Effective Health Care Program

Serum Free Light Chain Analysis for the Diagnosis, Management, and Prognosis of Plasma Cell Dyscrasias: Future Research Needs

Agency for Healthcare Research and Quality

Advancing Excellence in Health Care • www.ahrq.gov

Future Research Needs Paper

Number 23

Serum Free Light Chain Analysis for the Diagnosis, Management, and Prognosis of Plasma Cell Dyscrasias: Future Research Needs

Identification of Future Research Needs From Comparative Effectiveness Review No. 73

Investigators:
Madhumathi Rao, M.D., Ph.D.
Jenny L. Lamont, M.S.
Jeffrey Chan, B.S.
Thomas W. Concannon, Ph.D.
Raymond Comenzo, M.D.
Sara J. Ratichek, M.A.
Esther E. Avendano, B.A.

This report is based on research conducted by the Tufts Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10055-I). The findings and conclusions in this document are those of the author(s), who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care researchers and funders of research make well-informed decisions in designing and funding research and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of scientific judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical research and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials noted, for which further reproduction is prohibited without the specific permission of copyright holders.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.hhs.gov.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

An important part of evidence reports is to not only synthesize the evidence, but also to identify the gaps in evidence that limited the ability to answer the systematic review questions. AHRQ supports EPCs to work with various stakeholders to identify and prioritize the future research that is needed by decisionmakers. This information is provided for researchers and funders of research in these Future Research Needs papers. These papers are made available for public comment and use and may be revised.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The evidence reports undergo public comment prior to their release as a final report.

We welcome comments on this Future Research Needs document. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang M.D., M.P.H.
Director, EPC Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Elisabeth U. Kato, M.D., M.R.P.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the stakeholders for their contributions to this project.

Contributors

Contents

Appendixes
Appendix A. Search Strategy for Ongoing Studies
Appendix B. Lists of Ongoing Studies

Executive Summary

Background

Plasma cell dyscrasias (PCDs) are a group of clonal disorders characterized by the uninhibited expansion of a monoclonal population of malignant plasma cells.[1] Plasma cells arise from B cells in the bone marrow and produce immunoglobulins that constitute the body's normal humoral immune response. The immunoglobulin molecule is composed of a heavy chain and a light chain. Plasma cells normally produce light chains in excess that do not bind to heavy chains to form a complete immunoglobulin molecule and instead enter the bloodstream as free light chains (FLCs). In PCDs, each abnormally expanded clone of malignant plasma cells produce an excess of either intact immunoglobulin or FLCs of a single type called a monoclonal protein (M-protein) or paraprotein.

The serum FLC (SFLC assay (the Freelite™ Assay, The Binding Site Ltd., Birmingham, United Kingdom) was introduced in 2001 to measure the FLC component in particular.[2] The SFLC assay works by recognizing an epitope that is detectable only on light chains that are not bound to the heavy chain of the immunoglobulin molecule (i.e., FLCs) in the serum. This is the sole SFLC assay approved by the U.S. Food and Drug Administration. It detects low concentrations of FLCs and can measure the ratio of kappa chains to lambda chains.

It has been suggested that the SFLC assay could play an adjunctive role in screening, diagnosis, monitoring, and prognosis of PCDs in high-risk populations. The International Myeloma Working Group (IMWG) currently considers the SFLC assay to be an adjunct to traditional tests.[3] The assay could allow for quantitative monitoring of response and remission after treatment and provide prognostic information,[4,5] potentially reducing the need for frequent bone marrow biopsy for purposes of quantifying plasma cells, which is required as part of stringent monitoring for monoclonal gammopathy of undetermined significance (MGUS) progression to multiple myeloma (MM) or defining disease remission,[3] and potentially could be used in conjunction with serum protein electrophoresis (SPEP) and serum immunofixation electrophoresis (SIFE) to replace urine tests that require 24-hour collection (urine protein electrophoresis [UPEP] and urine immunofixation electrophoresis [UIFE]), which could simplify diagnosis and disease monitoring.[3,6] The SFLC assay may also be the only means of detecting a disease marker in some disease settings: nonsecretory MM, where SFLCs are often the only marker of the disease[7]; AL amyloidosis (systemic amyloidosis in which amyloid [A] proteins derived from immunoglobulin light chains [L] are deposited in tissue), where low monoclonal protein (M-protein) concentrations may not be detected by means of conventional techniques; and light chain MM, where the M-protein consists only of FLCs. These diagnostic applications have yet to be validated and standardized.[8] Thus, although the SFLC assay has been in use for a decade, it remains unclear how best to incorporate it into clinical practice.[9]

The 2012 Agency for Healthcare Research and Quality (AHRQ) comparative effectiveness review (CER) upon which the current Future Research Needs (FRN) project is based (available at www.effectivehealthcare.ahrq.gov/ehc/products/264/900/Serum-Free-Light-Chain_Draft-Report_20111215.pdf) reviewed pertinent publications through April 2012 regarding the role of the SFLC assay in combination with traditional testing by electrophoresis and immunofixation, compared with traditional testing alone, in the diagnosis and management of patients with PCDs. This CER examined evidence addressing five Key Questions.

- Key Question 1: Does adding the SFLC assay and the kappa/lambda ratio to traditional testing (SPEP, UPEP, SIFE, or UIFE), compared with traditional testing alone, improve the diagnostic accuracy for PCDs (MGUS, MM, nonsecretory MM, or AL amyloidosis) in undiagnosed patients suspected of having a PCD?
- Key Question 2: As compared with traditional tests, how well does the SFLC assay independently predict progression to MM in patients with MGUS?
- Key Question 3: In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay result in different treatment decisions as compared with traditional tests?
    - Does the use of the SFLC assay affect the management of patients by allowing for earlier institution of specific therapies?
    - Does the use of the SFLC assay influence the duration of treatment?
    - Does the use of the SFLC assay influence the type of treatment (e.g., radiation therapy)?
- Key Question 4: In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), is the SFLC assay better than traditional tests in indicating how the patient responds to treatment and of outcomes (overall survival, disease-free survival, remission, light chain escape, and quality of life)?
- Key Question 5: In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay reduce the need for other interventions (e.g., bone marrow biopsy)?

Findings of the CER

Results and conclusions from the CER are based on 15 comparative studies, most of which were retrospective. Overall, because of the small number of studies and their low methodological quality and considerable clinical heterogeneity, the strength of evidence was rated as insufficient for the superiority of the combination of the SFLC assay and traditional testing over traditional testing alone: specifically, the SLFC assay's impact on diagnostic accuracy in undiagnosed patients, as well as whether it is a better predictor of outcome or progression of MGUS to MM, a better indicator for therapeutic decisionmaking, or an appropriate substitute for other interventions.

Evidence Gaps

Table A summarizes the evidence gaps identified in our review, organized according to whether they relate to the population, intervention, comparator, outcome or study design (PICO-D). There is considerable clinical uncertainty regarding the applications of the SFLC assay both within and beyond the 2009 IMWG consensus guidelines.[3]

Table A. Evidence gaps affecting conclusions for the Key Questions

Key Questions	Evidence Gaps According to PICO Category

1. Does adding the SFLC assay and the kappa/lambda ratio to traditional testing (SPEP, UPEP, SIFE, or UIFE), compared with traditional testing alone, improve the diagnostic accuracy for PCDs (MGUS, MM, nonsecretory MM, or AL amyloidosis) in undiagnosed patients suspected of having a PCD?	**Population:**
	• Lack of direct comparisons of the SFLC assay with traditional testing in undiagnosed patients suspected of having a PCD.
	• Lack of studies of a nondiseased population as a comparison group (in order to accurately assess outcomes (e.g., false positives and true negatives).
	    o Studies of only patients with disease reflects the extreme end of the spectrum of disease severity and overestimates the proportion of patients with a positive result.
	
	Intervention/Comparator:
	• Lack of direct head-to-head comparisons of the SFLC assay alone with the SFLC assay plus traditional tests (e.g., to fully address whether SFLC measurement can replace UPEP or UIFE of 24-hour urine in diagnostic panels).
	
	Outcome:
	• Limited data regarding rates diagnostic accuracy of the SFLC assay.
	    o False positive rate in different settings.
	    o Effects of conditions such as polyclonal gammopathy or diminished kidney function on false positive rate
	    o Effects of factors such as antigen excess and technical variations in commercial assays on false negative rate.
	• Lack of consensus on suitable gold standard for PCD diagnosis.
	• Limited data on harms of testing.

2. As compared with traditional tests, how well does the SFLC assay independently predict progression to MM in patients with MGUS?	**Population:** • None. **Intervention/Comparator:** • Lack of direct head-to-head comparisons of the SFLC assay with traditional tests to predict progression from MGUS to MM. **Outcome:** • Lack of consensus on suitable metrics from the SFLC assay to predict progression of MGUS (e.g., absolute concentrations of the involved light chain (kappa or lambda), the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio).

Table A. Evidence gaps affecting conclusions for the Key Questions (continued)

Key Questions	Evidence Gaps According to PICO Category
3. In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay result in different treatment decisions as compared with traditional tests?	**Population:** • Lack of studies addressing the specific role of the SFLC assay in managing individual PCDs. **Intervention/Comparator:** • Lack of direct head-to-head comparisons of the SFLC assay and traditional tests (e.g., to fully address whether SFLC measurement can replace UPEP or UIFE of 24-hour urine in diagnostic panels to inform treatment decisions). **Outcome:** • Lack of consensus on suitable metrics from the SFLC assay to predict progression of diagnosed PCDs (e.g., absolute concentrations of the involved light chain (kappa or lambda), the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio). • Lack of studies of treatment decisions (including timing, duration, or type of treatment) based on results from the SFLC assay versus those from traditional tests.
4. In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), is the SFLC assay better than traditional tests in indicating how the patient responds to treatment and of outcomes (overall survival, disease-free survival, remission, light chain escape, and quality of life)?	**Population:** • Lack of studies addressing the specific role of the SFLC assay in assessing response and outcomes for patients with individual PCDs; nonsecretory MM, light chain MM, or AL amyloidosis. **Intervention or Comparator:** • Lack of direct head-to-head comparisons of the SFLC assay and traditional tests regarding patient response (e.g., to fully address whether SFLC measurement can replace UPEP or UIFE of 24-hour urine in diagnostic panels to measure response to treatment). **Outcome:** • Lack of consensus on suitable metrics from the SFLC assay to predict patient response and outcome (e.g., absolute concentrations of the involved light chain (kappa or lambda), the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio).
5. In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay reduce the need for other interventions (e.g., bone marrow biopsy)?	**Population:** • Lack of studies addressing the specific role of the SFLC assay in treatment decisions regarding individual PCDs: nonsecretory MM or AL amyloidosis. **Intervention/Comparator:** • Lack of direct head-to-head comparisons of the SFLC assay and traditional tests such as bone marrow biopsy or skeletal survey. **Outcome:** • Lack of consensus on suitable metrics from the SFLC assay to make treatment decisions for PCD patients (e.g., absolute concentrations of the involved light chain (kappa or lambda), the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio).

Abbreviations: AL amyloidosis = systemic amyloidosis in which amyloid [A] proteins derived from immunoglobulin light chains [L] are deposited in tissue; IFE = immunofixation electrophoresis; MGUS = monoclonal gammopathy of undetermined significance; MM = multiple myeloma; PCD = plasma cell dyscrasia; PICO-D = population, intervention, comparator, outcome, and study design; SFLC = serum free light chain; SIFE = serum immunofixation electrophoresis; SPEP=serum protein electrophoresis; UIFE = urine immunofixation electrophoresis; UPEP=urine protein electrophoresis.

Methods

Building on evidence gaps identified in the original CER (Table A above), the Evidence-based Practice Center (EPC) generated an initial list of FRN topics and solicited additional topics from a stakeholder panel (described below). Stakeholders were asked to consider four dimensions of need related to the proposed topic: importance, desirability of research, feasibility, and potential impact. These four dimensions come from the Effective Health Care Program Selection Criteria. The fifth dimension of the selection criteria, appropriateness, was evaluated by EPC program staff after submission of initial FRN topics. To inform the selection criterion of desirability of research, as well as avoidance of unnecessary duplication, we prepared results of an updated literature search in the ClinicalTrials.gov database and MEDLINE® to identify ongoing and newly published studies that were not published at the time of the last CER update. The studies identified in the searches were compared against the nominated FRN topics to assess if they would make future research on any nominated topic redundant, but none appeared to do so.

We adapted a Tufts-developed "7Ps" model of stakeholder engagement to identify individuals from seven stakeholder categories.[10] We designated a priori 10 stakeholders we would like to make up our proposed stakeholder panel, according to the 7Ps: two current patients/advocates to represent patients and the public; three providers (a clinical chemist/nominator, a hematologist–oncologist, and a nurse practitioner; hospital administrators are an important provider subgroup but were not represented because the SFLC assay is not a routine test in most hospitals); 0 purchasers (employers who purchase insurance policies were considered to share the payer perspective for this diagnostic test and therefore were not included as a separate stakeholder group); two payers (one private insurer and one Centers for Medicare and Medicaid Services employee); one policymaker; two principal investigators, researchers, or research funders; and one nonvoting product maker. All stakeholders completed a standard disclosure-of-interest form. The product maker was also asked to propose topics but did not participate in the discussion and prioritization of the topics. After the prespecified number of interested stakeholders was achieved in each category, AHRQ's Task Order Officer and the EPC jointly approved the panel.

The first round of Webinars was held in April and early May 2012. All but one stakeholder attended the first Webinar (held in replicate, on three different dates, to maximize participation), with one providing input by email in lieu of in-person participation. Taking the additional topics suggested by the stakeholder group during or immediately following the Webinars, we combined duplicate or similar FRN topics together and disseminated the revised list of topics, along with the list of possibly relevant ongoing studies (Appendix B), with an invitation to comment as to whether the nominated topics and supporting rationales were appropriately recorded and accounted for. These materials were sent to all stakeholders and discussed in the second set of Webinars, which were held (again in triplicate) in May 2012 to discuss the list and clarify or refine any of the topics. Eight of the 10 stakeholders attended the second Webinar. There was no third set of Webinars for this project, although we offered to hold one as well as to hold subgroup or one-on-one calls as needed. Following this second round, we further edited the topic list on the basis of stakeholder rationale. We sent this, along with minutes from all the Webinars, back to stakeholders, who were given 1 week to consider the input from stakeholders on other calls and to provide comments on research topics by email. Following this commentary period, we finalized the list and asked stakeholders to electronically indicate their top five topic choices using an online application (Zoomerang) to ensure a complete, structured response and voting

according to the AHRQ criteria for prioritization. The topics appeared in random order for each stakeholder to avoid selection bias.

The topics with the highest number of stakeholder endorsements were designated as the prioritized FRN topics. The EPC transformed the final list of FRN topics into answerable research questions using the standard "PICO-D" criteria (population, intervention, comparator, outcomes, study design). We discussed the pros and cons of various potential research designs and specifically considered the feasibility of the research questions focusing on potential sample size, time, and recruitment issues. Candidate study designs could differ across types of research needs.

Results

Research Needs

The FRN identification process led to the nomination of 16 topics (Table B). We considered the three topics with endorsement by at least 50 percent of the stakeholders (with 8 of the 10 participating stakeholders voting) as the highest priority FRN topics. All three are based directly on evidence gaps identified in the CER.

Table B. Nominated topics for future research needs for the SFLC assay in PCDs, according to rank*

Topic Rank†	Topic Questions	No. of Stakeholder Votes
	Prioritized future research needs topics	
1	Context of treatment decisions, monitoring, and outcomes: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in detecting disease status in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden)? [*Alternate text:* How accurately does the SFLC assay, compared with tests like electrophoresis or immunofixation, detect treatment response or relapse or disease activity?]	4
2	Context of treatment decisions, monitoring, and outcomes: How does the SFLC assay, either alone or in combination with traditional tests or as part of a definition of treatment response, compare with traditional tests in guiding treatment and management decisions (e.g., timing, duration, or type of treatment or use of adjunctive therapy or stem-cell transplantation) in patients with PCDs? a. Do these comparisons differ by individual PCD populations? b. How does the SFLC result, as a single result, change treatment or direct further investigation in individual PCDs? c. What are suitable metrics (i.e., the best measurements) from the SFLC assay to make treatment decisions for PCD patients (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio)? d. What is the optimal frequency of monitoring with the SFLC assay with and without other tests for guiding management decisions by disease status and by PCD?	5
3	Context of reducing need for other testing: What is the role of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in reducing the need for bone marrow examination by providing equivalent information for monitoring response to treatment, remission, or relapse in individual PCDs?	4

Table B. Nominated topics for future research needs for the SFLC assay in PCDs, according to rank* (continued)

Topic Rank†	Topic Questions	No. of Stakeholder Votes
	Other nominated future research needs topics	
4	Context of reducing need for other testing: What is the role of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in reducing the need for 24-hour urine testing in detection of light chains by providing equivalent information: a. for diagnosis of individual PCDs? b. for monitoring response to treatment, remission, or relapse in individual PCDs?	3
5[@]	Context of current practice How are clinicians adopting the SFLC assay currently? Is the test an add-on or does it replace an existing, more complicated or more invasive test? If so, are practitioners cutting back on use of other tests (bone marrow biopsy, 24-hour urine) as a result? What is the current frequency of SLFC testing in patients with PCDs?	3
6[@]	Context of standardization of best practice: What are the "best practices" (both technical and clinical) for use of the SFLC assay in different populations with regard to: a. laboratory testing (platform, lot, site) and reporting of results by laboratories b. clinical use in diagnosis, monitoring, and/or prognosis c. indication and prescribing/requisition patterns (e.g., frequency or timing of use)	3
7	Context of diagnostic testing: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests, in a high-risk population with regard to: a. conditions such as antigen excess and polyclonal gammopathy (which can result in false negatives and false positives, respectively) b. diminished kidney function and the need for different reference ranges in kidney failure c. technical variations in commercial assays, assay platform, or in house (hospital) vs. central lab testing d. resolving discordant SFLC and traditional testing results in different PCDs e. diagnostic reclassification based on SFLC testing	2
8	Context of diagnostic testing: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests, in individual PCDs with regard to: a. conditions such as antigen excess and polyclonal gammopathy (which can result in false negatives and false positives, respectively) b. diminished kidney function and the need for different reference ranges in kidney failure c. technical variations in commercial assays, assay platform, or in house (hospital) vs. central lab testing d. resolving discordant SFLC and traditional testing results in different PCDs e. diagnostic reclassification based on SFLC testing f. detecting response and remission, relapse, and assessment of tumor burden g. outcome prediction (response, disease-free survival, overall survival)	2

Table B. Nominated topics for future research needs for the SFLC assay in PCDs, according to rank* (continued)

Topic Rank†	Topic Questions	No. of Stakeholder Votes
9	Context of MGUS progression to MM How well does the SFLC assay predict progression to MM in patients with MGUS? a. What is its accuracy (strength of prediction as well as precision) as an independent predictor? b. Can it diagnose or detect disease progression earlier than traditional tests? c. Should the SFLC assay with or without other traditional testing be used in monitoring/surveillance and how often? d. Are there variant populations of MGUS where the SFLC is more relevant to predict progression to MM? e. What are suitable metrics (what are the best SFLC measurements) from the SFLC assay to predict progression of MGUS (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio)?	2
10	Context of treatment decisions, monitoring, and outcomes What is the role of the SFLC assay, relative to other tests, in providing prognostic information? a. How well does the SFLC assay predict outcomes such as treatment response, overall survival, or disease-free survival for individual PCD populations? b. Does the test result in improved health outcomes? c. What are suitable metrics from the SFLC assay to predict patient response and outcomes such as overall and disease-free survival (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio)?	2
11	Context of reducing need for other testing: Can the frequency of current SIFE, SPEP, and UPEP testing be decreased if SFLC testing is added?	2
12@	Context of cost effectiveness: What is the cost effectiveness of using the SFLC assay as an add-on to other testing?	2
13	Context of diagnostic testing: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests, in the general population with regard to: a. conditions such as antigen excess and polyclonal gammopathy (which can result in false negatives and false positives, respectively) b. diminished kidney function and the need for different reference ranges in kidney failure c. technical variations in commercial assays, assay platform, or in house (hospital) vs. central lab testing d. resolving discordant SFLC and traditional testing results in different PCDs e. diagnostic reclassification based on SFLC testing	1
14	Context of treatment decisions, monitoring, and outcomes What is the role of the SFLC assay, relative to other tests, in managing patients who have undergone stem-cell transplantation?	1
15@	Context of cost effectiveness: What are the factors driving both raw costs and other expenditures (e.g., total cost of usage at a given institution) in the SFLC assay? Is the cost variation an issue related to SFLC assay or in house use vs. outsourcing?	1

Table B. Nominated topics for future research needs for the SFLC assay in PCDs, according to rank* (continued)

Topic Rank†	Topic Questions	No. of Stakeholder Votes
16[@]	Context of newer tests: What are the newer tests available for the diagnosis, prognosis, and treatment monitoring of PCDs and how do they compare with the SFLC test with and without traditional tests?	1

**Polyclonal gammopathy* is seen in a diverse range of inflammatory and neoplastic diseases where clonal expansion occurs across various B-cell populations that produce more than one kind of immunoglobulin as part of a broad immune response unlike monoclonal gammopathy (or PCD) where a single neoplastic clone produces a single kind of immunoglobulin. Antigen excess is the condition in which SFLCs are present in great enough excess to form nonprecipitating immune complexes, which may alter the detection of the SFLCs and lead to inaccurate negative test results
† Prioritized topics (1–3) are ordered logically by clinical content. Other nominated topics are listed in the order they were prioritized by the stakeholder panel.@ Topics that were not addressed in the CER. Before initiating new research, a literature search and review is needed to confirm that any new research on the topic would not be duplicative
Abbreviations: MGUS=monoclonal gammopathy of undetermined significance; MM=multiple myeloma; PCD=plasma cell dyscrasia; SFLC=serum free light chain; SIFE=serum immunofixation electrophoresis; SPEP=serum protein electrophoresis; UIFE=urine immunofixation electrophoresis; UPEP=urine protein electrophoresis

The three top-ranked FRNs refer to the application of the SFLC assay in the contexts of diagnosis, treatment monitoring, and resource use. All three are of immediate and definitive clinical relevance and will impact the clinical management of patients with PCDs.

The three also have similar considerations for proposing a study design. The main reason for each topic being an evidence gap was a paucity of data, suggesting that a wide range of trial designs could be useful. A randomized controlled trial (RCT) would be ideal in all three cases, since RCTs inherently would yield the most robust data, but the low prevalence of PCDs and severity of the disease in many patients requiring monitoring both make it unethical to consider a trial in which some patients are randomized to undergo a potentially useful test and others are randomized to forego the test and the heterogeneity of the PCD testing population may make it less generalizable. It would not be feasible or ethical to compare SFLC testing with traditional testing in a controlled fashion, as current practice includes both sets of tests in a diagnostic panel. Moreover, the variable severity and stages of the disease in a referral population adds to the complexity of conducting a trial. Therefore, we considered more feasible study designs and propose two for each of the three topics, as described below. The two designs are not intended to be mutually exclusive, but rather complementary.

High-Priority Future Research Needs Topic 1

What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in detecting disease status in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden)?

Given the paucity of available evidence to completely answer important clinical questions relevant to the SFLC assay in the context of treatment decisions, monitoring, and outcomes, different study designs are reasonable. The following proposed study designs (Table C) address the diagnostic performance of SFLC testing and traditional testing in detecting disease status in various treatment contexts but differ in terms of their feasibility.

Table C. Potential research designs for topic 1

Considerations	Meta-Analysis of Individual Participant Data	Prospective Observational Study
Design description	Using completed or ongoing RCTs (e.g., STaMINA trial or others) or registry data as the source, where simultaneous SFLC and traditional testing has occurred, combine all treatment arms (within one or more RCTs or registry) into a single cohort of patients with a diagnosed PCD. Then ascertain the diagnostic accuracy of the SFLC assay against comparator(s), chosen a priori, to detect disease status in different treatment contexts (e.g., response and remission, relapse, or assessment of tumor burden).	Prospectively designed single-cohort studies consisting of patients receiving treatment (in the context of normal clinical practice) for a diagnosed PCD can be conducted to assess the diagnostic accuracy of the SFLC assay against comparator(s) to detect disease status in the different treatment contexts (e.g., response and remission, relapse, or assessment of tumor burden). Studies should have a priori calculation of the sample size needed for determination of the desired precision and should include inferences based on formal statistical testing of estimates of diagnostic accuracy.
Population	Patients with PCDs; each PCD could form a prespecified stratum or subgroup for analysis.	Patients with PCDs; each PCD could form a stratum or subgroup for analysis.
Intervention	SFLC testing and SFLC testing in combination with traditional testing.	SFLC testing and SFLC testing in combination with traditional testing.
Comparator	Traditional testing.	Traditional testing.
Outcomes	Measures of diagnostic accuracy (e.g., sensitivity, specificity, or predictive value) in different treatment contexts (e.g., response and remission, relapse, or assessment of tumor burden).	Measures of diagnostic accuracy (e.g., sensitivity, specificity, or predictive value) in different treatment contexts (e.g., response and remission, relapse, or assessment of tumor burden).
Timing	Trial duration/followup duration in cohorts or registries.	Trial duration/followup duration in cohorts.
Setting	Secondary analysis of preexisting data.	Prospective; clinical practice.
Advantages for producing a valid result	Benefits of large sample sizes and allowing for prespecified subgroup analyses but at the possible cost of bias from pooling varying patient groups; disadvantage of being a retrospective observational data analysis subject to inherent biases and limitations of component studies.	Relatively poor, given bias from lack of randomization.
Resource use, size, and duration	Resource use less than in an RCT but still moderate to high. Sample size not an issue, in spite of the relatively low prevalence of PCDs, because the cohorts would be from RCTs or registries. Duration would depend on acquisition of databases and subsequent analysis and would potentially be shorter than in a prospective cohort or clinical trial.	Resource use likely to be moderate. Size may be an issue due to low prevalence of PCDs. Study duration may vary due to natural history of condition.
Ethical, legal, and social issues	Meta-analysis uses existing data. With permission to use original study data, there should be no ethical issues.	Few ethical issues are likely to occur as the study involves components of normal patient care.
Availability of data/ability to recruit	Good; PCD is a rare disease but use of RCT or registry data would mitigate sample size limitations.	Poor, since PCD is a rare disease.

Abbreviations: PCD=plasma cell dyscrasia, RCT = randomized controlled trial, SFLC=serum free light chain

High-Priority Future Research Needs Topic 2

Context of treatment decisions, monitoring, and outcomes: How does the SFLC assay, either alone or in combination with traditional tests or as part of a definition of treatment response, compare with traditional tests in guiding treatment and management decisions (e.g., timing, duration, or type of treatment or use of adjunctive therapy or stem-cell transplantation) in patients with PCDs?

Because the main reason for this topic being an evidence gap is a lack of systematically and methodically generated data, we considered feasible study designs based on secondary data analysis, including creative use of survey designs, and propose two here (Table D). These two designs can be considered complementary.

Table D. Potential research designs for topic 2

Considerations	Meta-Analysis of Individual Participant Data	Physician Surveys
Design description	Using completed or ongoing RCTs (e.g., the STaMINA trial) or registry data as the source, where simultaneous SFLC and traditional testing has occurred, combine all treatment arms (within one or more RCTs or registry) into a single cohort of patients with a diagnosed PCD. In this composite study population ascertain whether the SFLC assay is equivalent to the comparator chosen a priori for the definition or prediction of treatment response and outcomes pre-specified in the parent study.	Using data from completed or ongoing RCTs (e.g., the STaMINA trial) or cohort studies among patients with a diagnosed PCD who have had simultaneous SFLC and traditional testing, design surveys for physicians to indicate their responses to a set of patient data with regard to response status, and therapeutic approach in the absence and presence of the SFLC data. Physicians could also be randomized as to whether they are given SFLC results before or after traditional results? Patient data may be categorized by individual PCDs and clinical stage of the disease depending on primary study.
Population	Patients with PCDs; each PCD could form a prespecified stratum or subgroup for analysis.	Physicians/providers.
Intervention	SFLC testing, metrics of SFLC testing, frequency of SFLC testing.	SFLC testing.
Comparator	Traditional testing, metrics of traditional testing, frequency of traditional testing (prespecified reference categories).	Traditional testing.
Outcomes	Response to treatment, remission, relapse, and light chain escape; overall survival and disease-free survival.	Diagnosis of response, remission or relapse; therapeutic approach proposed (timing, type and duration).
Timing	Trial duration/followup duration in cohorts or registries.	Trial duration/follow up duration in cohorts or registries.
Setting	Secondary analysis of preexisting data.	Patients with PCDs; each PCD could form a prespecified stratum or subgroup for analysis; clinical stage of the disease.
Advantages for producing a valid result	Benefits of large sample sizes and allowing for prespecified subgroup analyses but at the possible cost of bias from pooling varying patient groups; disadvantage of being a retrospective observational data analysis subject to inherent biases and limitations of component studies.	Feasible and inexpensive; appropriate methods of randomization of both participants and patient data would reduce bias. Disadvantages inherent to survey methodology including inflexibility and subjectivity. Dependent on the percentage of respondents, quality of participation and responses and the quality of the survey tool used.

Table D. Potential research designs for topic 2 (continued)

Considerations	Meta-Analysis of Individual Participant Data	Physician Surveys
Resource use, size, and duration	Resource use less than in an RCT but still moderate to high. Sample size not an issue, in spite of the relatively low prevalence of PCDs, because the cohorts would be from RCTs or registries. Duration would depend on acquisition of databases and subsequent analysis and would potentially be shorter than in a prospective cohort or clinical trial.	Resource use likely to be moderate; sample size would depend on number of respondents. Duration would depend on acquisition of databases and subsequent designing of survey and would potentially be shorter than a prospective cohort study or clinical trial.
Ethical, legal, and social issues	Would not be limiting.	Would not be limiting.
Availability of data/ability to recruit	Good; PCD is a rare disease but use of RCT or registry data will mitigate sample size limitations.	Dependent on number of respondents.

Abbreviations: PCD=plasma cell dyscrasia, RCT = randomized controlled trial, SFLC=serum free light chain

High-Priority Future Research Needs Topic 3

What is the role of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in reducing the need for bone marrow examination by providing equivalent information for monitoring response to treatment, remission, or relapse in individual PCDs?

We propose as two complementary study designs for this topic a prospective cohort study and a prospective observational study (Table E).

Table E. Potential research designs for topic 3

Considerations	Prospective Cohort Study	Prospective Observational Study
Design description	Using ongoing RCTs (e.g., the STaMINA trial or registry data as the source, combine all treatment arms (within one RCT or registry) into a single cohort of patients with a diagnosed PCD with requirement of a complete response to therapy. In this cohort (or across multiple cohorts from multiple, comparable RCTs or registries), ascertain whether the SFLC assay is equivalent to the gold standard (as defined in the RCT or chosen a priori if registry data are used) of treatment response.	Patients with a diagnosed PCD with requirement of a complete response to therapy (not including normal SFLC results) as well as a stringent complete response (including normal SFLC results), with physicians surveyed about what treatment decisions would be in the absence and presence of the SFLC data. Alternative: SFLC done whenever bone marrow done by someone blinded (possible since tests are ordered, not done in office); physicians could be randomized as to whether they are given SFLC results before or after traditional results.
Population	Patients with PCDs.	Physicians/providers.
Intervention	SFLC testing.	SFLC testing.
Comparator	Bone marrow examination.	Bone marrow examination.
Outcomes	Measures of diagnostic accuracy, e.g., sensitivity, specificity, predictive value in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden). Frequency of bone marrow examination.	Measures of diagnostic accuracy, e.g., sensitivity, specificity, predictive value in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden). Frequency of bone marrow examination.

Table E. Potential research designs for topic 3 (continued)

Considerations	Prospective Cohort Study	Prospective Observational Study
Timing	Trial duration/followup duration in cohorts or registries.	Trial duration/followup duration in cohorts or registries.
Setting	Prospective; clinical practice.	Prospective; clinical practice.
Advantages for producing a valid result	Arguably the optimal design when the treatment cannot ethically be assigned. Would carry most of the benefits of RCTs (or large registries) at the cost of possible bias from pooling of treatment arms.	Relatively poor, given bias due to lack of randomization and possible innate tendency of both physicians and patients to prefer more information to less.
Resource use, size, and duration	Resource use is less than in an RCT but still moderate to high. Size would not be an issue, in spite of the relatively low prevalence of PCDs, because the cohorts would be from RCTs or registries. Trial duration would be relatively short, since and the SLFC assay half-life is short (2 to 6 hours) and the median survival time is about 5 years .	Resource use is likely moderate. Size, and duration should not be an issue, since study would be observing normal clinical practice; sole difference is addition of SFLC assay for all patients rather than performed only at physician's discretion, but assay is affordable (~$25 to $100, anecdotally) and covered by insurance.
Ethical, legal, and social issues	Ethical issues should not be an issue.	Ethical issues should not be an issue.
Availability of data/ability to recruit	Moderate; PCD is a rare disease but use of RCT or registry data should mitigate this.	Poor, since PCD is a rare disease.

Abbreviations: PCD=plasma cell dyscrasia, RCT = randomized controlled trial, SFLC=serum free light chain

Discussion

The three top-ranked FRNs referred to the application of the SFLC assay in the contexts of diagnosis, treatment monitoring, and resource use. These contexts echoed several of the Key Questions in the original CER, albeit with several differences.

The study designs we proposed were guided by two major considerations. First, we had to take into account the reality that the SFLC assay was already a test in regular and routine use for the management of PCDs and that many providers believed that it added value to patient care and helped in decisionmaking. Hence it was difficult to visualize a situation where it would be possible to randomize patients to care in the absence of performing the assay. Second, it was also crucial to avoid the very limitations that the CER identified in the studies evaluated for our original Key Questions. In addition, since both the CER and discussion with stakeholders revealed that SFLC results are complementary, not prescriptive on their own, we focused the study designs to capture clinical preferences, although we acknowledge that patient preference would be another avenue of study.

The process was not without limitations. We were able to obtain input from only 8 of the 10 members of our stakeholder panel, and the total number of stakeholders recruited was restricted, thus limiting representation. The threshold for distinguishing high-priority topics from the other topics in the final ranking was somewhat arbitrary, namely four or more votes (i.e., prioritization by 50 percent or more of the eight voters). Also, despite formal planning, the selection of stakeholders, solicitation of contributions, facilitation of discussion, and synthesis of suggestions remain, to some degree, idiosyncratic. There are as of yet no accepted standard methods by which to assess the validity of procedures to synthesize diverse stakeholder viewpoints.

Conclusions

This report identifies three high-priority FRN topics to study the SFLC assay in PCD patients, which were identified by a stakeholder panel. These are all parts of or repeats of some of the CER Key Questions, for which there was little evidence. They are as follows:

1. Context of treatment decisions, monitoring, and outcomes: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in detecting disease status in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden)? [Alternate text: How accurately does the SFLC assay, compared with tests like electrophoresis or immunofixation, detect treatment response or relapse or disease activity?]

2. Context of treatment decisions, monitoring, and outcomes: How does the SFLC assay, either alone or in combination with traditional tests or as part of a definition of treatment response, compare with traditional tests in guiding treatment and management decisions (e.g., timing, duration, or type of treatment or use of adjunctive therapy or stem-cell transplantation) in patients with PCDs?
    a. Do these comparisons differ by individual PCD populations?
    b. How does the SFLC result, as a single result, change treatment or direct further investigation in individual PCDs?
    c. What are suitable metrics (i.e., the best measurements) from the SFLC assay to make treatment decisions for PCD patients (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio)?
    d. What is the optimal frequency of monitoring with the SFLC assay with and without other tests for guiding management decisions by disease status and by PCD?

3. Context of reducing need for other testing: What is the role of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in reducing the need for bone marrow examination by providing equivalent information for monitoring response to treatment, remission, or relapse in individual PCDs?

Abbreviations

AHRQ	Agency for Healthcare Research and Quality
AL amyloidosis	Systemic, or primary, amyloidosis in which amyloid [A] proteins derived from immunoglobulin light chains [L] are deposited in tissue (also called light-chain amyloidosis)
CER	Comparative Effectiveness Review
EPC	Evidence-based Practice Center
FLC	Free light chain
FRN	Future Research Needs
IMWG	International Myeloma Working Group
MGUS	Monoclonal gammopathy of undetermined significance
MM	Multiple myeloma
M-protein	Monoclonal protein (also called paraprotein)

PCD	Plasma cell dyscrasia
PICO (also PICOTS)	Populations, interventions (diagnostic tests/disease monitoring), comparators, outcomes (and timing and settings)
RCT	Randomized controlled trial
SFLC	Serum free light chain
SIFE	Serum immunofixation electrophoresis
SPEP	Serum protein electrophoresis
UIFE	Urine immunofixation electrophoresis
UPEP	Urine protein electrophoresis

References

1. Kyle RA, Rajkumar SV. Epidemiology of the plasma-cell disorders. Best Pract Res Clin Haematol. 2007;20:637-64.

2. Bradwell AR, Carr-Smith HD, Mead GP, et al. Highly sensitive, automated immunoassay for immunoglobulin free light chains in serum and urine. Clin Chem. 2001 Apr;47(4):673-80.

3. Dispenzieri A, Kyle R, Merlini G, et al. International Myeloma Working Group guidelines for serum-free light chain analysis in multiple myeloma and related disorders. Leukemia. 2009 Feb;23(2):215-24.

4. Mead GP, Carr-Smith HD, Drayson MT, et al. Serum free light chains for monitoring multiple myeloma. Br J Haematol. 2004 Aug;126(3):348-54.

5. Sanchorawala V, Seldin DC, Magnani B, et al. Serum free light-chain responses after high-dose intravenous melphalan and autologous stem cell transplantation for AL (primary) amyloidosis. Bone Marrow Transplant. 2005 Oct;36(7):597-600.

6. Katzmann JA. Serum free light chains: quantitation and clinical utility in assessing monoclonal gammopathies. Clin Lab News. 2006 Jun;June:12-4.

7. Kyle RA, Durie BGM, Rajkumar SV, et al. Monoclonal gammopathy of undetermined significance (MGUS) and smoldering (asymptomatic) multiple myeloma: IMWG consensus perspectives risk factors for progression and guidelines for monitoring and management. Leukemia. 2010 Jun;24(6):1121-7.

8. Rajkumar SV. Clinical features, laboratory manifestations, and diagnosis of multiple myeloma. In: Basow DS, eds. UpToDate. Waltham: UpToDate; 2011.

9. Whitlock EP, Lopez SA, Chang S, et al. AHRQ Series Paper 3: Identifying, selecting, and refining topics for comparative effectiveness systematic reviews: AHRQ and the Effective Health-Care program. J Clin Epidemiol. 2010;63:491-501.

10. Concannon T, Meissner P, Grunbaum J, et al. A new taxonomy for stakeholder engagement in patient-centered outcomes research. J Gen Intern Med. 2012:1-7.

Background

Plasma cell dyscrasias (PCDs) are a group of clonal disorders characterized by the uninhibited expansion of a monoclonal population of malignant plasma cells.[1] Multiple myeloma (MM) is the most common malignant plasma cell tumor, accounting for about 1 percent of all cancer types,[1] and the second most common hematologic malignancy in the United States. With an age-adjusted incidence rate of 5.5 cases per 100,000 population,[2] the American Cancer Society estimated that there were 19,900 new diagnoses and 10,790 deaths due to myeloma in 2007.[3] Although the median survival has improved to 5 years with current standards of treatment,[4] the annual costs of modern therapies can range from $50,000 to $125,000 per patient.[5,6]

Plasma cells arise from B cells in the bone marrow and produce immunoglobulins that constitute the body's normal humoral immune response. The immunoglobulin molecule is composed of a heavy chain and a light chain. Plasma cells normally produce light chains in excess that do not bind to heavy chains to form a complete immunoglobulin molecule and instead enter the bloodstream as free light chains (FLCs).

In PCDs, each abnormally expanded clone of malignant plasma cells produce an excess of either intact immunoglobulin or FLCs of a single type; either type of excess molecule is called a monoclonal protein (M-protein) or paraprotein. Measurement of M-proteins (either complete immunoglobulins or FLCs) is integral to diagnosing PCDs, monitoring disease response to therapy and adjusting treatment, and determining disease progression or relapse.

PCDs range in severity. The mildest and most common PCD is the precancerous monoclonal gammopathy of undetermined significance (MGUS), affecting approximately 3 percent of the general population 50 years of age or older. [1] MGUS can progress to asymptomatic MM (also called smoldering or indolent MM) or symptomatic MM. The M-proteins produced in MM are either intact immunoglobulins or FLCs or both. Rarer MM variants include light-chain MM (light chain MM, formerly known as Bence Jones myeloma), characterized by expanded FLC-producing clones, and oligosecretory or nonsecretory MM, in which few detectable light- or heavy-chain M-proteins are secreted. Other PCDs include systemic (primary) AL amyloidosis, also called light-chain amyloidosis (in which amyloid [A] proteins derived from immunoglobulin light chains [L] are deposited in tissue), as well as macroglobulinemia, solitary plasmacytoma, and plasma cell leukemia. AL amyloidosis can be a complication of MM but is often considered a distinct disorder related to a relatively stable, slow-growing plasma cell clone and organ dysfunction.

The International Myeloma Working Group (IMWG) recommends the following actions and tests for evaluation of a patient suspected of having a myeloma:[7,8] a complete history taking and physical examination; routine laboratory testing including serum protein electrophoresis (SPEP), serum immunofixation electrophoresis (SIFE), nephelometric quantitation of immunoglobulins, and measurement of serum FLCs (SFLCs); bone marrow aspiration and biopsy with immunophenotyping, conventional cytogenetics, and fluorescence in situ hybridization; and imaging. Thus, testing for M-protein is only one part—albeit an integral part—of a suite of tests done to diagnose PCDs.

M-protein measurement and typing are traditionally achieved through the use of SPEP and/or urine protein electrophoresis (UPEP) and SIFE and/or urine immunofixation electrophoresis (UIFE), plus immunoglobulin quantification. These traditional tests have relatively low sensitivity, especially regarding concentrations of SFLCs. This lack of sensitivity leads to

undetected cases of PCDs that involve excess FLCs. It is likely that up to 3 percent of cases of nonsecretory MM, light chain MM, or AL amyloidosis are not detected by traditional tests.[9] To increase the chance of detection of FLCs in urine, 24-hour urine collection has been recommended, along with procedures to concentrate urine samples. Yet these adaptations can be cumbersome for patients and providers, affecting compliance and test accuracy.

The SFLC assay (Freelite™ Assay, The Binding Site Ltd., Birmingham, United Kingdom) was introduced in 2001 to measure the FLC component in particular.[10] The SFLC assay works by recognizing an epitope that is detectable only on light chains that are not bound to the heavy chain of the immunoglobulin molecule (i.e., FLCs) in the serum. This is the sole SFLC assay approved by the U.S. Food and Drug Administration and is classified as an immunoglobulin light chain–specific immunological test system. It measures kappa and lambda light chains separately and detects low concentrations of FLCs—less than 1 mg/dL in serum and less than 200 mg/day in urine.[11] The other main advantage is the ability to measure the ratio of kappa chains to lambda chains, for which the diagnostic range is 0.26–1.65.[12] An abnormal ratio provides a useful index of clonality, as clonal disorders produce disproportionately high concentrations of a single type of light chain. In a given case of PCD, if kappa chains are in excess, the kappa/lambda ratio is greater than 1.65; if the lambda chains are in excess, the ratio is less than 0.26.

It has been suggested that the SFLC assay could play an adjunctive role in screening, diagnosis, monitoring, and prognosis of PCDs in high-risk populations. The IMWG currently considers the SFLC assay to be an adjunct to traditional tests.[13] The assay could allow for quantitative monitoring of response and remission after treatment and provide prognostic information,[14,15] potentially reducing the need for frequent bone marrow biopsy for purposes of quantifying plasma cells, which is required as part of stringent monitoring for MGUS progression to MM or defining disease remission.[13] It could also potentially be used in conjunction with SPEP and SIFE to replace urine tests that require 24-hour collection (UPEP and UIFE), which could simplify diagnosis and disease monitoring,[12,13] although proper studies have yet to be undertaken to explore this potential benefit. The SFLC assay may also be the only means of detecting a disease marker in some disease settings: nonsecretory MM, where SFLCs are often the only marker of the disease[16]; AL amyloidosis, where low M-protein concentrations may not be detected by means of conventional techniques; and light chain MM, where the M-protein consists only of FLCs. Thus, in addition to detecting a wider spectrum of PCDs than traditional tests, the assay may help detect earlier stages of the disease, and because of the short half-life of SFLCs (2 to 6 hours, vs. 21 days for complete immunoglobulins[17]), the assay may also help detect relapses and treatment failures earlier than by reliance on M-protein concentrations alone. These diagnostic applications have yet to be validated and standardized.[9] Thus, although the SFLC assay has been in use for a decade, it remains unclear how best to incorporate it into clinical practice.[18]

The current Future Research Needs (FRN) project was launched upon the completion of an Agency for Healthcare Research and Quality (AHRQ) comparative effectiveness review (CER) and builds on the evidence gaps identified in that review. The present report describes the development of a stakeholder-prioritized list of research needs for the use of the SFLC assay in PCDs, along with a measured consideration of the advantages and disadvantages of various potential research designs, in order to help researchers and funders develop future research proposals or solicitations.

Scope of the CER

The 2012 CER upon which the current FRN report is based, Serum Free Light Chain Analysis for the Diagnosis, Management, and Prognosis of Plasma Cell Dyscrasias (available at www.effectivehealthcare.ahrq.gov/ehc/products/264/900/Serum-Free-Light-Chain_Draft-Report_20111215.pdf) was sponsored by AHRQ and conducted by the Tufts Evidence-based Practice Center (EPC). It reviewed pertinent publications through April 2012 regarding the role of the SFLC assay in combination with traditional testing by electrophoresis and immunofixation, compared with traditional testing alone, in the diagnosis and management of patients with PCDs. This CER examined evidence addressing five Key Questions.

- Key Question 1: Does adding the SFLC assay and the kappa/lambda ratio to traditional testing (SPEP, UPEP, SIFE, or UIFE), compared with traditional testing alone, improve the diagnostic accuracy for PCDs (MGUS, MM, nonsecretory MM, or AL amyloidosis) in undiagnosed patients suspected of having a PCD?
- Key Question 2: As compared with traditional tests, how well does the SFLC assay independently predict progression to MM in patients with MGUS?
- Key Question 3: In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay result in different treatment decisions as compared with traditional tests?
  - Does the use of the SFLC assay affect the management of patients by allowing for earlier institution of specific therapies?
  - Does the use of the SFLC assay influence the duration of treatment?
  - Does the use of the SFLC assay influence the type of treatment (e.g., radiation therapy)?
- Key Question 4: In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), is the SFLC assay better than traditional tests in indicating how the patient responds to treatment and of outcomes (overall survival, disease-free survival, remission, light chain escape, and quality of life)?
- Key Question 5: In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay reduce the need for other interventions (e.g., bone marrow biopsy)?

Findings of the CER

Results and conclusions from the CER are based on 15 comparative studies. Three retrospective studies evaluated the SFLC assay as an adjunct to traditional testing in populations suspected of having a PCD. Three other retrospective studies of AL amyloidosis and eight studies of MM (six of which were retrospective) evaluated either baseline or post-treatment concentrations of SFLC or monoclonal protein as predictors of clinical outcomes. Finally, one study explored whether the use of the SFLC assay compared with traditional testing would reduce the need for bone marrow examination in assessing response to treatment.

Overall, because of the small number of studies and their low methodological quality and considerable clinical heterogeneity, the strength of evidence was rated as insufficient for the superiority of the combination of the SFLC assay and traditional testing over traditional testing alone: specifically, the SLFC assay's impact on diagnostic accuracy in undiagnosed patients, as well as whether it is a better predictor of outcome or progression of MGUS to MM, a better indicator for therapeutic decisionmaking, or an appropriate substitute for other interventions.

Currently, there is a lack of 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. Its precise role and optimal use across the spectrum of PCDs and clinical settings still needs to be defined. Potential areas where its benefit may be seen are in diagnosis and prognosis, monitoring of therapy, and aiding treatment decisions. Future research should focus on standardization of patient inclusion criteria, testing of diagnostic and disease monitoring algorithms, and defining outcome and response definitions. We discuss this in more detail in the next section.

Evidence Gaps

Table 1 summarizes the evidence gaps identified in our review, organized according to whether they relate to the population, intervention, comparator, outcome or study design (PICO-D). There is considerable clinical uncertainty regarding the applications of the SFLC assay both within and beyond the 2009 IMWG consensus guidelines.[13] Areas of uncertainty span the comparative effectiveness of the adjunctive role of the assay for the diagnosis of PCDs and the adjunctive and independent role of the assay in therapeutic decisions and monitoring, recognition of response and remission, and in predicting clinical outcomes and prognosis among patients with diagnosed PCDs. While comparative effectiveness was a common theme across all the CER's Key Questions, clinical effectiveness is also a measure clinicians use to gauge the value of a test and therefore we applied this consideration for the purposes of identifying FRN topics. The currently available data do not completely answer important clinical questions relevant to patient management; further research is needed to help elucidate these issues.

Table 1. Evidence gaps affecting conclusions for the Key Questions

Key Questions	Evidence Gaps According to PICO Category

1. Does adding the SFLC assay and the kappa/lambda ratio to traditional testing (SPEP, UPEP, SIFE, or UIFE), compared with traditional testing alone, improve the diagnostic accuracy for PCDs (MGUS, MM, nonsecretory MM, or AL amyloidosis) in undiagnosed patients suspected of having a PCD?	**Population:**
	• Lack of direct comparisons of the SFLC assay with traditional testing in undiagnosed patients suspected of having a PCD.
	• Lack of studies of a nondiseased population as a comparison group (in order to accurately assess outcomes (e.g., false positives and true negatives).
	    o Studies of only patients with disease reflects the extreme end of the spectrum of disease severity and overestimates the proportion of patients with a positive result.
	
	Intervention/Comparator:
	• Lack of direct head-to-head comparisons of the SFLC assay alone with the SFLC assay plus traditional tests (e.g., to fully address whether SFLC measurement can replace UPEP or UIFE of 24-hour urine in diagnostic panels).
	
	Outcome:
	• Limited data regarding rates diagnostic accuracy of the SFLC assay.
	    o False positive rate in different settings.
	    o Effects of conditions such as polyclonal gammopathy or diminished kidney function on false positive rate
	    o Effects of factors such as antigen excess and technical variations in commercial assays on false negative rate.
	• Lack of consensus on suitable gold standard for PCD diagnosis.
	• Limited data on harms of testing.

Table 1. Evidence gaps affecting conclusions for the Key Questions (continued)

Key Questions	Evidence Gaps According to PICO Category
2. As compared with traditional tests, how well does the SFLC assay independently predict progression to MM in patients with MGUS?	**Population:** • None. **Intervention/Comparator:** • Lack of direct head-to-head comparisons of the SFLC assay with traditional tests to predict progression from MGUS to MM. **Outcome:** • Lack of consensus on suitable metrics from the SFLC assay to predict progression of MGUS (e.g., absolute concentrations of the involved light chain (kappa or lambda), the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio).
3. In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay result in different treatment decisions as compared with traditional tests?	**Population:** • Lack of studies addressing the specific role of the SFLC assay in managing individual PCDs **Intervention/Comparator:** • Lack of direct head-to-head comparisons of the SFLC assay and traditional tests (e.g., to fully address whether SFLC measurement can replace UPEP or UIFE of 24-hour urine in diagnostic panels to inform treatment decisions). **Outcome:** • Lack of consensus on suitable metrics from the SFLC assay to predict progression of diagnosed PCDs (e.g., absolute concentrations of the involved light chain (kappa or lambda), the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio). • Lack of studies of treatment decisions (including timing, duration, or type of treatment) based on results from the SFLC assay versus those from traditional tests.
4. In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), is the SFLC assay better than traditional tests in indicating how the patient responds to treatment and of outcomes (overall survival, disease-free survival, remission, light chain escape, and quality of life)?	**Population:** • Lack of studies addressing the specific role of the SFLC assay in assessing response and outcomes for patients with individual PCDs; nonsecretory MM, light chain MM, or AL amyloidosis. **Intervention or Comparator:** • Lack of direct head-to-head comparisons of the SFLC assay and traditional tests regarding patient response (e.g., to fully address whether SFLC measurement can replace UPEP or UIFE of 24-hour urine in diagnostic panels to measure response to treatment). **Outcome:** • Lack of consensus on suitable metrics from the SFLC assay to predict patient response and outcome (e.g., absolute concentrations of the involved light chain (kappa or lambda), the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio).

Table 1. Evidence gaps affecting conclusions for the Key Questions (continued)

Key Questions	Evidence Gaps According to PICO Category
5. In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay reduce the need for other interventions (e.g., bone marrow biopsy)?	**Population:** • Lack of studies addressing the specific role of the SFLC assay in treatment decisions regarding individual PCDs: nonsecretory MM or AL amyloidosis. **Intervention/Comparator:** • Lack of direct head-to-head comparisons of the SFLC assay and traditional tests such as bone marrow biopsy or skeletal survey. **Outcome:** • Lack of consensus on suitable metrics from the SFLC assay to make treatment decisions for PCD patients (e.g., absolute concentrations of the involved light chain (kappa or lambda), the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio).

Abbreviations: AL amyloidosis = systemic amyloidosis in which amyloid [A] proteins derived from immunoglobulin light chains [L] are deposited in tissue; IFE = immunofixation electrophoresis; MGUS = monoclonal gammopathy of undetermined significance; MM = multiple myeloma; PCD = plasma cell dyscrasia; PICO-D = population, intervention, comparator, outcome, and study design; SFLC = serum free light chain; SIFE = serum immunofixation electrophoresis; SPEP=serum protein electrophoresis; UIFE = urine immunofixation electrophoresis; UPEP=urine protein electrophoresis.

Methods

Identification of Evidence Gaps

We used an iterative process with a stakeholder panel to identify FRN topics for prioritization. Building on evidence gaps identified in the original CER, the EPC generated an initial list of FRN topics (Table 1 above) and solicited additional topics from the stakeholder panel through conference calls and emails.

Criteria for Prioritization

Stakeholders (described in the next section) were asked to consider four dimensions of need related to the proposed topic. These four dimensions come from the Effective Health Care Program selection criteria.

- Importance
  - Represents a significant disease burden, large proportion or priority population.
  - Is of high public interest; affects health care decisionmaking, outcomes, or costs for a large proportion of the U.S. population or for a priority population in particular.
  - Was nominated/strongly supported by one or more stakeholder groups.
  - Represents important uncertainty for decisionmakers.
  - Incorporates issues around both clinical benefits and potential clinical harms.
  - Represents important variation in clinical care, or controversy in what constitutes appropriate clinical care.
  - Represent high costs to consumers, patients, health care systems or payers; due to common use, high unit costs, or high associated costs.
- Desirability of research
  - Would not be redundant (i.e., the proposed topic is not already covered by available or soon-to-be available high quality systematic review by AHRQ or others).
- Feasibility
  - Effectively uses existing research and knowledge by considering adequacy of research for conducting a systematic review, and newly available evidence.
- Potential impact
  - Potential for significant health impact, significant economic impact, potential change, potential risk from inaction, addressing inequities and vulnerable populations, and/or addressing a topic with clear implications for resolving important dilemmas in health and health care decisions made by one or more stakeholder groups.

The fifth dimension of the selection criteria, appropriateness, was evaluated by EPC program staff after submission of initial FRN topics.

- Appropriateness:
  - Represents a health care drug, intervention, device, technology or health care system/setting available (or soon to be available) in the United States.
  - Relevant to 1013 enrollees (Medicare, Medicaid, CHIP, other Federal health care programs.

- Represents one of the priority conditions designated by the U.S. Department of Health and Human Services (HHS).

Use of the Effective Health Care Program prioritization criteria was emphasized repeatedly throughout the prioritization process, including during discussion, nomination, and final topic selection. Upon the close of stakeholder prioritization, we identified the top five topics as those most frequently endorsed by stakeholders in their top five selections. The final ranked list was emailed to stakeholders.

To inform the selection criterion of desirability of research, as well as avoidance of unnecessary duplication, we prepared results of updated literature search in the ClinicalTrials.gov database and MEDLINE to identify ongoing and newly published studies that were not published at the time of the last CER update. Protocols of retrieved entries were reviewed for use of interventions and outcomes relevant to the Key Questions of the CER. The search strategy is provided in Appendix A and the identified studies are listed in Appendix B. The studies identified in the searches were compared against the nominated FRN topics to assess if they would make future research on any nominated topic redundant, but none appeared to do so.

Engagement of Stakeholders, Researchers, and Funders

Although researchers and funders of research are the primary audience for FRN documents, the EPC solicits input from other stakeholders as well when identifying high-priority research gaps and FRN topics. Stakeholders are selected to provide broad expertise and a breadth of perspectives, as well as input on the kind of information that is helpful in health care decisionmaking. These stakeholders are engaged throughout the future research process. Their role is to (1) review the preliminary list of evidence gaps and possible future research topics derived from them, (2) to nominate additional topics to the list, (3) discuss topic nominations, and (4) participate in prioritization of the FRN topics. Stakeholders are not involved in translating the gaps into research questions and study designs or composing or reviewing the report.

Prespecification of the Stakeholder Panel

We adapted a Tufts-developed "7Ps" model of stakeholder engagement to identify individuals from seven stakeholder categories.[19] This model is designed to build a panel representing the full range of stakeholders who may use research evidence in health care and public health decisionmaking. The 7Ps are:

1. **Patients and the public:** This group represents current and potential consumers of patient-centered health care and population-focused public health research. This group also includes caregivers, family members, and patient advocacy organizations, all of whom represent the interests of consumers.
2. **Providers:** This group includes individuals (e.g., nurses, physicians, and other providers of care and support services) and organizations (e.g., hospitals, clinics, community health centers, community-based organizations, pharmacies, emergency medical services agencies, skilled nursing facilities, schools) that provide care to patients and populations.
3. **Purchasers:** This group includes employers, the self-insured, government, and other entities responsible for underwriting the costs of health care.

4. **Payers:** This group represents insurers, Medicare and Medicaid, individuals with deductibles, and others responsible for reimbursement for interventions and episodes of care.
5. **Policymakers:** This group includes organizations such as the White House, Health and Human Services, Congress, States, professional associations, and intermediary groups that collate and distribute information to policymakers.
6. **Principal investigators, researchers, and research funders.**
7. **Product makers** (i.e., drug and device manufacturers): representatives were asked to participate in the topic nomination process but not in topic refinement or final prioritization of topics.

These categories are not mutually exclusive. Any stakeholder may wear several hats and may be responsible for different types of decisions. For example, some health care purchasers are also payers; and conversely, some payers also provide care. Patients and their advocates may be providers or employers with policy-making responsibilities, and so on.

In addition, each of these seven stakeholder types may be focused on applying comparative effectiveness research at the patient level or at the population level. Patient-level decisions include questions pertaining to what treatment would be best for a given patient at a particular time. Population-level decisions include questions pertaining to what services, resources, policies, or other alternatives are best for groups of patients and entire communities that are connected by a practice setting, geography, clinical domain, or other cluster. To be patient-centered, decisions made about groups of patients must recognize both the diversity of needs across populations and the heterogeneity of individuals within populations. Each stakeholder was asked to make patient-centered recommendations, whether they represent a category that is focused on treating patients or on treating populations.

We designated a priori the numbers of stakeholders we would like to represent each group in our proposed stakeholder panel (Table 2), to achieve balance across stakeholder categories and to cover a range of technical and personal expertise in the group as a whole. We invited individuals who have previously served in advisory roles for the EPC's SFLC CER. EPC team members also identified contacts in their professional networks. We conducted searches in Google, MEDLINE, Web of Science, and other publication databases as needed to identify active SFLC researchers to fill the provider and principal investigator categories.

Table 2. Prespecified composition of stakeholder panel for the SFLC future research needs document*

Category	Subcategory	No. of Stakeholders
Patients and the public	Current patients/patient advocates	2
Providers	Clinician – Clinical Chemist/Nominator	3
	Clinician – Hematologist–Oncologist	
	Clinician – Nurse practitioner	
Payers	Private insurer	2
	Centers for Medicare and Medicaid Services [Federal employee]	
Policymakers	Health Resources and Services Administration researcher	1
Principal investigators/ researchers	Clinical researchers	2
Product maker [nonvoting]	Binding Site Ltd.	1^
TOTAL		11

Abbreviation: SFLC = serum free light chain.
* In the provider category, hospital administrators are an important provider subgroup but were not represented in this stakeholder panel because the SFLC assay is not a routine test in most hospitals. In the purchaser category, employers who purchase insurance policies were considered to share the payer perspective for this diagnostic test and therefore were not included as a separate stakeholder group.^ Product maker was contacted to submit FRN topics but did not provide a response.

Invitation and Solicitation of Topics

An invitation letter was sent to our candidates for representing each of the six stakeholder categories (excluding purchasers, for this project). To comply with the Paperwork Reduction Act requirements, the maximum number of nonfederal employee stakeholders is nine. All stakeholders completed a standard disclosure-of-interest form and were told that these forms would be made available to all other stakeholders. Along with the invitation letter, we distributed pertinent portions of the executive summary of the CER and a description of the FRN project to the invited stakeholders. The purpose of the project and expectations for input from the stakeholders were outlined clearly in the invitation letter. After the prespecified number of interested stakeholders was achieved in each category (Table 2), AHRQ's Task Order Officer and the EPC jointly approved the panel.

The first round of Webinars, held in April and May 2012, served to introduce members, describe the purpose and process of FRN topic development, and review Key Questions and FRN topics proposed in the CER. During and after the first Webinar, stakeholders identified additional topics for nomination as well as comments. We synthesized these into the existing list of topics (including combining duplicate or similar FRN topics together into one topic). All but one stakeholder attended the first Webinar (held in replicate, on three different dates, to maximize participation), with one providing input by email in lieu of in-person participation.

Topics nominated by stakeholders were incorporated into the topic list along with supporting rationale, which we condensed from the discussion and subsequent emails. We combined duplicate or similar FRN topics together and disseminated the revised list of topics, along with the list of possibly relevant ongoing studies (Appendix B), with an invitation to comment as to whether the nominated topics and supporting rationales were appropriately recorded and accounted for.

These materials were sent to all stakeholders and discussed in the second set of Webinars, which were held (again in triplicate) in May 2012 to discuss the list and clarify or refine any of the topics. The topics could have included new questions identified during the conduct of the

CER that could not be addressed in that context. Eight of the 10 stakeholders attended the second Webinar. There was no third set of Webinars for this project, although we offered to hold one as well as to hold subgroup or one-on-one calls as needed to accommodate stakeholder schedules and to allow all stakeholders to provide input.

Following this second round, we further edited the topic list on the basis of stakeholder rationale. We sent this, along with minutes from all the Webinars, back to stakeholders, who were given 1 week to provide comments by email. Following this commentary period, we finalized the list and asked stakeholders to electronically indicate their top five topic choices using an online application (Zoomerang) to ensure a complete, structured response and voting according to the AHRQ criteria for prioritization. The topics appeared in random order for each stakeholder to avoid selection bias. The topics with the highest number of stakeholder endorsements were designated as the prioritized FRN topics. Stakeholders from industry were not included in the prioritization process.

The EPC attempted to use natural breaks in the rankings of the topics across categories, rather than strictly defining the numbers of topics that fell into each category. The final ranked list was emailed to stakeholders.

Research Question Development and Research Design Considerations

The EPC transformed the final list of FRN topics into answerable research questions using the standard "PICO-D" criteria (population, intervention, comparator, outcomes, study design). We discussed the pros and cons of various potential research designs and specifically considered the feasibility of the research questions focusing on potential sample size, time, and recruitment issues. To do so, we followed the structure laid out in the Future Research Needs for the Comparison of Percutaneous Coronary Interventions With Bypass Graft Surgery in Nonacute Coronary Artery Disease as a sample FRN document.[20] As needed, the EPC consulted with individual stakeholders for assistance in making decisions regarding appropriate study designs.

Candidate study designs could differ across types of research needs. Effectiveness or efficacy of treatments are most definitively addressed in randomized trials and secondarily addressed in well-conducted nonrandomized comparative observational studies. In contrast, eliciting patient preferences can be meaningfully performed with nonexperimental designs (e.g., in a survey). Furthermore, observational studies may be most appropriate to enhance generalizability and determine effectiveness as opposed to efficacy alone. Each final FRN topic was assessed as to context of the research question and as to whether evaluation of efficacy or effectiveness is of greater need. Regardless of study design, PICO criteria were proposed.

Studies that do not require new data collection are in principle feasible, provided that access to existing data can be obtained or has already been granted. An analysis of an existing registry, a standard meta-analysis, or a meta-analysis of individual patient data can be conducted in a limited time frame. The feasibility of such studies generally does not depend on the desired sample size. The feasibility of trials (or other studies requiring accrual of primary data) may be infeasible if they are too expensive or complex to conduct, would require too long a followup, or would rely on information or data that is not yet available or would be difficult to obtain. It is important to note that randomized trials are among the most expensive research designs.

Results

Research Needs

The FRN identification process led to the nomination of 16 topics (Table 3). We considered the 3 topics with endorsement by at least 50 percent of the stakeholders (with 8 of the 10 stakeholders voting) as the highest priority FRN topics. All three are based directly on evidence gaps identified in the CER.

Table 3. Nominated topics for future research needs for the SFLC assay in PCDs, according to rank*

Topic Rank†	Topic Questions	No. of Stakeholder Votes
	Prioritized future research needs topics	
1	Context of treatment decisions, monitoring, and outcomes: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in detecting disease status in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden)? [*Alternate text:* How accurately does the SFLC assay, compared with tests like electrophoresis or immunofixation, detect treatment response or relapse or disease activity?]	4
2	Context of treatment decisions, monitoring, and outcomes: How does the SFLC assay, either alone or in combination with traditional tests or as part of a definition of treatment response, compare with traditional tests in guiding treatment and management decisions (e.g., timing, duration, or type of treatment or use of adjunctive therapy or stem-cell transplantation) in patients with PCDs? a. Do these comparisons differ by individual PCD populations? b. How does the SFLC result, as a single result, change treatment or direct further investigation in individual PCDs? c. What are suitable metrics (i.e., the best measurements) from the SFLC assay to make treatment decisions for PCD patients (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio)? d. What is the optimal frequency of monitoring with the SFLC assay with and without other tests for guiding management decisions by disease status and by PCD?	5
3	Context of reducing need for other testing: What is the role of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in reducing the need for bone marrow examination by providing equivalent information for monitoring response to treatment, remission, or relapse in individual PCDs?	4
	Other nominated future research needs topics	
4	Context of reducing need for other testing: What is the role of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in reducing the need for 24-hour urine testing in detection of light chains by providing equivalent information: a. for diagnosis of individual PCDs? b. for monitoring response to treatment, remission, or relapse in individual PCDs?	3

Table 3. Nominated topics for future research needs for the SFLC assay in PCDs, according to rank* (continued)

Topic Rank†	Topic Questions	No. of Stakeholder Votes
5[@]	Context of current practice How are clinicians adopting the SFLC assay currently? Is the test an add-on or does it replace an existing, more complicated or more invasive test? If so, are practitioners cutting back on use of other tests (bone marrow biopsy, 24-hour urine) as a result? What is the current frequency of SLFC testing in patients with PCDs?	3
6[@]	Context of standardization of best practice: What are the "best practices" (both technical and clinical) for use of the SFLC assay in different populations with regard to: a. laboratory testing (platform, lot, site) and reporting of results by laboratories b. clinical use in diagnosis, monitoring, and/or prognosis c. indication and prescribing/requisition patterns (e.g., frequency or timing of use)	3
7	Context of diagnostic testing: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests, in a high-risk population with regard to: a. conditions such as antigen excess and polyclonal gammopathy (which can result in false negatives and false positives, respectively) b. diminished kidney function and the need for different reference ranges in kidney failure c. technical variations in commercial assays, assay platform, or in house (hospital) vs. central lab testing d. resolving discordant SFLC and traditional testing results in different PCDs e. diagnostic reclassification based on SFLC testing	2
8	Context of diagnostic testing: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests, in individual PCDs with regard to: a. conditions such as antigen excess and polyclonal gammopathy (which can result in false negatives and false positives, respectively) b. diminished kidney function and the need for different reference ranges in kidney failure c. technical variations in commercial assays, assay platform, or in house (hospital) vs. central lab testing d. resolving discordant SFLC and traditional testing results in different PCDs e. diagnostic reclassification based on SFLC testing f. detecting response and remission, relapse, and assessment of tumor burden g. outcome prediction (response, disease-free survival, overall survival)	2

Table 3. Nominated topics for future research needs for the SFLC assay in PCDs, according to rank* (continued)

Topic Rank†	Topic Questions	No. of Stakeholder Votes
9	Context of MGUS progression to MM How well does the SFLC assay predict progression to MM in patients with MGUS? a. What is its accuracy (strength of prediction as well as precision) as an independent predictor? b. Can it diagnose or detect disease progression earlier than traditional tests? c. Should the SFLC assay with or without other traditional testing be used in monitoring/surveillance and how often? d. Are there variant populations of MGUS where the SFLC is more relevant to predict progression to MM? e. What are suitable metrics (what are the best SFLC measurements) from the SFLC assay to predict progression of MGUS (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio)?	2
10	Context of treatment decisions, monitoring, and outcomes What is the role of the SFLC assay, relative to other tests, in providing prognostic information? a. How well does the SFLC assay predict outcomes such as treatment response, overall survival, or disease-free survival for individual PCD populations? b. Does the test result in improved health outcomes? c. What are suitable metrics from the SFLC assay to predict patient response and outcomes such as overall and disease-free survival (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio)?	2
11	Context of reducing need for other testing: Can the frequency of current SIFE, SPEP, and UPEP testing be decreased if SFLC testing is added?	2
12[@]	Context of cost effectiveness: What is the cost effectiveness of using the SFLC assay as an add-on to other testing?	2
13	Context of diagnostic testing: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests, in the general population with regard to: a. conditions such as antigen excess and polyclonal gammopathy (which can result in false negatives and false positives, respectively) b. diminished kidney function and the need for different reference ranges in kidney failure c. technical variations in commercial assays, assay platform, or in house (hospital) vs. central lab testing d. resolving discordant SFLC and traditional testing results in different PCDs e. diagnostic reclassification based on SFLC testing	1
14	Context of treatment decisions, monitoring, and outcomes What is the role of the SFLC assay, relative to other tests, in managing patients who have undergone stem-cell transplantation?	1
15[@]	Context of cost effectiveness: What are the factors driving both raw costs and other expenditures (e.g., total cost of usage at a given institution) in the SFLC assay? Is the cost variation an issue related to SFLC assay or in house use vs. outsourcing?	1

Table 3. Nominated topics for future research needs for the SFLC assay in PCDs, according to rank* (continued)

Topic Rank†	Topic Questions	No. of Stakeholder Votes
16[@]	Context of newer tests: What are the newer tests available for the diagnosis, prognosis, and treatment monitoring of PCDs and how do they compare with the SFLC test with and without traditional tests?	1

**Polyclonal gammopathy* is seen in a diverse range of inflammatory and neoplastic diseases where clonal expansion occurs across various B-cell populations that produce more than one kind of immunoglobulin as part of a broad immune response unlike monoclonal gammopathy (or PCD) where a single neoplastic clone produces a single kind of immunoglobulin. *Antigen excess* is the condition in which SFLCs are present in great enough excess to form nonprecipitating immune complexes, which may alter the detection of the SFLCs and lead to inaccurate negative test results.

† Prioritized topics (1–3) are ordered logically by clinical content. Other nominated topics are listed in the order they were prioritized by the stakeholder panel.

[@] Topics that were not addressed in the CER. Before initiating new research, a literature search and review is needed to confirm that any new research on the topic would not be duplicative.

Abbreviations: MGUS=monoclonal gammopathy of undetermined significance; MM=multiple myeloma; PCD=plasma cell dyscrasia; SFLC=serum free light chain; SIFE=serum immunofixation electrophoresis; SPEP=serum protein electrophoresis; UIFE=urine immunofixation electrophoresis; UPEP=urine protein electrophoresis.

The three top-ranked FRNs refer to the application of the SFLC assay in the contexts of diagnosis, treatment monitoring, and resource use. All three are of immediate and definitive clinical relevance and will impact the clinical management of patients with PCDs. Questions also emerged with respect to the diagnostic and prognostic role of the assay, especially in comparison with nontraditional testing. An example is the use of positron emission tomography to measure plasma cell tumor load.[21] We assessed the potential of these issues for topic nomination, although none were top-rated.

The three top-ranked FRNs also have similar considerations for proposing a study design. The main reason for each topic being an evidence gap was a paucity of data, suggesting that a wide range of trial designs could be useful. A randomized controlled trial (RCT) would be ideal in all three cases, since RCTs inherently would yield the most robust and generalizable data, but the low prevalence of PCDs and severity of the disease in many patients requiring monitoring both make it unethical to consider a trial in which some patients are randomized to undergo a potentially useful test and others are randomized to forego the test. It would not be feasible or ethical to compare SFLC testing with traditional testing in a controlled fashion, as current practice includes both sets of tests in a diagnostic panel. Moreover, the variable severity and stages of the disease in a referral population adds to the complexity of conducting a trial. Therefore, we considered more feasible study designs and propose two for each of the three topics, as described below. The two designs are not intended to be mutually exclusive, but rather complementary.

High-Priority Future Research Needs Topic 1

What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in detecting disease status in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden)?

Background

Insufficient evidence resulting from the lack of direct head-to-head comparisons regarding the diagnostic utility of the SFLC assay (vs. traditional tests) to inform treatment decisions led to

the identification of this FRN topic from the SFLC CER. Key Question 4 of the CER was, "In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), is the SFLC assay better than traditional tests in indicating how the patient responds to treatment and of outcomes (overall survival, disease-free survival, remission, light chain escape, and quality of life)?" The CER included 11 studies in patients with AL amyloidosis and MM that evaluated the SFLC assay and traditional testing in parallel and examined their relationship to clinical outcomes in PCDs.[15,22-30,31] No direct comparisons between the SFLC assay and traditional tests were performed. Three studies (all quality C) were conducted in patients with AL amyloidosis[15,26, 28] and 8 (3 quality B; 5 quality C) in patients with MM.[22-25,27,29-31] The CER found insufficient evidence for SFLC response as a better predictor of survival than M-protein response in both AL amyloidosis and in MM, as well as for the other outcomes specified. Although SFLC response to therapy appeared to be a consistent predictor of outcomes in AL amyloidosis, there was no evidence that the SFLC assay is superior to traditional tests, as direct comparisons were unavailable. Similarly, there was no evidence to ascertain whether SFLC response was a better predictor of outcomes than traditional tests in MM.

Building on research gaps identified from this Key Question, the aim of this FRN topic is to assess how accurately the SFLC assay, compared with traditional tests, detects treatment response, relapse, or disease activity.

Discussion among stakeholders related to several aspects of the use of the SFLC assay in the context of informing treatment. In consideration of the short half-life of light chains, stakeholders mentioned the potential ability of the assay to make a determination as to whether a specific therapy is working (possibly before traditional testing is able to). Another stakeholder was more cautious, noting that findings from a single test, particularly the SFLC assay, should not alone be used to change treatment, as its increased sensitivity may capture biochemical changes that are not clinically relevant. Other discussion mentioned adjunct tests, such as PET-CT in patients with nonsecretory MM, used with the SFLC assay to adjust treatment. Finally, one stakeholder proposed that a head-to-head study of treatment decisions, informed with and without SFLC results, could potentially be done to ascertain whether results of the assay would affect decisionmaking. However, across the board, stakeholders acknowledged the lack of information on the role of SFLC data to assess disease status.

Proposed Study Designs

Given the paucity of available evidence to completely answer important clinical questions relevant to the SFLC assay in the context of treatment decisions, monitoring, and outcomes, different study designs are reasonable. The following proposed study designs (Table 4) address the diagnostic performance of SFLC testing and traditional testing in detecting disease status in various treatment contexts but differ in terms of their feasibility.

Table 4. Potential research designs for topic 1

Considerations	Meta-Analysis of Individual Participant Data	Prospective Observational Study
Design description	Using completed or ongoing RCTs (e.g., STaMINA trial or others in Appendix B) or registry data as the source, where simultaneous SFLC and traditional testing has occurred, combine all treatment arms (within one or more RCTs or registry) into a single cohort of patients with a diagnosed PCD. Then ascertain the diagnostic accuracy of the SFLC assay against comparator(s), chosen a priori, to detect disease status in different treatment contexts (e.g., response and remission, relapse, or assessment of tumor burden).	Prospectively designed single-cohort studies consisting of patients receiving treatment (in the context of normal clinical practice) for a diagnosed PCD can be conducted to assess the diagnostic accuracy of the SFLC assay against comparator(s) to detect disease status in the different treatment contexts (e.g., response and remission, relapse, or assessment of tumor burden). Studies should have a priori calculation of the sample size needed for determination of the desired precision and should include inferences based on formal statistical testing of estimates of diagnostic accuracy.
Population	Patients with PCDs; each PCD could form a prespecified stratum or subgroup for analysis.	Patients with PCDs; each PCD could form a stratum or subgroup for analysis.
Intervention	SFLC testing and SFLC testing in combination with traditional testing.	SFLC testing and SFLC testing in combination with traditional testing.
Comparator	Traditional testing.	Traditional testing.
Outcomes	Measures of diagnostic accuracy (e.g., sensitivity, specificity, or predictive value) in different treatment contexts (e.g., response and remission, relapse, or assessment of tumor burden).	Measures of diagnostic accuracy (e.g., sensitivity, specificity, or predictive value) in different treatment contexts (e.g., response and remission, relapse, or assessment of tumor burden).
Timing	Trial duration/followup duration in cohorts or registries.	Trial duration/followup duration in cohorts.
Setting	Secondary analysis of preexisting data.	Prospective; clinical practice.
Advantages for producing a valid result	Benefits of large sample sizes and allowing for prespecified subgroup analyses but at the possible cost of bias from pooling varying patient groups; disadvantage of being a retrospective observational data analysis subject to inherent biases and limitations of component studies.	Relatively poor, given bias from lack of randomization.
Resource use, size, and duration	Resource use less than in an RCT but still moderate to high. Sample size not an issue, in spite of the relatively low prevalence of PCDs, because the cohorts would be from RCTs or registries. Duration would depend on acquisition of databases and subsequent analysis and would potentially be shorter than in a prospective cohort or clinical trial.	Resource use likely to be moderate. Size may be an issue due to low prevalence of PCDs. Study duration may vary due to natural history of condition.
Ethical, legal, and social issues	Meta-analysis uses existing data. With permission to use original study data, there should be no ethical issues.	Few ethical issues are likely to occur as the study involves components of normal patient care.
Availability of data/ability to recruit	Good; PCD is a rare disease but use of RCT or registry data would mitigate sample size limitations.	Poor, since PCD is a rare disease.

Abbreviations: PCD=plasma cell dyscrasia, RCT = randomized controlled trial, SFLC=serum free light chain.

High-Priority Future Research Needs Topic 2

Context of treatment decisions, monitoring, and outcomes:

How does the SFLC assay, either alone or in combination with traditional tests or as part of a definition of treatment response, compare with traditional tests in guiding treatment and management decisions (e.g., timing, duration, or type of treatment or use of adjunctive therapy or stem-cell transplantation) in patients with PCDs?

a. Do these comparisons differ by individual PCD populations?
b. How does the SFLC result, as a single result, change treatment or direct further investigation in individual PCDs?
c. What are suitable metrics (i.e., the best measurements) from the SFLC assay to make treatment decisions for PCD patients (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio)?
d. What is the optimal frequency of monitoring with the SFLC assay with and without other tests for guiding management decisions by disease status and by PCD?

Background

The SFLC CER revealed this topic as an evidence gap by finding insufficient evidence for one of the Key Questions that was similar to this topic:
"Key Question 3: In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay result in different treatment decisions as compared with traditional tests? Does the use of the SFLC assay affect the management of patients by allowing for earlier institution of specific therapies? Does the use of the SFLC assay influence the duration of treatment? Does the use of the SFLC assay influence the type of treatment (e.g., radiation therapy)?"

We found no studies that formally compared the use of the SFLC assay with traditional tests to determine whether either led to different decisions with regard to timing, duration, or type of treatment. Although two studies described instances where SFLC testing was used for therapeutic decisionmaking, they were in largely uncontrolled settings and adequate inferential analysis was not possible.[32,33] They did not fulfill the eligibility criteria for inclusion in the CER either. We therefore rated the strength of evidence as insufficient for this question.

Our stakeholders noted that the Key Question 3 was indeed a persistent evidence gap and remains a powerful question for both patients and providers, all of whom would like to understand the value of SFLC testing in the context of treatment and prognosis of PCDs. The degree of paraprotein reduction with treatment is one of the standard criteria used to define the completeness of hematologic response. Stakeholders indicated the need to both evaluate the independent role of SFLC testing and its role in comparison to other available tests that measure paraprotein. The biologic and other characteristics of the tumor often determine the type of paraprotein produced, which must be remembered while making the evaluation or comparison of the effectiveness of SFLC testing versus traditional testing across PCDs. The utility of SFLC testing in the context of treatment and prognosis will be predicated upon the significance of change in light chain production for disease response and outcome. Notwithstanding, SFLC testing has the highest sensitivity for detecting light chain production by tumors.

Bearing these aspects of SFLC testing in mind, the stakeholder panel offered additional qualifications to the original Key Question 3. They asked whether evaluation and comparison of

SFLC testing with traditional testing would differ by individual PCD populations, with specific reference to MM, nonsecretory MM, or AL amyloidosis, and queried the value of a single SFLC result in changing treatment or directing further investigation in individual PCDs. There was also discussion about the best metrics applicable to the SFLC assay to make treatment decisions for PCD patients (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio). Finally, an important aspect of SFLC testing was the definition for the optimal frequency of monitoring with and without other tests for guiding management decisions by disease status and by PCD.

Proposed Study Designs

Because the main reason for this topic being an evidence gap is a lack of systematically and methodically generated data, we considered feasible study designs based on secondary data analysis, including creative use of survey designs, and propose two here (Table 5). These two designs can be considered complementary. The first design addresses and compares the relationship between the results of SFLC testing and traditional testing to response definitions and treatment outcomes in an observational setting, using an individual patient data meta-analytic approach. The second design uses a survey using data from preexisting trials to query physicians blinded to the management or clinical outcome status of the patient to obtain their responses with respect to therapeutic decisionmaking.

Table 5. Potential research designs for topic 2

Considerations	Meta-Analysis of Individual Participant Data	Physician Surveys
Design description	Using completed or ongoing RCTs (e.g., the STaMINA trial in Appendix B) or registry data as the source, where simultaneous SFLC and traditional testing has occurred, combine all treatment arms (within one or more RCTs or registry) into a single cohort of patients with a diagnosed PCD. In this composite study population ascertain whether the SFLC assay is equivalent to the comparator chosen a priori for the definition or prediction of treatment response and outcomes pre-specified in the parent study.	Using data from completed or ongoing RCTs (e.g., the STaMINA trial in Appendix B) or cohort studies among patients with a diagnosed PCD who have had simultaneous SFLC and traditional testing, design surveys for physicians to indicate their responses to a set of patient data with regard to response status, and therapeutic approach in the absence and presence of the SFLC data. Physicians could also be randomized as to whether they are given SFLC results before or after traditional results? Patient data may be categorized by individual PCDs and clinical stage of the disease depending on primary study.
Population	Patients with PCDs; each PCD could form a prespecified stratum or subgroup for analysis	Physicians/providers.
Intervention	SFLC testing, metrics of SFLC testing, frequency of SFLC testing	SFLC testing.
Comparator	Traditional testing, metrics of traditional testing, frequency of traditional testing (prespecified reference categories)	Traditional testing.
Outcomes	Response to treatment, remission, relapse, and light chain escape; overall survival and disease-free survival	Diagnosis of response, remission or relapse; therapeutic approach proposed (timing, type and duration).
Timing	Trial duration/followup duration in cohorts or registries	Trial duration/follow up duration in cohorts or registries.
Setting	Secondary analysis of preexisting data	Patients with PCDs; each PCD could form a prespecified stratum or subgroup for analysis; clinical stage of the disease.
Advantages for producing a valid result	Benefits of large sample sizes and allowing for prespecified subgroup analyses but at the possible cost of bias from pooling varying patient groups; disadvantage of being a retrospective observational data analysis subject to inherent biases and limitations of component studies	Feasible and inexpensive; appropriate methods of randomization of both participants and patient data would reduce bias. Disadvantages inherent to survey methodology including inflexibility and subjectivity. Dependent on the percentage of respondents, quality of participation and responses and the quality of the survey tool used.
Resource use, size, and duration	Resource use less than in an RCT but still moderate to high. Sample size not an issue, in spite of the relatively low prevalence of PCDs, because the cohorts would be from RCTs or registries. Duration would depend on acquisition of databases and subsequent analysis and would potentially be shorter than in a prospective cohort or clinical trial	Resource use likely to be moderate; sample size would depend on number of respondents. Duration would depend on acquisition of databases and subsequent designing of survey and would potentially be shorter than a prospective cohort study or clinical trial.
Ethical, legal, and social issues	Would not be limiting	Would not be limiting.
Availability of data/ability to recruit	Good; PCD is a rare disease but use of RCT or registry data will mitigate sample size limitations	Dependent on number of respondents.

Abbreviations: PCD=plasma cell dyscrasia, RCT = randomized controlled trial, SFLC=serum free light chain

High-Priority Future Research Needs Topic 3

What is the role of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in reducing the need for bone marrow examination by providing equivalent information for monitoring response to treatment, remission, or relapse in individual PCDs?

Background

The SFLC CER revealed this topic as an evidence gap by finding insufficient evidence for one of the Key Questions that was a broader version of this topic, focused on any diagnostic testing, not only bone marrow biopsy: "Key Question 5: In patients with an existing diagnosis of PCD (MM, nonsecretory MM, or AL amyloidosis), does the use of the SFLC assay reduce the need for other diagnostic tests (e.g., bone marrow biopsy)?" Bone marrow examination is of particular interest because they can be cumbersome to perform and uncomfortable for patients, causing considerable noncompliance among physicians.

The evidence for this Key Question was deemed insufficient because only one small, retrospective cohort study) addressed the Key Question; it did, however, address bone marrow examination in particular as a gold standard and concluded that bone marrow examination could not be replaced by the SFLC assay.[34] Specifically, the authors found that in 29 MM patients in whom the SFLC assay was performed, normalization of the SFLC ratio did not guarantee the achievement of complete remission on bone marrow examination (as defined by the gold standard of 5 percent or fewer plasma cells, 10 percent of patients with SFLC ratio normalization did not have complete remission). The study did not address the frequency of use of bone marrow examination for monitoring.

More information on this topic is anticipated from the results of the STaMINA trial of patients with light-chain MM (in which enrollment began in June 2010) by the Center for International Blood and Marrow Transplant Research (ClincialTrials.gov number NCT01109004; see Appendix B), which is prospectively collecting serum samples for light chain analysis (with normal levels required as part of a "stringent complete response" to treatment), along with flow-cytometry measurement of bone marrow and traditional tests for M protein.

Our stakeholders noted that this Key Question was indeed in and of itself an evidence gap and that it remained a powerful question for both patients and providers, all of whom would like to reduce the need for painful, expensive, invasive testing for the monitoring of response to treatment, remission, or relapse in patients with any PCD. Practitioners suggested that they did not anticipate that the SFLC would replace bone marrow biopsy for purposes of diagnosis at least, although it might reduce the need for biopsy later; one patient stakeholder gave anecdotal evidence that SLFC assay can be used alone in monitoring of the light-chain PCDs (nonsecretory MM, AL amyloidosis, or light chain MM).

Another reason for this being an evidence gap is that many of the available studies did not explicitly address comparative effectiveness of the SFLC assay versus traditional tests, instead of just clinical effectiveness of the SFLC assay.

Proposed Study Designs

We propose as two complementary study designs for this topic a prospective cohort study and a prospective observational study (Table 6).

Table 6. Potential research designs for topic 3

Considerations	Prospective Cohort Study	Prospective Observational Study
Design description	Using ongoing RCTs (e.g., the STaMINA trial in Appendix B) or registry data as the source, combine all treatment arms (within one RCT or registry) into a single cohort of patients with a diagnosed PCD with requirement of a complete response to therapy. In this cohort (or across multiple cohorts from multiple, comparable RCTs or registries), ascertain whether the SFLC assay is equivalent to the gold standard (as defined in the RCT or chosen a priori if registry data are used) of treatment response.	Patients with a diagnosed PCD with requirement of a complete response to therapy (not including normal SFLC results) as well as a stringent complete response (including normal SFLC results), with physicians surveyed about what treatment decisions would be in the absence and presence of the SFLC data. Alternative: SFLC done whenever bone marrow done by someone blinded (possible since tests are ordered, not done in office); physicians could be randomized as to whether they are given SFLC results before or after traditional results.
Population	Patients with PCDs.	Physicians/providers.
Intervention	SFLC testing.	SFLC testing.
Comparator	Bone marrow examination.	Bone marrow examination.
Outcomes	Measures of diagnostic accuracy, e.g., sensitivity, specificity, predictive value in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden). Frequency of bone marrow examination.	Measures of diagnostic accuracy, e.g., sensitivity, specificity, predictive value in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden). Frequency of bone marrow examination.
Timing	Trial duration/followup duration in cohorts or registries.	Trial duration/followup duration in cohorts or registries.
Setting	Prospective; clinical practice.	Prospective; clinical practice.
Advantages for producing a valid result	Arguably the optimal design when the treatment cannot ethically be assigned. Would carry most of the benefits of RCTs (or large registries) at the cost of possible bias from pooling of treatment arms.	Relatively poor, given bias due to lack of randomization and possible innate tendency of both physicians and patients to prefer more information to less.
Resource use, size, and duration	Resource use is less than in an RCT but still moderate to high. Size would not be an issue, in spite of the relatively low prevalence of PCDs, because the cohorts would be from RCTs or registries. Trial duration would be relatively short, since and the SLFC assay half-life is short (2 to 6 hours) and the median survival time is about 5 years.	Resource use is likely moderate. Size, and duration should not be an issue, since study would be observing normal clinical practice; sole difference is addition of SFLC assay for all patients rather than performed only at physician's discretion, but assay is affordable (~$25 to $100, anecdotally) and covered by insurance.
Ethical, legal, and social issues	Ethical issues should not be an issue.	Ethical issues should not be an issue.
Availability of data/ability to recruit	Moderate; PCD is a rare disease but use of RCT or registry data should mitigate this.	Poor, since PCD is a rare disease.

Abbreviations: PCD=plasma cell dyscrasia, RCT = randomized controlled trial, SFLC=serum free light chain.

Discussion

On the basis of the 2012 SFLC CER and our discussion with stakeholders, we identified 16 potential research areas, 3 of which were ranked as high-priority areas of future research. The recommendations for priority topics for future research were generated based on a stakeholder-driven nomination and review process. We followed the recently developed 7Ps taxonomy that was designed to aid researchers in the identification, recruitment, and engagement of stakeholders. Our stakeholder panel represented a broad range of perspectives, across all major stakeholder categories identified in this taxonomy.

Across the diversity of representation of different disciplines by stakeholders, a uniform consensus was the complementary role of the SFLC assay in the diagnosis and management of PCDs. Hence the need for future research was mainly directed toward optimizing the use of the assay in various contexts, patient groups, and disease stages.

The three top-ranked FRNs referred to the application of the SFLC assay in the contexts of diagnosis, treatment monitoring, and resource use. These contexts echoed several of the Key Questions in the original CER, albeit with several noteworthy differences. In the diagnostic context, emphasis was placed upon the diagnostic performance of the assay in detecting response, remission, and relapse and assessment of tumor burden in patients with diagnosed disease who are in some stage of therapeutic intervention and followup or disease management. Our original Key Questions addressed the use of the assay in patients with undiagnosed disease and there appeared to be a consensus among stakeholders that the assay had an adjunctive role to the established approach of needing a panel of tests to diagnose PCDs. In the context of therapeutic management of PCD, the prioritized FRN addressed the role of the SFLC assay in guiding treatment and management decisions. The scope of the research question appreciated the heterogeneous nature of PCDs and the likelihood that the application of the assay could differ by individual PCD. Finally, the context of utilizing the SFLC assay in a judicious way to reduce resource utilization, with specific reference to bone marrow examination, extended the understanding of its diagnostic role in patients with known PCD to define disease status such as response to treatment, remission, or relapse.

The study designs we proposed were guided by two major considerations. First, we had to take into account the reality that the SFLC assay was already a test in regular and routine use for the management of PCDs and that many providers believed that it added value to patient care and helped in decisionmaking. Hence it was difficult to visualize a situation where it would be possible to randomize patients to care in the absence of performing the assay. Second, it was also crucial to avoid the very limitations that the CER identified in the studies evaluated for our original Key Questions. Designs had to ensure adequate power to answer questions of interest and minimize biases. Study questions had to address the issues of heterogeneity of PCDs and be explicit and objective in the comparisons desired so that appropriate statistical testing could be applied. In addition, since both the CER and discussion with stakeholders revealed that SFLC results are complementary, not prescriptive on their own, we focused the study designs to capture clinical preferences, although we acknowledge that patient preference would be another avenue of study.

The process was not without limitations. We were able to obtain input from only 8 of the 10 members of our stakeholder panel, and the total number of stakeholders recruited was restricted, thus limiting representation. The threshold for distinguishing high-priority topics from the other topics in the final ranking was somewhat arbitrary, namely four or more votes (i.e., prioritization by 50 percent or more of the eight voters). Also, despite formal planning, the selection of

stakeholders, solicitation of contributions, facilitation of discussion, and synthesis of suggestions remain, to some degree, idiosyncratic. There are as of yet no accepted standard methods by which to assess the validity of procedures to synthesize diverse stakeholder viewpoints. We believe that future methods work may be necessary to establish a formal process for validation, certification, and peer review of FRN rankings.

Conclusions

This report identifies three high-priority FRN topics to study the SFLC assay in PCD patients, which were identified by a stakeholder panel. These are all parts of or repeats of some of the CER Key Questions, for which there was little evidence. They are as follows:

1.   Context of treatment decisions, monitoring, and outcomes: What is the diagnostic performance of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in detecting disease status in different treatment contexts (e.g., response and remission, relapse, and assessment of tumor burden)? [Alternate text: How accurately does the SFLC assay, compared with tests like electrophoresis or immunofixation, detect treatment response or relapse or disease activity?]

2.   Context of treatment decisions, monitoring, and outcomes: How does the SFLC assay, either alone or in combination with traditional tests or as part of a definition of treatment response, compare with traditional tests in guiding treatment and management decisions (e.g., timing, duration, or type of treatment or use of adjunctive therapy or stem-cell transplantation) in patients with PCDs?

   a.   Do these comparisons differ by individual PCD populations?

   b.   How does the SFLC result, as a single result, change treatment or direct further investigation in individual PCDs?

   c.   What are suitable metrics (i.e., the best measurements) from the SFLC assay to make treatment decisions for PCD patients (e.g., absolute concentrations of the involved light chain [kappa or lambda], the absolute difference between the concentrations of the kappa and lambda chains, and the SFLC kappa-to-lambda ratio)?

   d.   What is the optimal frequency of monitoring with the SFLC assay with and without other tests for guiding management decisions by disease status and by PCD?

3.   Context of reducing need for other testing: What is the role of the SFLC assay either alone or in combination with traditional tests, compared with traditional tests alone, in reducing the need for bone marrow examination by providing equivalent information for monitoring response to treatment, remission, or relapse in individual PCDs?

References

1. Kyle RA, Rajkumar SV. Epidemiology of the plasma-cell disorders. Best Pract Res Clin Haematol. 2007;20:637-64.

2. Ries LAG, Harkins D, Krapcho M, et al. SEER cancer statistics review, 1975-2003. NCI, Bethesda, MD; 2005. Available at: http://seer.cancer.gov/csr/1975_2003. Accessed January 30, 2012.

3. Jemal A, Siegel R, Ward E, et al. Cancer Statistics, 2007. CA Cancer J Clin. 2007;57(1):43-66. doi: 10.3322/canjclin.57.1.43.

4. Katzel JA, Hari P, Vesole DH. Multiple myeloma: charging toward a bright future. CA Cancer J Clin. 2007;57(5):301-8. doi: 10.3322/CA.57.5.301.

5. Cook R. An economic perspective on treatment options in multiple myeloma. Managed Care Oncol. 2007; (Spring):10-2.

6. Messori A, Trippoli S, Santarlasci B. Pharmacotherapy of multiple myeloma: an economic perspective. Expert Opin Pharmacother. 2003 Apr 1;4(4):515-24. doi: 10.1517/14656566.4.4.515.

7. Kyle RA, Rajkumar SV. Criteria for diagnosis, staging, risk stratification and response assessment of multiple myeloma. Leukemia. 2009;23(1):3-9.

8. Palumbo A, Sezer O, Kyle R, et al. International Myeloma Working Group guidelines for the management of multiple myeloma patients ineligible for standard high-dose chemotherapy with autologous stem cell transplantation. Leukemia. 2009 Jun 4;23(10):1716-30.

9. Rajkumar SV. Clinical features, laboratory manifestations, and diagnosis of multiple myeloma. In: Basow DS, eds. UpToDate. Waltham, MA: 2011.

10. Bradwell AR, Carr-Smith HD, Mead GP, et al. Highly sensitive, automated immunoassay for immunoglobulin free light chains in serum and urine. Clin Chem. 2001 Apr;47(4):673-80.

11. Rajkumar SV. Recognition of monoclonal proteins. In: Kyle RA, Connor RF, eds. UpToDate. Waltham, MA: 2011.

12. Katzmann JA. Serum free light chains: quantitation and clinical utility in assessing monoclonal gammopathies. Clin Lab News 2006 Jun;12-4.

13. Dispenzieri A, Kyle R, Merlini G, et al. International Myeloma Working Group guidelines for serum-free light chain analysis in multiple myeloma and related disorders. Leukemia. 2009 Feb;23(2):215-24.

14. Mead GP, Carr-Smith HD, Drayson MT, et al. Serum free light chains for monitoring multiple myeloma. Br J Haematol. 2004 Aug;126(3):348-54.

15. Sanchorawala V, Seldin DC, Magnani B, et al. Serum free light-chain responses after high-dose intravenous melphalan and autologous stem cell transplantation for AL (primary) amyloidosis. Bone Marrow Transplant. 2005 Oct;36(7):597-600.

16. Kyle RA, Durie BGM, Rajkumar SV, et al. Monoclonal gammopathy of undetermined significance (MGUS) and smoldering (asymptomatic) multiple myeloma: IMWG consensus perspectives risk factors for progression and guidelines for monitoring and management. Leukemia. 2010 Jun;24(6):1121-7.

17. Bradwell AR. Serum free light chain measurements move to center stage. Clin Chem. 2005 May;51(5):805-7.

18. Whitlock EP, Lopez SA, Chang S, et al. AHRQ Series Paper 3: Identifying, selecting, and refining topics for comparative effectiveness systematic reviews: AHRQ and the Effective Health-Care program. J Clin Epidemiol. 2010;63:491-501.

19. Concannon T, Meissner P, Grunbaum J, et al. A new taxonomy for stakeholder engagement in patient-centered outcomes research. J Gen Intern Med. 2012:1-7.

20. Trikalinos TA, Dahabreh IJ, Wong J, et al. Future Research Needs for the Comparison of Percutaneous Coronary Interventions with Bypass Graft Surgery in Nonacute Coronary Artery Disease: Identification of Future Research Needs from Comparative Effectiveness Review No. 9 [Internet]. Rockville, MD: Agency for Healthcare Research and Quality. September 2010. Available at: www.effectivehealthcare.ahrq.gov. Accessed January 30, 2012.

21. Elliott BM, Peti S, Osman K, et al. Combining FDG-PET/CT with laboratory data yields superior results for prediction of relapse in multiple myeloma. Eur J Haematol. 2011 Apr;86(4):289-98.

22. Dispenzieri A, Zhang L, Katzmann JA, et al. Appraisal of immunoglobulin free light chain as a marker of response. Blood. 2008 May 15;111(10):4908-15.

23. Giarin MM, Giaccone L, Sorasio R, et al. Serum free light chain ratio, total kappa/lambda ratio, and immunofixation results are not prognostic factors after stem cell transplantation for newly diagnosed multiple myeloma. Clin Chem. 2009 Aug;55(8):1510-6.

24. Khoriaty R, Hussein MA, Faiman B, et al. Prediction of response and progression in multiple myeloma with serum free light chains assay: corroboration of the serum free light chain response definitions. Clin Lymphoma Myeloma Leuk. 2010 Feb;10(1):E10-3.

25. Kroger N, Asenova S, Gerritzen A, et al. Questionable role of free light chain assay ratio to determine stringent complete remission in multiple myeloma patients. Blood. 2010 Apr 22;115(16):3413-4;

26. Kumar SK, Dispenzieri A, Lacy MQ, et al. Changes in serum-free light chain rather than intact monoclonal immunoglobulin levels predicts outcome following therapy in primary amyloidosis. Am J Hematol. 2011 Mar;86(3):251-5.

27. Kyrtsonis MC, Vassilakopoulos TP, Kafasi N, et al. Prognostic value of serum free light chain ratio at diagnosis in multiple myeloma. Br J Haematol. 2007 May;137(3):240-3.

28. Lachmann HJ, Gallimore R, Gillmore JD, et al. Outcome in systemic AL amyloidosis in relation to changes in concentration of circulating free immunoglobulin light chains following chemotherapy. Br J Haematol. 2003 Jul;122(1):78-84.

29. van RF, Bolejack V, Hollmig K, et al. High serum-free light chain levels and their rapid reduction in response to therapy define an aggressive multiple myeloma subtype with poor prognosis. Blood. 2007 Aug 1;110(3):827-32.

30. Paiva B, Martinez-Lopez J, Vidriales MB, 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.

31. Dytfeld D, Griffith KA, Friedman J, et al. Superior overall survival of patients with myeloma achieving very good partial response or better to initial treatment with bortezomib, pegylated liposomal doxorubicin, and dexamethasone, predicted after two cycles by a free light chain- and M-protein-based model: extended follow-up of a phase II trial. Leuk Lymphoma. 2011 Jul;52(7):1271-80.

32. Cohen AD, Zhou P, Chou J, et al. Risk-adapted autologous stem cell transplantation with adjuvant dexamethasone +/- thalidomide for systemic light-chain amyloidosis: results of a phase II trial. Br J Haematol. 2007 Oct;139(2):224-33.

33. Hutchison CA, Bradwell AR, Cook M, et al. Treatment of acute renal failure secondary to multiple myeloma with chemotherapy and extended high cut-off hemodialysis. Clin J Am Soc Nephrol. JASN 2009 Apr;4(4):745-54.

34. Chee CE, Kumar S, Larson DR, et al. The importance of bone marrow examination in determining complete response to therapy in patients with multiple myeloma. Blood. 2009 Sep 24;114(13):2617-8.

Abbreviations

AHRQ	Agency for Healthcare Research and Quality
AL amyloidosis	Systemic, or primary, amyloidosis in which amyloid [A] proteins derived from immunoglobulin light chains [L] are deposited in tissue (also called light-chain amyloidosis)
CER	Comparative Effectiveness Review
EPC	Evidence-based Practice Center
FLC	Free light chain
FRN	Future Research Needs
IMWG	International Myeloma Working Group
MGUS	Monoclonal gammopathy of undetermined significance
MM	Multiple myeloma
M-protein	Monoclonal protein (also called paraprotein)
PCD	Plasma cell dyscrasia
PICO (also PICOTS)	Populations, interventions (diagnostic tests/disease monitoring), comparators, outcomes (and timing and settings)
RCT	Randomized controlled trial
SFLC	Serum free light chain
SIFE	Serum immunofixation electrophoresis
SPEP	Serum protein electrophoresis
UIFE	Urine immunofixation electrophoresis
UPEP	Urine protein electrophoresis

Appendix A. Search Strategy for Ongoing Studies

Databases: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations <January 31, 2012>, EBM Reviews - Cochrane Database of Systematic Reviews <2005 to January 31, 2012>, EBM Reviews - Cochrane Central Register of Controlled Trials <1st Quarter 2012>, Ovid MEDLINE(R) without Revisions <1996 to January Week 4 2012>
Last run 1/31/2012

1 Immunoglobulin Light Chain*.mp. or exp Immunoglobulin Light Chains/
2 monoclonal light chain*.mp.
3 serum free light chain*.mp.
4 immunoglobulin-free light chain*.mp.
5 Bence Jones protein.mp. or exp Bence Jones Protein/
6 1 or 2 or 3 or 4 or 5
7 limit 6 to English language [Limit not valid in the Cochrane databases; records were retained]
8 limit 7 to yr="2000 -Current"
9 remove duplicates from 8

Appendix B. Lists of Ongoing Studies

Ongoing Research on the SFLC Assay in ClinicalTrials.gov as of July 7, 2012*

ClinicalTrials.gov Number	Title	Recruitment	Patient Population	Enrollment
NCT00089167	Melphalan, Thalidomide, and Dexamethasone in Treating Patients With Newly Diagnosed, Previously Untreated Primary Systemic Amyloidosis	Completed	MM and plasma cell neoplasm	NR
NCT00301275	Assessing Free Immunoglobulin Light Chains in Patients With Myeloma	Completed	MM	NR
NCT00416897	Dexamethasone and Chemotherapy With or Without Plasma Exchange in Patients With Newly Diagnosed Multiple Myeloma and Acute Kidney Failure	Recruiting	MM and plasma cell neoplasm	280
NCT00478075	Samarium Sm 153 Lexidronam Pentasodium and Bortezomib in Treating Patients With Relapsed or Refractory Multiple Myeloma	Completed	MM and plasma cell neoplasm	50
NCT00499577	Stem Cell Transplant, Chemotherapy, and Biological Therapy in Treating Patients With High-Risk or Refractory Multiple Myeloma	Completed	MM and plasma cell neoplasm	56
NCT00621400	Lenalidomide in Combination With Melphalan and Dexamethasone in Newly-diagnosed Light-chain (AL)-Amyloidosis	Completed	Amyloidosis	27
NCT00910078	Serum Free Light Change in Multiple Myeloma	Completed	MM	NR
NCT00972712	Protocol of the Combination of Bortezomib and Tipifarnib for Relapsed or Refractory Multiple Myeloma	Active, not recruiting	MM	35
NCT00637767	High-Dose Melphalan With or Without Radiolabeled Monoclonal Antibody in Treating Patients With Multiple Myeloma Undergoing an Autologous Stem Cell Transplant	Recruiting	MM and plasma cell neoplasm	90
NCT00681044	High-Dose Melphalan and Stem Cell Transplant in Treating Patients With Immunoglobulin Deposition Disease or Light-Chain Deposition Disease	Recruiting	MM and plasma cell neoplasm	30
NCT00821301	Study of Fluphenazine in Relapsed or Relapsed-and-Refractory Multiple Myeloma	Recruiting	MM	30
NCT00881920	Autologous T Lymphocytes Engrafted With a Chimeric Antigen Receptor Targeting the Kappa Light Chain	Recruiting	Lymphoma Myeloma Leukemia	18
NCT01084837	Study of Induction Treatment With Velcade and Dexamethasone for Previously Untreated Patients With Multiple Myeloma and Renal Failure	Recruiting	MM	60
NCT01194791	Lendexal in Patients With Primary Systemic Amyloidosis (AL) Newly Diagnosed	Recruiting	Primary systemic amyloidosis	30
NCT01208818	Studies in Patients With Multiple Myeloma and Renal Failure Due to Myeloma Cast Nephropathy	Recruiting	Chronic renal failure with uremic nephropathy	284
NCT01277016	A Trial for Systemic Light-chain (AL) Amyloidosis	Recruiting	AL amyloidosis	110
NCT01383759	Bortezomib/Dexamethasone (BD), Followed By Autologous Stem Cell Transplantation and Maintenance Bortezomib/Dexamethasone For the Initial Treatment of Monoclonal Immunoglobulin Deposition Disease (MIDD) Associated With Multiple Myeloma and AL Amyloidosis	Recruiting	Light chain and heavy chain deposition disease Monoclonal Immunoglobulin deposition disease Amyloidosis	35

ClinicalTrials.gov Number	Title	Recruitment	Patient Population	Enrollment
NCT01423344	Clinical Evaluation of the Serum Free Light Chain Analysis	Recruiting	MM	30
NCT01109004**	Stem Cell Transplant With Lenalidomide Maintenance in Patients With Multiple Myeloma (BMT CTN 0702)	Recruiting	MM	750
NCT00120263	Trial of Plasma Exchange for Acute Renal Failure at the Onset of Myeloma	Completed	MM Acute Renal Failure	92
NCT00135187	Study of Combination Therapy With VELCADE, Doxil, and Dexamethasone (VDd) in Multiple Myeloma	Completed	MM	30
NCT00344526	Intensive Versus Conventional Treatment in Patients With Primary Amyloidosis	Completed	Primary systemic amyloidosis	100
NCT00726466	Study of Efalizumab Combined With Intravitreal Ranibizumab in the Treatment of Age-Related Macular Degeneration	Withdrawn	AMD CNV	10
NCT00787969	Rituximab, Cladribine, and Temsirolimus in Treating Patients With Newly Diagnosed Mantle Cell Lymphoma	Active, not recruiting	Lymphoma	74
NCT00981097	Study of Blood and Tissue Samples From Patients With Aggressive Non-Hodgkin B-Cell Lymphoma or Hodgkin Lymphoma	Recruiting	Lymphoma Nonneoplastic Condition	50
NCT01067313	Continuous Hemodialysis With an Enhanced Middle Molecule Clearance Membrane	Completed	Septic Shock	24
NCT01237054	Imaging in MGUS, SMM and MM	Completed	MM Smoldering MM MGUS	31
NCT01308073	Stability of Middle Molecule Clearance	Recruiting	Continuous Hemodialysis	15
NCT01541527	Non Neutralizing Antibodies: Prevalence and Characterization	Recruiting	Hemophilia A	300

Abbreviations: AL amyloidosis = systemic amyloidosis in which amyloid [A] proteins derived from immunoglobulin light chains [L] are deposited in tissue, AMD = Age-Related Macular Degeneration, CNV = choroidal neovascularization, MGUS = Monoclonal Gammopathy of Undetermined Significance, MM = multiple myeloma, NR = not reported.

* Search terms were as follows: [Immunoglobulin Light Chain* OR monoclonal light chain* OR serum free light chain* OR immunoglobulin-free light chain* OR Bence Jones protein*]. We also conducted a search of published randomized controlled trials but did not find any of interest that were published after the CER was completed.

** Identified by a reviewer of the SFLC comparative effectiveness review rather than through this ClinicalTrials.gov search. After adding this, because it was not captured in search, we ran an alternative search strategy crossing four new terms (Multiple myeloma* OR Plasma cell neoplasms* OR Amyloidosis* OR Paraproteinemias*) with the original terms (Immunoglobulin Light Chain* OR monoclonal light chain* OR serum free light chain* OR immunoglobulin-free light chain* OR Bence Jones protein*). This alternative search yielded only 23 studies—less than the updated original and still missing the NCT01109004 study. Thus we report the results of our original search in this table.