ID#	Council/ Public	Commenter	Submitter Organization	On Behalf of	Question/Topic
1	Public	Natalie Landman	Arizona State University	Denis Cortese	General Comments
2	Member/ Health Professionals Council	Rebecca Hancock	American Academy of Ophthalmology		Defining Episodes

3	Member/ Health Professionals Council	Rebecca Hancock	American Academy of Ophthalmology	Rebecca Hancock	Introduction
4	Member/ Health Professionals Council	Rebecca Hancock	American Academy of Ophthalmology	Rebecca Hancock	Understanding Episode Groupers
5	Member/ Health Professionals Council	Rebecca Hancock	American Academy of Ophthalmology	Rebecca Hancock	Submitting Episode Groupers for Evaluation

6	Member/ Health Professionals Council	Rebecca Hancock	American Academy of Ophthalmology	Rebecca Hancock	Evaluating Episode Groupers for Endorsement
7	Member/ Health Professionals Council	Rebecca Hancock	American Academy of Ophthalmology	Rebecca Hancock	Conclusion/Next Steps
8	Member/ Health Professionals Council	Rebecca Hancock	American Academy of Ophthalmology	Rebecca Hancock	Considerations for NQF Endorsement of Episode Groupers

9	Member/ Health Professionals Council	Rebecca Hancock	American Academy of Ophthalmology	Rebecca Hancock	General Comments
10	Member/ Supplier Industry Council	Jennifer Van Meter	PhRMA		Submitting Episode Groupers for Evaluation
11	Member/ Supplier Industry Council	Jennifer Van Meter	PhRMA		General Comments

12	Member/ Provider Organizations Council	Mary Wheatley	AAMC		Introduction
14	Member/ Provider Organizations Council	Michael Phelan	Cleveland Clinic	Michael Phelan	General Comments
15	Member/ Health Plan Council	Carmella Bocchino	America's Health Insurance Plans		Defining Episodes

16	Member/ Health Professionals Council	Koryn Rubin	American Medical Association		General Comments
17	Member/ Health Professionals Council	Koryn Rubin	American Medical Association		Submitting Episode Groupers for Evaluation
18	Member/ Health Plan Council	Carmella Bocchino	America's Health Insurance Plans		Submitting Episode Groupers for Evaluation

19	Member/ Health Plan Council	Carmella Bocchino	America's Health Insurance Plans		Evaluating Episode Groupers for Endorsement
20	Member/ Health Plan Council	Carmella Bocchino	America's Health Insurance Plans		Considerations for NQF Endorsement of Episode Groupers
21	Member/ Health Plan Council	Carmella Bocchino	America's Health Insurance Plans		General Comments

22	Member/ Health Professionals Council	Koryn Rubin	American Medical Association		Evaluating Episode Groupers for Endorsement
23	Member/ Health Professionals Council	Koryn Rubin	American Medical Association		Introduction
24	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Introduction

25	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Introduction
26	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Defining Episodes
27	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Understanding Episode Groupers

28	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Understanding Episode Groupers
29	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Understanding Episode Groupers
30	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Understanding Episode Groupers

31	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Submitting Episode Groupers for Evaluation
32	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Evaluating Episode Groupers for Endorsement
33	Member/ Provider Organizations Council	Ellen Schwalenstocker	Children's Hospital Association		Consideration for NQF Endorsement of Episode Groupers

34	Member/ Health Professionals Council	Georgia Hearn	American Society of Nuclear Cardiology	Georgia Hearn	General Comments
35	Member/ Quality Measurement, Research and Improvement Council	Donald Casey	The American College of Medical Quality	Don Casey	Evaluating Episode Groupers for Endorsement
36	Member/ Quality Measurement, Research and Improvement Council	Donald Casey	The American College of Medical Quality	Don Casey	Evaluating Episode Groupers for Endorsement

37	Member/ Quality Measurement, Research and Improvement Council	Donald Casey	The American College of Medical Quality	Don Casey	Considerations for NQF Endorsement of Episode Groupers
38	Member/ Quality Measurement, Research and Improvement Council	Donald Casey	The American College of Medical Quality	Don Casey	Submitting Episode Groupers for Evaluation
39	Member/ Quality Measurement, Research and Improvement Council	Donald Casey	The American College of Medical Quality	Don Casey	Understanding Episode Groupers

40	Member/ Quality Measurement, Research and Improvement Council	Donald Casey	The American College of Medical Quality	Don Casey	Understanding Episode Groupers
41	Member/ Quality Measurement, Research and Improvement Council	Donald Casey	The American College of Medical Quality	Don Casey	Understanding Episode Groupers
42	Member/ Quality Measurement, Research and Improvement Council	Donald Casey	The American College of Medical Quality	Don Casey	Understanding Episode Groupers

43	Member/ Quality Measurement, Research and Improvement Council	Donald Casey	The American College of Medical Quality	Don Casey	General Comments
44	Member/ Health Professionals Council	Eileen Hagan	American College of Cardiology	Joseph Drozda	General Comments
45	Member/ Health Professionals Council	Eileen Hagan	American College of Cardiology	Joseph Drozda	General Comments

46	Member/ Purchaser Council	Corette Byrd	CMS		General Comments
47	Member/ Purchaser Council	Corette Byrd	CMS		General Comments
48	Member/ Purchaser Council	Corette Byrd	CMS		General Comments

49	Member/ Purchaser Council	Corette Byrd	CMS		General Comments
50	Member/ Purchaser Council	Corette Byrd	CMS		General Comments
51	Member/ Purchaser Council	Corette Byrd	CMS		General Comments

National Quality Forum

Comment Report for Episode Grouper Evaluation Criteria Report

Comments received as of 5/13/2014

Comment
There are a number of key elements in this report: 1. Patient-centered episode. This is the heart of the matter, the need to create payment models that revolve around the patient and reward quality and efficiency. Although some NQF panelists support a provider-centered approach, this would defeat one of the key purposes of bundling payments: encouraging more integration/coordination of care. 2. Flexibility. The report suggests that development of useful groupings should be based on the specific purpose for which they will be used. We agree. There is no "one size fits all" approach. One of the big problems with Medicare has been the approach that everyone should do things in the one manner that CMS prescribes.
The Academy agrees with the expert panel that an acute episode would have a defined period of time for the episode. In general, the Academy believes that the episode for a surgical procedure like cataract surgery should be the global period, which is often 90 days. Chronic conditions can persist without end, and thus, in ophthalmology, the goal becomes preventing further vision loss and total blindness. Because of the nature of chronic conditions, we agree that it is difficult to define a period of time for an episode, but believe that defining the episode on an annual basis may make sense. We understand the challenges associated with identifying and attributing costs for a patient with multiple chronic conditions. However, it is important that before a grouper is endorsed by NQF or implemented by any entity that there be an adequate way to account for multiple co-occurring conditions. Further, for many ophthalmology procedures needed because of a patient's chronic eye disease, medical management is a significant portion of the patient care. Management of complications and exacerbation of diseases are appropriate medical care and necessarily cost more, yet episode grouper

On behalf of the American Academy of Ophthalmology (the Academy), we thank the expert panel and NQF for their work on this important issue. We appreciate the opportunity to comment on the draft report for evaluating episode groupers. The Academy is the world's largest association of eye physicians and surgeons—Eye M.D.s—with 20,000 members in the US.

The Academy supports efforts to measure quality and improve performance and recently launched IRISTM, the nation's first comprehensive eye disease clinical registry, which is a centralized data repository and reporting tool that uses EHRs to collect patient data and perform analyses. Participants can access national and inter-practice benchmark reports on demand, which validate the quality of care ophthalmologists provide and pinpoint opportunities for improvement. Registries like IRISTM will lead to an accurate understanding of the value of healthcare services, and will drive action for improvement by providing timely, transparent and relevant feedback to providers.

The Academy agrees that transparency needs to be a key component of the evaluation of episode groupers. In order to understand episode groupers and thus accurately evaluate the level of transparency, it is crucial to understand the challenges associated with appropriately grouping ophthalmologic episodes. Below the Academy highlights some of these challenges:

In eye care, variation related to use of imaging, choice of injected drug, and frequency required of injections is due to the complexities associated with eye disease, and not a misuse of resources or inadequate provision of care by a provider.

The costs of many ophthalmologic medications, particularly anti-vegf agents to treat several chronic retina diseases, vary drastically, and the choice of biologic is often limited by state rule and is not reflective of the physician's decision on which treatment would be most cost effective. The

The Academy believes that information including the length of the episode and risk adjustment methodologies should be submitted to NQF, along with evidence that the grouper has shown valid results and accurate attribution after testing on real world administrative claims data in a large population sample. Additionally, the Academy agrees that in submitting an episode grouper for evaluation, required information should include a grouper's specific purpose and intended use, as well as limitations of the grouper. If any of the issues presented in our previous comments are present, this should be stated upfront. The Academy also agrees that the developer's clinical logic should also be required, including the approach for assigning individual services or claims to an episode.

Acknowledging that it is difficult to assign claims to an episode for patients with multiple co-occurring conditions, transparency around this is key. The Academy shares the panel's concern that limitations resulting from the use of claims data make it difficult to adequately risk adjust.

We agree that validity is an important factor that should be incorporated into the evaluation criteria for episode groupers. This criterion should ensure that the grouper will not inappropriately compare and categorize physicians across subspecialties. We agree that the validity criterion should verify that episodes are clinically appropriate and that claims are assigned in alignment with the actual clinical course. In addition to reliability and validity, the Academy further encourages the inclusion of clinical relevance as a factor in the evaluation criteria for episode groupers, as we believe many clinical measures currently included in physician profiling programs are not important or useful. The Academy supports the feasibility criterion and believes this factor for evaluation should examine a grouper's ability to account for laterality and accurately adjust for severity. It may be an oversight, but we fail to see mention of the need to evaluate the fraction of variation explained by the grouper model or the coefficient of determination known as R squared. This is key in any evaluation.
While the panel indicates that commercial sector groupers have been in the market for a number of years, we reiterate that we know of no private payer that has successfully implemented an episode grouper for any chronic ophthalmology disease. In fact, the providers who have been able to review their patient-level data from commercial groupers have found extreme inaccuracies and frank mistakes for many ophthalmology conditions. Given the complexities inherent with applying episode groupers to ophthalmology, we encourage the inclusion of appropriate stakeholders in the development of any public as well as private episode groupers prior to the endorsement of any grouper.
While the Academy understands that it may not be feasible to require that a grouper allow for national comparisons, we stress that it is similarly not reasonable for groupers to only allow for comparisons on a small scale basis, as there are not enough cases to account for variances and enable fair comparisons, and also it is not reasonable for groupers to allow for comparisons across subspecialties as patient populations vary significantly. The Academy agrees with the panel that selecting individual measures for application without considering how costs are assigned to other occurring conditions may be misleading, and supports the panel's notion that the CMS Episode Grouper undergo NQF endorsement at both the grouper level, episode level, and at the individual level prior to selection for use by the MAP. We again stress that we do not believe that the eye care groupers reviewed in the Brandeis project embodied the transparency that we hope will be required for NQF endorsement, and do not believe that the eye care groupers reviewed should be endorsed.

The Academy has a number of concerns that must be addressed prior to the endorsement of any episode grouper. The Academy highlights the inherent weakness of claims data to translate to accurate clinical information. ICD codes do not have enough granularity or laterality, and do not capture important information for chronic diseases, such as controlled vs. uncontrolled, or exacerbation of a previously controlled condition. Further, physicians often evaluate conditions during office visits that will not be coded, as they have never been instructed or required to code on every condition. This is different than coding in a hospital-setting. In order to properly risk-adjust, physicians will need to be instructed to code with every available diagnosis that is being evaluated at every visit. Other concerns around episode groupers include the sources for data, such as drug and hospital service costs, are not under physician control, yet groupers often evaluate physicians with that data. Additionally, we are concerned because physicians have no control over the care provided to a patient before a patient is referred to them, but groupers often unfairly attribute the costs of care received prior to the final physician.
The Pharmaceutical Research and Manufacturers of America (PhRMA) appreciates that the Expert Panel developed a detailed list of information that developers must submit for evaluation of an episode grouper for endorsement. One of the requirements is for the vendor to outline the maintenance process to be conducted to ensure that an episode grouper keeps pace with standards of care and clinical guidelines. We agree that this process is critically important as it is essential for ensuring that groupers keep pace with changes in medical practice that can result from an evolving evidence base and advances in diagnostics, treatments, services and delivery models within the episode of care being grouped for evaluation. The vendor should have a publicly defined, responsive, flexible maintenance process that will allow for updates and a means for ensuring that their customers receive such updates.
While the Pharmaceutical Research and Manufacturers of America (PhRMA) recognizes that this NQF project focuses on how to appropriately evaluate episode groupers, we note that creation of an episode grouper is challenging. As noted by NQF, the field of episode grouper development is still evolving and the NQF report itself noted “the many challenges to expanding evaluation, and potentially endorsement, beyond individual measures to episode groupers.” We appreciate the considerable work NQF has done to identify many important technical aspects of grouper development, and believe that when episode groupers are developed, they should be grounded in sound process and methods. This includes ensuring the methods and evidence base used in defining both the episode and grouper are transparent; time dimensions of the clinical episode are clear and objective; accounting for variations in the treatment population; and linking the episode to relevant, endorsed measures of clinical quality.

The AAMC appreciates the opportunity to comment on the NQF Proposed Recommendations for Evaluating Episode Groupers. Episode grouper design and measurement is complex and the expert panel summarized many of the key issues to evaluating groupers. In general, the AAMC supports the recommendations in the report, with a few suggested revisions. The AAMC believes that the methodology by which an episode grouper handles complex patients with multiple concurrent episodes should be an explicit part of the evaluation criteria for episode groupers. While NQF proposes to use a "patient-centric episodes of care framework," there is always an inherent tension between holistically reviewing a patient's clinical history versus analyzing an episode of care. By design, episodes are disease or procedure-centric, documenting only a portion of the clinical activity that occurs to the patient over a specific period of time. As the report notes, patients can often have more than one concurrent episode. Developers should be required to document how the episode grouper treats complex patients with multiple episodes, as well as document any considerations and limitations for analyzing this population. Similarly, the NQF steering
This is a very complex and important space. Overall we agree with the episode must be specific to the use case. One consideration not addressed is the difference between medical condition groupers and procedural focused grouper. We expected this distinction to be discussed in the clinical logic section lines 255-269 and we would recommend some elaboration. One concern is that it can be harder to assign services (attribute) to a medical episode if multiple medical episodes are occurring concurrently. We did not see the issue of attribution addressed adequately since patients may be seen in multiple providers or health systems within a grouper, which is then all billed to the same payer. The episode grouper also needs some risk adjustment perhaps including, social demographic factors, severity of illness, and or even patient specific characteristics like noncompliance, understanding that more complex patients will cost more. Ideally, there should be offsetting quality metrics and or appropriateness measures to assure standard of care are followed. Patient-centric episodes are preferable to provider-centric episodes because physicians in a health care system work together to deliver "a series of
Defining the episode in terms of patient centric measures should be articulated more clearly and occur earlier in this section of the document. As written, a population health diagram is used initially and it is not until later in the document that the patient centric nature of the episode is described.

The AMA commends the episode grouper taskforce for putting together a thoughtful report. We are glad to see the report cite the administrative burden episode groupers pose on physicians.

The AMA would like to see refinements to the introduction that deals with the policy landscape. Specifically, we are concerned with the report's over simplification of CMS' Physician Value Based Modifier (VBM) program requirements. The report states that the Value Based Physician Modifer rewards physicians who successfully participate in the PQRS program with an an upward adjustment in their Medicare payment. However, successful participation in the PQRS program is only one component of the VBM calculation and even those who successfully participate in PQRS could still see a negative or neutral adjustment in their Medicare payments. Also, while we agree with the report's conclusion that episode groupers are likely to become an important part of the physician feedback reports that underpin the VBPM, it should be recognized that the VBPM adjustments in any given year are based on performance two years earlier. Thus any adjustments that employ episode groupers could not occur prior to 2017. The QRUR reports

The AMA would have liked for the panel and the report to provide greater more direction on how NQF will address risk adjustment and patient severity when evaluating an episode grouper. The current report provides no real direction and we foresee this area hindering the ability to properly evaluate a grouper when groupers comes forward to NQF for endorsement. We understand the need for variation based on the type of episode constructed; however, some thought about how this should occur would greatly assist with future steering committee evaluation work. Otherwise, we foresee this as a major stumbling block. At the very least an expanded discussion of the various approaches that could be used and the pros and cons of each depending on intended use would be helpful.

We recommend adding to the list of attributes that the episode grouper is clinically sound, reflecting anticipated elements in the natural course of that episode. This concept does not appear until later in this section and should be considered earlier.

It is critical that the episode grouper logic and definitions be transparent. In particular, we recommend that the list of elements required for submission for evaluation of an episode grouper include any limitations of the episode grouper, such as whether the provider attribution logic can be used to group providers together by practice for profiling and accountability purposes.

Testing the episode grouper should include information on specificity and sensitivity of the grouper in identification of the targeted population.

We support the principle that “output should be actionable and usable for performance and quality improvement” but recommend revising this principle to explicitly list affordability in addition to performance and quality improvement.
We are supportive of endorsement at the episode grouper level but recommend caution as endorsement will provide market advantages to individual vendors. Endorsement should be based on clear and robust standards in the areas identified by the document and multiple vendors should be considered.
The Expert Panel was convened to recommend criteria by which episode groupers should be evaluated for NQF endorsement. We agree with this approach rather than one that prescribes a specific episode grouper. Due to the challenges associated with chronic conditions, such as defining the clean period and the role of team-based care, we recommend that NQF consider the need for a multiple chronic condition episode grouper. Given that one of the primary goals of implementing episode groupers is to improve outcomes, we recommend that the report emphasize the need for outcomes measures derived from episode groupers.

Evaluation Criteria- Reliability: We agree with the statement on page 12 that episode developers should "demonstrate how the episode grouper performs across data sets of variable size and multiple data sets." We are concerned that the criteria summary (top of page 14) says that groupings should be repeatable and consistent "a high proportion of the time when assessed with the same data in the same time period."

Evaluation Criteria-Validity: Experience with the spate of inaccurate press reports triggered by Medicare's recent public release of 2012 claims data for individual physicians highlighted the importance of laying out a data set's limitations, the possible consequences of those limitations and the steps that have been taken to address the limitations. We are pleased that NQF has focused on this issue and hope that it will continue to do so as episode groupers come in for evaluation.

Evaluation Criteria-Usability: The release of individual physician data also underscores the importance of the report's conclusion that in evaluating a

The AMA would like to see refinements to the introduction that deals with the policy landscape. Specifically, we are concerned with the report's over simplification of CMS' Physician Value Based Modifier (VBM) program requirements. The report states that the Value Based Physician Modifer rewards physicians who successfully participate in the PQRS program with an an upward adjustment in their Medicare payment. However, successful participation in the PQRS program is only one component of the VBM calculation and even those who successfully participate in PQRS could still see a negative or neutral adjustment in their Medicare payments. Also, while we agree with the report's conclusion that episode groupers are likely to become an important part of the physician feedback reports that underpin the VBPM, it should be recognized that the VBPM adjustments in any given year are based on performance two years earlier. Thus any adjustments that employ episode groupers could not occur prior to 2017. The QRUR reports also are not timely which makes it extremely difficult for physicians to change behavior to improve care and reduce costs.

On behalf of more than 220 member children's hospitals across the country, the Children's Hospital Association (the Association) appreciates the opportunity to comment on the proposed recommendations for evaluating episode groupers. We agree that there remains a great deal to learn about evidence-based approaches, highlighting the importance of multi-stakeholder review. Although the focus of the report appears to be on the use of episode groupers to assess utlization and costs, the Association agrees with other commenters on the importance of assessing utlization and costs in combination with quality measures, particularly outcomes measures. Assessing utilization and costs in isolation is likely to result in unintended consequencs, especially for vulnerable populaitons and individuals with multiple co-occurring conditions.

The Children's Hosptal Association strongly agrees with the last sentence of the introduction regarding the need for testing the use of a grouper if applied to populatons other than it was developed for (e.g., "across the lifespan"). We believe this point merits further emphasis. In addition, information regarding the data sources and characteristics of the population used to test the episode grouper should be required for submissions, considered in the evaluation of and included with the discussion of the intended use (and limitaitons of) the grouper. Experience has shown that existing approaches cannot simply be extended to other populations without additional evaluation and input from approopriate experts.
The Children's Hospital Association agrees that a major challenge in defining episodes is "determining when and how to attribute claims for the treatment of conditions that occur as complications of the underlying condition or procedures."
The Children's Hospital Association further agrees that there are challenges associated with assigning claims to condition-specific episodes when multiple co-occuring conditions are present as well as associated difficulties of maintaining adequate sampe size and accounting for patient severity and risk. The epidemiology of disease differs in children as compared to adults, and relatively small numbers of cases are concentrated in specific conditions. A small group of children, defined as children with medical complexity are characterized as having complex acute and chronic conditions, numerous comorbidities and a broad range of mental health and psychosocial needs. These children have diagnoses that are multiple and varied, are usually under the care of multiple pediatric specialists and require additional services within and outside of the care system. Although children with medical complexity comprise only about 6 percent of children enrolled in Medicaid, they account for approximately 40 percent of costs. Because their conditions vary widely, condition-specific approaches may not be useful or feasible; however, mechanisms for identifying groups such as this and aggregating informaiton on costs and quality (such as measures related to care coordination) would be very useful.

The Children's Hosptial Association further agrees that there are challenges associated with assigning claims to condition-specific episodes when multiple co-occuring conditions are present as well as associated difficulties in maintaining adequate sample size and accounting for patient severity and risk. The epidemiology of disease varies in children as compared to adults with a lower percentage of children experiencing chronic conditions as compared to adults and relatively small numbers of cases concentrated in specific conditions.
A small number of children, defined as children with medical complexity, are characterized as having complex acute and chronic conditions, numerous comorbidities and a broad range of mental health and psychosocial needs These children have diagnoses that are multiple and varied, are usually under the care of multiple pediatric specialists and require additional services within and outside of the health care system. Although children with medical complexity comprise only about 6 percent of children enrolled in Medicaid, they account for approximately 40 percent of costs. Because their clinical conditions vary widely, condition-specific approaches may not be useful or feasible; however, mechanisms for identifying groups such as this and aggregating costs and quality measures (such as measures related to care coordination) would be very useful.
As noted above, the Children's Hospital Association believes it is essential that detailed information regarding the limitations of the grouper as well as its intended purpose be included with the descriptive invormation included with the submission. The intended purpose should include a discussion of the specific population to be included in the application as well as the specific use (e.g., public reporting, payment). Information regarding testing should fully describe the characteristics of the population that was sampled. Complete information regarding the database(s) used to develop the grouper methodology should also be provided. In addition, the documentation should note explicitly whether or not the grouper uses a risk adjustment methodology as well as describing the methodology if used.

As noted above, the Children's Hospital Association believes it is essential that detailed information regarding the limitations of the grouper as well as its intended purpose be included with the descriptive information. The intended purpose should include a discussion of the specific population to be included in the application as well as the specific program use (e.g., public reporting, payment). Information regarding testing should fully describe the characteristics of the population that was sampled. Complete information regarding the database(s) used to develop the grouper methodology should also be provided. In addition, the documentation should note explicitly whether or not the grouper uses a risk adjustment methodology as well as describing the methodology if used.
It will be important for the National Quality Forum to translate the panel discussion of core principles into more defined criteria and requirements for endorsement. The Children's Hospital Association agrees with the panel's discussion on validity, particularly with regard to disclosure of the limitations of the grouper methodology.
As noted above, the Children's Hospital Association agrees that it is important that the output of episode groupers include both quality and cost signals.

National Quality Forum

1030 15th St NW

Washington, DC 20005

Re: Comment on the Proposed Recommendations for Episode Grouper

In 2008, AHRQ published online a report on a systematic review of health care efficiency measures, which was not referenced in the draft report, but should be.

Included in this report was an important list of issues identified by various stakeholders for whom efficiency measurement (which includes using episode groupers) is of critical importance.

Issues of greatest concern to most stakeholders are related to:
1. Data aggregation and quality: which organizational entity should provide, clean, and aggregate data files; will data be easily accessible; are data complete and populated correctly for evaluation; are complete, accurate encounter data available for capitated payment arrangements?
2. Cost calculation: whether to use standardized costs vs. actual costs (it is especially complicated in regions in which providers are heavily capitated

In addition, the Committee is also encouraged to review and specifically consider the published article by CMS (Thomas, et al, Health Care Financing Review, Fall 2009 Volume 31 Number 1 p 51-61) which provides a useful summary of clinician feedback on the most obvious and important decisions that must be faced by Medicare to use grouped claims data as the foundation for a physician performance measurement system

One question that looms large and unaddressed in the Committee's report is how the transition from ICD-9 to ICD-10 will be or is handled and how does this transition get measured against historical performance of the grouper methods developed using ICD-9? Some detailed quantitative information comparing the relative performance of both systems should be provided.
While different methodologies from the various commercial groupers available limit comparability of results among users of these various systems, there should still be an evaluation of each method against some sort of gold standard or benchmark. Additionally, there should be transparent comparisons between these methods that include statistical validity, accuracy, precision and impact of implementation.
Claims-based groupers should be cross-validated against clinical data (e.g. from electronic health records). For many clinical conditions, ICD-9 codes are notoriously unreliable, inaccurate and often do not contain adequate information for risk adjustment. For example, ICD-9 codes for heart failure do not contain information about the stage of heart failure according to ACC/AHA criteria, or New York Heart Association (NYHA) functional class.

Another example of this challenge is the combination of information captured through a comprehensive geriatric assessment multidimensional frailty score of individual patients. Such additional types of these data would provide invaluable and clinically validated information for better risk adjustment.
Episode construction should include assumptions about the most currently available nationally available clinical practice guidelines, appropriate use criteria and patient preference sensitive interventions and technologies. Each episode category should include an analysis and cost of those treatments and diagnostic evaluations with the highest quality of evidence and strength of recommendation. For example, the 2013 ACC-AHA Heart Failure guideline includes a series of strong recommendations based upon high quality of evidence ("1A") that includes achievement of delivery of guideline directed medical therapy for several medication classes as well as evaluation and appropriate usage of imaging studies, cardiac resynchronization therapy, cardiac defibrillator implantation and left ventricular assistive devices.
Not mentioned in this report is the issue of how/if socioeconomic status should be addressed and evaluated in the context of a condition-specific episode. Those individuals with or without supplemental insurance coverage, proportionately high out-of-pocket expenses for co-pays and deductibles related household income should also be considered and evaluated where feasible.

Lastly, we believe that each member of the committee should publicly disclose any relevant relationships with industry (financial and non-financial) relative to this topic.

The requirements for submitting episode groupers and the criteria for evaluating them seem reasonable. That said, the framework described in the report is complex and conflicted and leaves the reader confused about whether episodes of care or episode groupers are the focus.

Key areas needing further discussion and emphasis:

1.The principal purpose of groupers is to track costs of care around an episode of illness and are actually, then, "disease or condition centric "as opposed to provider or patient centric. Perhaps we are confusing this concept with attribution. Grouper methodologies can be patient-centric if they appropriately handle multiple episodes involving single patients, e.g., concurrent episodes or sequential/related episodes. For instance, a patient could have a chronic episode of diabetes and suffer a STEMI (an acute episode). Those 2 conditions play off each other in ways that directly implicate both the quality and costs of care. As far as attribution goes, we support advocating for such at levels above that of the individual practitioner, for example group

Key areas needing further discussion and emphasis (continued):

2. Episode groupers must include clinical data from electronic health recordsor registries as well as claims data.

3. Quality measures and, in particular, outcomes must be integral to any grouper

4. Grouper methodology must be transparent and updated regularly to align with clinical guidelines

5. Groupers intended for public reporting or value-based payment must be fully tested and vetted prior to use.

CMS is very appreciative to NQF for taking on this challenging task and for being considerate and purposeful in the selection of committee members, preparation of the materials and discussion guides, management of the Steering Committee meetings and ultimately in the writing of this report. We know this is a complex topic, and appreciate the extensive thoughtful discussion of the steering committee, and the attempt to describe the many nuances to the recommendations in the report. We believe this process has resulted in a strong initial foray into designing criteria to evaluate episode based performance measures. We, however, have a few comments we wanted to highlight (several are embedded in the attached document).

General Comments

1. In general, it is very challenging to comment on the evaluation criteria without getting a better sense of what the proposed endorsement process will be for episode groupers. For example, will there be a call for episode groupers? What is the timing of this or will they be submitted on a flow basis? Will there be a requirement to submit the specifications and systems documentation around the Episode Groupers? What are the specific documentation requirements? These questions are important because they are keys to understanding how the evaluation criteria will be applied and

2. A reader unfamiliar to the area wouldn't understand all the challenges surrounding this area based on this report. We are not sure how to address this concern; perhaps it would help to include more information that would reflect better, and more concretely the complications folks have encountered while trying to develop and/or use a grouper. A good start to addressing this concern would be to illuminate the limited nature of claims-derived data – that the data are sparse, non-standard, and often do not specify a diagnosis that is the stimulus for the treatment or service that is being billed. So, the grouping software has to make some arbitrary assumptions, or leave large amounts of claims ungrouped.

3. The report should make clear that patients, and particularly Medicare patients, do not have a single episode at a time. With multiple contemporary episodes, it's not always clear which episode a service or treatment is intended for. And some treatments may target multiple episodes simultaneously. Also, the combination of co-existing episodes may cause a physician to treat the patient in a different manner, because of the sensitivity to some complications that the co-existing episodes create. This includes triaging the most critical clinical target, choosing another modality because of the side-effects of a typical treatment, etc.

4. We suggest warning the reader about the potential consequences of uninformed or too literal interpretations of the physician profiling reports that groupers can create. In arraying physicians from most expensive, on average, per episode to least expensive, the consumer of these reports needs to recognize that physician discretion and preferred practice style is only one driver of the underlying costs differences revealed in the report. Differences in average patient risk and severity, rates of patient compliance, engagement and medical literacy, differences in patient preferences for invasive or aggressive treatment, or less costly treatments can all drive the observed differences in average episode costs. The claims data are completely devoid of any information about patient preferences and characteristics, and are only somewhat more helpful in documenting degrees of patient risk and severity. Because of this concern, it is imperative that the document be more concrete in what is meant by validity, and how, in each use case, one should test or assess validity.

5. We agree that the importance of the measure to report is not totally relevant; however, we do not agree that there should **not** be some mention of the importance of the condition-specific or procedure-specific episodes that are being reviewed.

Technical Comments

1. Overall, we recommend using an active tense (present tense) voice when writing the report and incorporating a technical writing style.

2. It is recommended that key terms are defined early on in the report so the reader has a better understanding of what is being said. Further, references should be used to support general statements and the criteria for consensus development and endorsement that are referenced throughout the report - this helps validate the report's content.

3. On p. 2 under the Introduction section, "several projects" are mention but the projects are not listed. It is recommended that the projects should either be listed or included in an appendix.

Response
Thank you for your comments. The Panel agrees with the challenges outlined in this comment. In particular, the linking of costs and quality continues to be an important issue and the Panel recognizes the potential implications of solely measuring and reporting costs in the absence of a quality signal.
Thank you for your comment. The committee appreciates the challenges to developing eye care groupers outlined by the commenter. The Panel recommends that developers take this into consideration when developing episode groupers for eye care and other conditions.

Thank you for your comment. The Panel recognizes the challenges of developing groupers as well as implementing them for payment and comparative purposes. While there are important implications for the application of groupers, the purvue of this Panel is limited to understanding the complexities of all types of groupers and the implications for evaluating and endorsing them. Further planning is required before NQF begins to evaluate epiosode
Thank you for your comment. The committee appreciates the challenges to developing eye care groupers outlined by the commenter. The Panel recommends that developers take this into consideration when developing episode groupers for eye care and other conditions.
Thank you for your comment. The Panel recognizes the complexities of condition-specific episode construction and recommends that developers of episode groupers strongly consider these comments. The Panel will continue to consider the these challenges in ongoing discussions.

Thank you for your comment. The Panel will consider this input as we seek to further clarify and explain recommendations for criteria and evaluation.
The Panel also recognizes the potential implications for influencing the market with the endorsement of an episode grouper. NQF staff and the Panel will continue to consider these challenges prior to accepting any groupers for endorsement.
Thank you for your comment. The Panel recognizes the challenges of developing groupers as well as implementing them for payment and comparative purposes. While there are important implications for the application of groupers, the purvue of this Panel is limited to understanding the complexities of all types of groupers and the implications for evaluating and endorsing them. Further planning is required before NQF begins to evaluate epiosode

The Panel acknowledges the limitations in using claims data to define clinical episodes of care and encourages the development and exploration of groupers that enable the use of clinical data. Based on the current landscape of groupers currently in use, administrative claims is currently the sole data source for defining episodes. The Panel recognizes that both public and private episode groupers have presented challenges for demonstrating statistical
The Panel agrees that the maintenance of the episode grouper is key to ensuring ongoing validity of the tool and its output.
The Panel agrees that transparency is key to the evaluation, endorsement and usability of episode groupers and is aware of the challenges with implementing and reporting results for accountability applications such as reimbursement.

Thank you for your comment. NQF staff and the Panel will consider these comments as revisions and further clarity is provided in the report. The Panel agrees that the groupers' approach to handling multiple episodes should be explicit and transparent. Also, the challenge posed by the transition to ICD-10 codes is one that is not limited to episode groupers, but applies to other administrative claims-based measures as well. The Panel recommends that developers take
The Panel agrees that procedure-focused episodes is an important approach to measuring episodes that should be addressed. Attribution is always a challenging issue particularly when dicussed in the context of costs and more work is required to explore the key principles and implications for attributing costs for specific applications.
Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.

Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
The Panel agrees that risk and severity adjustments are an important aspect of episode grouping that should be addressed. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
The Panel agrees that clinical validity, testing and transparency are key guiding principles that should be considered when evaluating groupers. The Panel will consider these comments as revisions are made to the report to provide further clarity.

Thank you for your comment. Linking costs and quality continues to be an important issue. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
The Panel also recognizes the potential implications for influencing the market with the endorsement of an episode grouper. NQF staff and the Panel will continue to consider these challenges prior to accepting any groupers for endorsement.
Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.

Thank you for your comment. The Panel recognizes that episode groupers have presented challenges for demonstrating statistical reliability and validity and that ongoing work is required to explore these issues and identify ptoential approaches to mitigate these issues. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
Thank you for your comment. Linking costs and quality continues to be an important issue. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.

Thank you for your comment. The Panel agrees that validity, testing and transparency are key guiding principles that should be considered when evaluating groupers. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
Thank you for your comment.
Thank you for your comment. The Panel agrees that risk and severity adjustments are an important aspect of episode grouping that should be addressed. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.

Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.

Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
Thank you for your comment. The Panel agrees that transparency is key to the evaluation, endorsement and usability of episode groupers. Further planning is required before NQF begins to evaluate epiosode groupers for endorsement.
Thank you for your comment. Linking costs and quality continues to be an important issue.

Thank you for your comments. The Panel agrees with the challenges outlined in this comment. In particular, the linking of costs and quality continues to be an important issue and the Panel recognizes the potential implications of solely measuring and reporting costs in the absence of a quality signal.
Thank you for your comment. The Panel recognizes that episode groupers have presented challenges for demonstrating statistical reliability and validity and that ongoing work is required to explore these issues and identify ptoential approaches to mitigate these issues. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
Thank you for your comment. NQF and the Panel will further consider relevant literature to support its recommendations.

Thank you for your comment. The challenges posed by the transition to ICD-10 codes is one that is not limited to episode groupers, but applies to other administrative claims-based measures as well. The Panel recommends that developers take into consideration these challenges as they develop episode groupers.
The Panel agrees that transparency is an important aspect of episode grouper development, evaluation and endorsement. Much of the Panel's discussion focused on the challenge of the variation in approaches to grouping episodes of care, all of which may be appropriate given the rationale for the logic. Futher exploration is required to determine if comparison to a gold standard of episode grouping is feasible.
The Panel acknowledges the limitations in using claims data to define clinical episodes of care and encourages the development and exploration of groupers that enable the use of clinical data. Based on the current landscape of groupers currently in use, administrative claims is currently the sole data source for defining episodes.

Thank you for your comment. The Panel agrees that clinical validity, testing and transparency are key guiding principles that should be considered when evaluating groupers.
Thank you for your comment. The Panel agrees that the use of clinical guidelines in the development of episode groupers is important and should be integrated where possible. However, many groupers do not solely count claims or resources used based on “appropriate” care that would be outlined in guidelines, but rather sort all eligible claims and attribute them based on their relevance to the clinical condition.
The Panel acknowledges the challenges of accounting for SES when determining the cost of an episode. Current work with NQF exploring this issue will serve as guidance for future NQF policy to address this issue. Futher consideration will need to be given to the implications those recommendations to episode grouper evaluation.

NQF's policy for all Panels and Committees requires that members disclose any potential conflicts of interest in writing and orally. Disclosures for this Panel were made orally during the in person meeting and can be found in the transcripts for this meeting.
Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.
The Panel acknowledges the limitations in using claims data to define clinical episodes of care and encourages the development and exploration of groupers that enable the use of clinical data. Based on the current landscape of groupers currently in use, administrative claims is currently the sole data source for defining episodes. The Panel also agrees that the maintenance of the episode grouper is key to ensuring ongoing validity of the tool and its

Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report to provide further clarity. Based on the recommendations from the panel, NQF staff will further refine the process for reviewing and evaluting groupers prior to accepting groupers for evaluation.
Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report to provide further clarity.
Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.

Thank you for your comment. The Panel is aware of the challenges with implementing and reporting results for accountability applications. The output should be clearly defined and future Panels should consider the grouper outputs and any untintended consequences.
Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report to provide further clarity.
Thank you for your comment. NQF staff and the Panel will consider these comments as revisions are made to the report and further clarity is provided in the report.