveness of the SLFC assay is not in argument, this comparative effectiveness review addresses the *comparative* effectiveness of its diagnostic performance. It is likely that studies using diagnostic samples will yield inflated estimates and Key Question 1 sought evidence for the SFLC assay's diagnostic performance in populations without preexisting disease, as was suggested by the Technical Expert Panel. The FLC ratio has been suggested as part of a panel of tests for the diagnostic evaluation for a variety of symptoms and laboratory abnormalities such as anemia, hypercalcemia, or renal failure (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2904571/pdf/nihms211720.pdf). An increased FLC ratio without IgH positivity may also be associated with renal disorders and therefore may not be specific for patients with monoclonal gammopathies. The authors estimated that 23% of patients with an increased FLC ratio without IgH positivity either have or will develop renal disease. If LC-MGUS is identified in this context, they recommend (as for MGUS) further workup with bone marrow examination with clonality testing, imaging studies, renal biopsy, and/or Congo red staining of fat or other tissue as clinically appropriate to exclude a diagnosis of MM, amyloidosis, or related conditions. If such further workup is negative, they suggest repeating the SFLC assay in 6 months and yearly thereafter.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Methods	The search strategies are explicitly stated, and the search results clearly depicted by the figure. The excluded papers are appropriately listed in the Appendix. The inclusion and exclusion criteria for KQs 2-5 are justifiable. For KQ1, the restriction to only studies involving undiagnosed patients excludes several studies involving "already-diagnosed" patients known to have a PCD, where the relative value added by the sFLC assay over traditional testing is reported. Despite their biases, these studies have still informed clinical practice, and could be analyzed as a separate subset under KQ1, where their strengths and weaknesses could be addressed during the results and discussion similar to the other studies reviewed.	Undiagnosed patients were selected as the relevant population for Key Question 1 by the Technical Expert Panel. This was done to allow for a comparison of test accuracy among patients that were not preselected as having disease. We understand that studies that were not included in this review have already informed clinical practice, but the purpose of this comparative effectiveness review was to perform a comparative effectiveness analysis among tests in the population of interest. Thank you for your comment.
Peer Reviewer #3	Methods	The outcome measures listed on page 6 are all appropriate, with the exception of KQ5. As stated above, the need for 24 hour urine collection could be considered as a relevant outcome. "Hospital stays" may be an unrealistic outcome to measure for a laboratory test (would performing a urine immunofixation be expected to shorten a hospital stay?). Bone scans should be corrected to bone surveys.	We thank the reviewer for evaluating the relevance of the outcomes we have compiled. We agree that there are multiple variables and steps influencing the length of hospital stay and the contribution of a given diagnostic laboratory test would likely be insignificant. We have removed it from the list of outcome measures; and indeed, we found no literature that addressed the relationship. While elimination of the need for 24-hour urine collection could well be considered an outcome for Key Question 5, we found that the papers that addressed it did so in a diagnostic testing framework, not as an explicit outcome. Hence these papers were reviewed under the framework of Key Question but were ultimately not eligible for inclusion, as they were performed only in patients with preexisting diagnoses of PCDs. We have changed "bone scan" throughout the report to ''skeletal survey,'' as suggested. Thank you.
Peer Reviewer #5	Methods	I realize that this is beyond the time line for the article search but this paper is very valuable. It points out the SLFC ratio is not reliable for determining remission status. Paiva et al Comparison of immunofixation, serum free light chain, and immunophenotyping for response evaluation and prognostication in multiple myeloma. J Clin Oncol. 2011 Apr 20;29(12):1627-33. Epub 2011 Mar 14.	This paper was identified by our updated search and indeed has been included in the comparative effectiveness review. Thank you.
Peer Reviewer #8	Methods	The selection of manuscripts/ publications used for evidence is well stated.	Thank you.
Peer Reviewer #9	Methods	Okay	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Results	Details are clearly presented. Characteristics of studies are clearly described. Key messages are explicit and applicable. Tables are descriptive and add clarity to the report. The inclusion of a discussion of the different types of bias in a study was particularly helpful.	Thank you.
Peer Reviewer #2	Results	The analysis indicates that there are no publications that compared the FLC with traditional tests to determine if FLC predicts for progression to MM. In our study of MGUS progression (Ref 64 in this study) we have compared FLC to a number of other predictors. The 3 prognostic factors that retained significance in univariate analysis were M-spike size, heavy chain isotype, and rFLC; they each increased the risk ~2.5-fold. In addition, they were all independent prognosticators in multivariate analysis. This large study relied on retrospective clinical data, and the weakness of the study was that most MGUS patients don't get a bone marrow biopsy. Although that means that some factors are not available, it also means that only real-world factors were evaluated.	We thank the reviewer for pointing out this study (http://bloodjournal.hematologylibrary.org/content/106/3/812.full.pdf), which was included in our full-text review. We agree that this was a large retrospective study of patients with MGUS in whom the SFLC ratio as well as M protein levels and the heavy chain isotype were independent prognostic factors for progression to MM. However, no direct comparisons of the predictive ability of these tests were carried out, limiting any comparative inferences.
Peer Reviewer #2	Results	One conclusion of the study on the use of FLC for assessing risk of MGUS progression to MM (Ref 64) is that 40% of MGUS patients have a very low risk of progression (0.1%/yr). With this extremely low life-time risk, it is not necessary to do periodic laboratory assessments of M-spike progression. Followup testing was omitted as an intervention in the definition of KQ5, but the use of FLC, SPEP (M-spike), and IFE (heavy chain isotype) reduces the number of MGUS patients needing periodic phlebotomy, testing, and associated anxiety.	We thank the reviewer for calling attention to this potential use of the SFLC ratio in a composite model including M protein levels and the heavy chain isotype for risk stratification of MGUS patients. It is clear that the addition of these two variables refines prediction, although a comparative evaluation is not available. The frequency of follow up testing in the various risk groups is certainly an area for more research.
Peer Reviewer #3	Results	The authors provide good analysis of the included studies, with clear discussion of the major findings and clear justification for the quality assessments for each study. The tables are appropriate and well-done.	Thank you.
Peer Reviewer #4	Results	Table 4 It would be helpful to add 95% CI for sensitivity and specificity. In addition to total sample size, it would be helpful to add the number of patients with the diagnosis.	95% confidence intervals were not provided in these papers for sensitivity and specificity, but we have included intervals from the studies that provided enough data to calculate them. The number of patients with the diagnosis is included in Table 3 under the column labeled "PCD prevalence."
Peer Reviewer #4	Results	Table 7 Add P-value for study Sanchorawala, 2005?	P values have been added where available.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Results	Amyloidosis is a very rare and often difficult disorder to diagnose. SFLC are useful for establishing B cell clonality when the workup for a monoclonal protein is negative. So, when amyloidosis is found but the tissue staining is negative for light chains as well as the protein and urine being negative or equivocal for an M Spoke, the SFLC can be helpful for establishing the diagnosis and treatment. This however is clinical opinion and there is no concrete prospective data that definitively supports this approach. Otherwise, the data is analyzed well and presented in a logical manner.	Thank you.
Peer Reviewer #8	Results	Most of the relevant studies have been included. The assessment of relevance of the individual manuscript does not follow a consistent pattern.	Thank you. We have listed the reasons articles were deemed not relevant (were excluded) in Appendix B, as well as clarified throughout what exactly our inclusion criteria were.
Peer Reviewer #9	Results	Okay	Thank you.
Peer Reviewer #1	Discussion/ Conclusion	The major finding was an improved sensitivity as compared with traditional methods without any findings regarding patient outcome. The clearly defined key questions and outlined requirements for a study to evaluate those key questions will translate easily into new research.	Thank you.
Peer Reviewer #3	Discussion/ Conclusion	The implications of the findings are stated clearly and the limitations of the studies are well-discussed. The future research section is well-constructed and lays out a framework for more definitive studies to address the key questions.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Discussion/ Conclusion	Several studies that were excluded from the review could merit more discussion in this section, given their impact on clinical practice. For KQ2 this includes the Dispenzieri Lancet 2010 paper and the Rajkumar Blood 2005 paper, which used the FLC assay to define a new category of MGUS and define a new risk-stratification scheme for MGUS, respectively. For KQ4, this includes the AL amyloidosis studies exploring the prognostic value of FLC levels (Dispenzieri Blood 2006; Kumar Blood 2010). The importance of the FLC assay in assessing response in AL amyloidosis has recently been validated in a large retrospective international cohort (Palladini et al, ASH 2010; Abstract #1364), and forms the basis for the new International Society of Amyloidosis Consensus Response criteria.	We agree on the importance of many articles that were ultimately excluded; however, we were working within the defined boundaries of the Key Questions and the scope of this comparative effectiveness review. All excluded studies are listed, along with a reason for exclusion, in Appendix B. We did not include abstracts, per the a priori exclusion criteria. Such material, once published would lend itself to updates and fulfillment of future research needs.
Peer Reviewer #5	Discussion/ Conclusion	This is well written. I am not sure I found it in the discussion but a lot of tests are adopted in medicine and only when widely used, evaluated for utility. From a clinician's standpoint, SFLC are useful when establishing a diagnosis but not necessarily for following a patient. The IMWG criteria for progression and relapse only use SFLC testing in the absence of any other signs of M protein. Of note, it might be worth mentioning the "Heavy Light" test that is being developed the the Binding Site as a future light chain test for evaluation. Finally, BMT CTN 0702 will be prospectively collecting serum samples for light chain analysis along with flow cytometry measurement of bone marrow with the traditional tests for M protein. This should provide some useful information for the role or not of SFLC in monitoring MM patients.	We thank the reviewer for providing information on the BMT CTN 0702 study and appreciate the comment that many tests in current use are often objectively evaluated for utility after they come into general use. The goals of such assessment should then include a refinement of the indications for their use that could lead to recommendations to optimize their use. Such evaluations should take into consideration different clinical settings and phases of disease as well as different disease populations and subgroups.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Discussion/ Conclusion	The conclusions are rather vague. It is clear that the role of the assay still remains to be better defined in most of the conditions. However, its utility as a prognostic marker has been shown in several studies in various stages. It's utility in non secretory myeloma or light chain myeloma is clearly shown. Finally, it is the most useful test in patients with amyloidosis.	We appreciate the reviewer's comments and have refined the conclusions. The purpose of the comparative effectiveness review was not to cast doubt on the utility of the assay—which, as pointed out by the reviewer, has been shown in several studies. The comparative effectiveness review focused on specific areas where comparative performance of the assay might be relevant. Disease categories such as NSMM, LCMM, and amyloidosis are particularly relevant for such a comparative review where the findings would have clinical relevance, as again pointed out by the reviewer. However there was a paucity of such comparative data in these categories, indicating a future research need.
Peer Reviewer #9	Discussion/ Conclusion	Okay	Thank you.
Karen H. van Hoeven, The Binding Site (Public Reviewer)	Conclusion	Thank you for the opportunity to comment on the draft Comparative Effectiveness Review on Serum Free Light Chain (SFLC) Analysis. Key conclusions were, "We did not find sufficient evidence to determine whether the addition of the SFLC assay to traditional testing would increase the diagnostic accuracy of PCD or whether it would help prognosticate the disease course." Because of the narrowly framed key questions, the number of publications available to analyze was reduced by > 99% (from 2819 to 13 citations). Moreover, > 95% of the articles selected for full screening were rejected, and none of the accepted articles were grade A according to the Quality Assessment EPC Program definitions. It is disappointing that this report − that was intended to organize knowledge and evaluate evidence − rejected the overwhelming majority of the available evidence because of very narrowly defined parameters, and therefore the findings were inconclusive.	We agree on the importance of many articles that were ultimately excluded; however, we were working within the defined boundaries of the Key Questions and the scope of this comparative effectiveness review. All excluded studies are listed, along with a reason for exclusion, in Appendix B.
Karen H. van Hoeven, The Binding Site- (Public Reviewer)	Conclusion	Another key conclusion of the AHRQ report was that future research is needed. However, given that the IMWG and NCCN guidelines have significantly changed and improved clinical practice with the incorporation of sFLC analysis, it will be nearly impossible to design a screening study that compares the SFLC assay with urine protein electrophoresis without introducing substantial bias.	Thank you for this comment. While we are aware that the SFLC assay is increasingly incorporated into clinical practice, but we believe that it would be possible to design a screening study that compares the SFLC assay to urine protein electrophoresis, if both are done in an individual patient as part of a research protocol.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Clarity and Usability	The report is particularly well structured and organized. Main points are clearly presented. Since the findings were inconclusive the conclusions cannot be use to inform policy and/or practice decisions but can guide the design of future research.	Thank you.
Peer Reviewer #3	Clarity and Usability	The report is well-structured, organized clearly, and easy to follow.	Thank you.
Peer Reviewer #5	Clarity and Usability	Yes. It would be worth mentioning that SFLC may be part of the original work up and there is no utility for using to measure response if other tests are available for clonality. The test may be used only in the absence of other measures of M protein clonality in blood, urine or bone marrow. Agree emphatically for the need for prospective trials and long term followup regarding the use of SFLC and likely the new Heavy Light test.	Thank you.
Peer Reviewer #8	Clarity and Usability	Reasonable well organized	Thank you.
Peer Reviewer #9	Clarity and Usability	well organized	Thank you.
Rex Astles (Public Reviewer)	References	The reference numbers don't match the citations.	Thank you. We have checked the reference list. There are separate lists (per requirements) for the Executive Summary and for the main report.
Peer Reviewer #1	General	Quality of the Report: Superior Number of Hours Spent to Review the Report: 12 hours	Thank you.
Peer Reviewer #1	General	Even though the finding is one of insufficient data, AACC Evidence Based Medicine Subcommittee finds this study to be quite meaningful in identifying what is currently lacking in studies. Key questions are appropriate and explicitly stated. Shareholders were adequately involved in the formation of key questions. It is tempting to wish that different comparators had been selected, ones that would have allowed greater leeway in literature inclusion and grading. However, this would have compromised evidence based medicine principles. The clarity of the key questions and strict evaluation of literature will serve as a guide to better research planning in this area in the future. The target population and audience are explicitly defined.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	General	The introduction is very good and shows an excellent understanding of plasma cell proliferative diseases. I think, however, there are a few flaws in the study design and results, and therefore the conclusions as well.	We have reviewed the methods, results, and conclusions for possible flaws and have corrected them.
Peer Reviewer #2	General	Our clinical practice has incorporated FLC assays into testing for PCD. I understand the approach to evidence based conclusions, but the structure of the study did not allow the benefits of serum FLC assays to become evident.	As mentioned above, the purpose of this comparative effectiveness review was not to evaluate the clinical effectiveness of the SFLC assay but its comparative effectiveness in defined populations and settings. The test has been widely incorporated into diagnostic testing protocols for PCD and assessment of its effectiveness, but this use was outside the scope of the review.
Peer Reviewer #3	General	Quality of the Report: Good Number of Hours Spent to Review the Report: 7	Thank you.
Peer Reviewer #3	General	This systematic review seeks to assess the role of the serum free light chain (sFLC) assay in the diagnosis, prognosis, and management of plasma cell dyscrasias (PCDs): in particular, to determine the additional value gained by adding the sFLC assay to standard testing compared to standard testing alone. The audience/target population for the report is well-defined. The key questions are narrowly-drawn but appropriate and explicitly stated, with the exception of KQ5, where the "other interventions" are not clearly defined. For example, a reduction in the need for 24 hour urine collections for UPEP and urine immunofixation is one of the benefits claimed for the sFLC assay (as the authors note), but this "intervention" is not included as one of the outcome measures for KQ5, despite several studies in the literature addressing this question.	We thank the reviewer for these comments. We agree that with regard to Key Question 5, "other interventions" could include the need for 24-hour urine collection for UPEP and UIFE. However, as we explain in the review, we found that the papers that addressed this question did so in a diagnostic testing framework. Thus, we reviewed them in the context of Key Question 1; since the studies were in patients with preexisting diagnoses of PCDs rather than those not yet diagnosed, the articles were ultimately not eligible for inclusion.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	General	The authors utilize very stringent criteria to identify studies worthy of inclusion, with 13 of 290 potentially relevant studies ultimately reviewed. This underscores their main point that there is a paucity of high-quality data on the comparative effectiveness of these assays, and reflects their dedication to maintain the report's focus on comparative effectiveness, rather than effectiveness in general. However, this narrow focus somewhat undermines their stated objective on page ES-3: "to summarize the existing literature regarding the role of SFLC testing in the diagnosis, management, and prognosis of patients with PCDs." By excluding a number of studies that have demonstrated utility for the sFLC assay in these settings, they give the reader the impression that the assay provides little to no value in managing these patients, which is not reflective of real-world clinical practice (particularly for light-chain myeloma and AL amyloidosis). This limits the applicability and clinical relevance of the reported findings.	We thank the reviewer for highlighting the key goal of this comparative effectiveness review and understand that the consequence was the demonstration of a paucity of high-quality data on the one hand but a narrowed focus on the other. The criteria used for including studies were that the studies be comparative rather than simply assess the utility of the SFLC assay in different settings. The purpose of the report was to provide a summary of comparative effectiveness, as per AHRQ guidelines. The report notes that there is a paucity of data for comparative effectiveness despite the substantial body of data supporting the use of the SFLC assay for clinical use in general. The applicability and clinical relevance of the report include raising awareness about the lack of comparative data so the clinician or provider will interpret the test with this caveat. We anticipate that with the rapid evolution of PCD management, the need for comparative evaluation of tests in different clinical settings will gain importance.
Peer Reviewer #4	General	There was no meta-analysis conducted in this report due to heterogeneity in the designs, populations and comparisons in the included studies. Also the number of studies for each KQ is small, too and the quality of the studies is generally poor. I agree with the author's decision on not conducting any meta-analysis.	Thank you.
Peer Reviewer #4	General	To report the quality of the studies, it would be clearer to report "Good", "Fair" and "Poor", instead of Quality A, B and C.	Our EPC has always used a three-level grading system, with the grades denoted as either good, fair, or poor, or A, B, and C (respectively). Our experience is that the use of A, B, and-C elicits a less emotional response from readers and reviewers. In contrast, for strength of the body of evidence, we use a four-level system (high, moderate, low, and insufficient). We have retained these uses in this report.
Peer Reviewer #5	General	Quality of the Report: Superior Number of Hours Spent to Review the Report: 4	Thank you.
Peer Reviewer #5	General	The report is clinically meaningful. There has been a rush to use serum free light chains for diagnosis and monitoring of patients with plasma cell disorders. The questions are important and the key questions are stated explicitly.	Thank you.
Karen H. van Hoeven, The Binding Site	General	The International Myeloma Working Group (IMWG) in 2009 was composed of almost 100 of the world's most experienced, knowledgeable and authoritative	We thank the public reviewer for highlighting the important role played by the guidelines and IMWG members in streamlining the diagnostic, therapeutic, and monitoring

Commentator & Affiliation	Section	Comment	Response
(Public Reviewer)		hematologists/oncologists, many of whom devote their practice exclusively to the diagnosis and treatment of multiple myeloma and other plasma cell disorders. Thirty-three IMWG members co-authored the "International Myeloma Working Group guidelines for serum free light chain analysis in multiple myeloma and related disorders."1 After evaluating the available evidence, recognizing the importance of cost effectiveness, acknowledging the impracticalities of 24-h urine collection, reviewing pathophysiologic principles such as the median SFLC concentration at which overflow "Bence Jones" proteinuria occurs, and supplementing these data with their clinical observations based on experience, the IMWG guidelines concluded: §☐ The SFLC assay in combination with serum protein electrophoresis and serum immunofixation yielded high sensitivity when screening for multiple myeloma and related disorders and eliminated the need for 24-h urine studies for diagnoses other than light chain amyloidosis §☐ The SFLC assay was of major prognostic value in virtually all plasma cell disorders §☐ The SFLC assay was useful for quantitative monitoring of oligosecretory plasma cell disorders and reduced the need for frequent bone marrow biopsies in this setting §☐ Periodic urinary or SFLC measurements should be performed to detect light chain escape §☐ SFLC analysis should be performed as part of the criteria for stringent complete response. The National Comprehensive Cancer Network incorporated many similar statements into their 2011 guidelines.	approaches to PCDs and underscoring the utility of the SFLC assay in practice. Our comparative effectiveness review should not be construed in any way as a rebuttal of or challenge to those guidelines. Rather, it sought to answer a set of questions with a relatively narrow focus that compared the SFLC assay with traditional testing in very specific clinical settings where there may have been a value in making the comparison, by virtue of the biological basis of the assay. These questions were vetted by panels of Key Informants and Technical Experts who assisted in identifying the areas of focus. We found insufficient evidence to address those areas, indicating a need for targeted research in the future. We also found that much of the available research did not meet stringent reporting standards; this finding should inform the conduct of future studies. With regard to the individual IMWG guidelines, we respectfully respond to the comments: *§ The SFLC assay in combination with serum protein electrophoresis and serum immunofixation yielded high sensitivity when screening for multiple myeloma and related disorders and eliminated the need for 24-h urine studies for diagnoses other than light chain amyloidosis.* In this comparative effectiveness review, we identified only three studies of the added value of the FLC assay in undiagnosed populations compared to traditional testing. The emphasis on undiagnosed populations is relevant, since the guideline refers to *screening* for PCDs. Moreover, the guideline does not invoke any comparator and refers to the utility of the assay. *§ The SFLC assay was of major prognostic value in virtually all plasma cell disorders.* The guideline statement refers to utility of the assay and is noncomparative; thus it was not addressed in our comparative effectiveness review. *§ The SFLC assay was useful for quantitative monitoring of oligosecretory plasma cell disorders and reduced the need for frequent bone marrow biopsies in this setting* This guideline is noncomparative; thus it was not addressed in our comparative effectiveness review. Also, the role of the SFLC assay relative to bone marrow examination is not established and is opinion based.

Commentator & Affiliation	Section	Comment	Response
			§ Periodic urinary or SFLC measurements should be performed to detect light chain escape. This is a very relevant guideline given the changes in disease behavior in response to chemotherapy. It is clearly an area where future research is critical. *§ SFLC analysis should be performed as part of the criteria for stringent complete response.* This guideline is a noncomparative statement; thus it was not addressed in our comparative effectiveness review. It is most likely opinion based, as we did not find evidence that complete response with or without the SFLC ratio criteria is prognostic for progression-free survival or overall survival or that stringent complete response correlates with bone marrow response.
Karen H. van Hoeven, The Binding Site (Public Reviewer)	General	Even before the widespread adoption of these guidelines into clinical practice in the U.S. and abroad, the initial screening blood sample was accompanied by a urine sample in (at most) 52% of instances.[2] In one U.S. center, the estimate was 35%.3 With the implementation of these guidelines, these estimates are now likely to be substantially lower. Moreover, two studies recently independently reported that urine protein electrophoresis was inaccurate in 20% to 30% of samples.[4-5] Several U.S. laboratories severely restrict access to urine protein electrophoresis based on the accumulated data and their experience with the inaccuracies of the test.	We thank the public reviewer for highlighting the practical difficulties associated with obtaining a urine sample and the tremendous value that the FLC assay will bring if it could obviate the need for urine collection. Although the comparative diagnostic performance of the SFLC assay and UPEP or UIFE has been shown in patients with preexisting disease, it has not yet been shown in undiagnosed populations, where the danger of false negatives for the SFLC assay has not been thoroughly vetted. More study is needed in this regard.
Peer Reviewer #8	General	Quality of the Report: Good Number of Hours Spent to Review the Report: 4	Thank you.
Peer Reviewer #8	General	The report does address the important questions. However, the report fails to highlight the clinical utility of FLC assay in patients with very small amounts of intact immunoglobulin or only light chains and in patients with amyloidosis where it may be the most relevant marker.	We appreciate these comments and have refined our conclusions. The purpose of the comparative effectiveness review was not to cast doubt on the utility of the assay—which, as pointed out by the reviewer, has been shown in several studies. Rather, the comparative effectiveness review focused on specific areas where comparative performance of the assay might be relevant. Disease categories such as NSMM, LCMM, and amyloidosis are particularly relevant for identifying what findings would have clinical relevance. However, there was a paucity of comparative data for these PCD categories, indicating an area of future research need.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #9	General	Quality of the Report: Good Number of Hours Spent to Review the Report: 1	Thank you.
Peer Reviewer #9	General	Okay	Thank you.