



TO: GI/GU Steering Committee

FR: Suzanne Theberge, MPH; Reva Winkler, MD, MPH; Ashlie Wilbon, MPH

SU: Staff Notes on Evaluation of GI/GU Measures

DA: March 25, 2013

The purpose of this memo is to prepare Committee members for the April 3 & April 8 Measure Evaluation Conference Calls.

### Memo Table of Contents

1. Action steps
2. General comments on measure evaluation
3. Suggested discussion questions
4. NQF member comments – general
5. Measure specific comments – NQF staff & NQF members

Please complete the [preliminary evaluation survey](#) online by 9:00am ET on Monday, April 1.

### ACTION STEPS

1. Review the measure evaluation forms posted [on SharePoint](#). **Please evaluate all six measures.**
2. Review the staff comments and the NQF member comments on the measures in this memo
3. Complete the [preliminary evaluation surveys](#) by **9:00am ET on Monday, April 1.**
4. For lead discussants, prepare your measure introduction (see discussion questions below)
  - **0098:** Urinary Incontinence: Assessment, Characterization, and Plan of Care for Urinary Incontinence in Women Aged 65 Years and Older – an administrative measure (NCQA): ***Alayne Markland***
  - **2065:** GI Hemorrhage Mortality Rate (IQI #18) (AHRQ): ***Liliana Bordeianou***
  - **0658:** Endoscopy/polyp surveillance: Appropriate follow-up interval for normal colonoscopy (AMA-PCPI): ***Johannes Koch / eSpecifications: Zahid Butt***
  - **0659:** Endoscopy/Poly Surveillance: Colonoscopy Interval for Patients with a History of Adenomatous Polyps- Avoidance of Inappropriate Use (AMA-PCPI): ***Phillip Schoenfeld / eSpecifications: Zahid Butt***
  - **0622:** GERD - Upper Gastrointestinal Study in Adults with Alarm Symptoms (ActiveHealth): ***John Morton***
  - **0635:** Chronic liver disease - Hepatitis A vaccination (ActiveHealth): ***Rick Luetkemeyer/ Zahid Butt***
5. For those not assigned as lead discussants, be prepared to share your thoughts and comments on the measures.
6. Attend conference calls on April 3 and April 8 to discuss and decide on recommendations for each measure. **Please be at a computer during the calls** so that you can submit your votes.

**Wednesday, April 3, 2013, 3:00 PM – 5:00 PM ET**

Phone: (888) 799-5160

Confirmation Code: 18046334

Web Link: <http://ngf.commpartners.com/se/Rd/Mt.aspx?463850>



**Monday, April 8, 2013, 3:00 PM – 5:00 PM ET**

Phone: (888) 799-5160

Confirmation Code: 18068069

Web Link: <http://ngf.commpartners.com/se/Rd/Mt.aspx?354478>

## **GENERAL GUIDANCE FROM NQF STAFF**

### *NQF Staff and Committee review of submissions*

NQF staff has performed initial reviews of the measure submission materials. If the information was found to be incomplete or non-responsive, the developer was given an opportunity to revise their submission. Committee members are NOT expected to do any additional research or literature reviews when evaluating measures. Please evaluate the measures using the information provided in the submission forms. If you know of additional information, please bring it up during the conference calls for discussion by the entire Committee.

### *Pre-evaluation member comments*

At the start of this project, comments were solicited from NQF membership on maintenance measure in use in addition to the newly submitted measures. Members may submit comments on any issues or concerns about the measures in use and/or indicate support for the measures for consideration by the Committee. The measures received 32 comments; these comments are included with the measure-specific staff notes below.

### *Importance*

While the Committee is not re-reviewing the Importance criterion, additional information on evidence and/or measure gaps for some measures was requested by the Committee in stage one; these requests were noted in the developers' checklist and are indicated below in the measure-specific notes below. If the measure passes all other criteria and is recommended by the Committee for endorsement, the Committee should review these additions and make sure they are adequate.

### *EHR specifications*

Two of the measures are submitted with eSpecifications (0658 and 0659). The criteria for evaluation of eSpecifications require a crosswalk of the EHR specifications with the original specifications for alignment and use of the Quality Data Model (QDM). NQF's HIT department has done the crosswalk review for the Committee. The results are noted in the measure specific comment tables below.

### *Level of analysis and data source(s)*

NQF endorsement applies only to the level of analysis and data source(s) that have been tested.

### *CSAC Guidance for Measure Construction*

NQF's Consensus Standards Approval Committee (CSAC) has identified preferred measure construction practices that should be considered by Committee evaluating measures:

- Avoid measures that can be met primarily through documentation without evaluation of the quality of the activity (e.g., satisfied with a checkbox, date or code) such as assessment completed; care plan created; or instruction, advice, counseling or teaching given.



- It is preferable to measure teaching/counseling from the patient perspective – i.e., intermediate outcomes of the knowledge gained or experience.
- Consider the impact of missing data. Generally, missing data should not be specified as an exclusion or implicitly limits inclusions (e.g., percent of patients with normal lab values is often specified so that the denominator only includes patients who had the test.)
- Exclusions should be supported by the evidence or supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion.
- Statistical risk models should not include factors related to disparities of care.
- Measures should be specified with the broadest applicability (target populations. Settings, levels of analysis) as supported by the evidence.
- Avoid measures where improvement decreases the denominator population (e.g., denominator – patients who received diagnostic test; numerator – patients who inappropriately received the diagnostic test. With improvement, the denominator will decrease.)

**Committee presentation and discussion of measures during Committee evaluation calls**

Lead discussants should start by stating the measure title and description. Summarize each criterion using the questions below and the results of the Committee preliminary evaluation surveys (no more than 3-5 minutes for each criterion of Reliability, Validity, Feasibility and Usability and use). After each criterion, the entire Committee will discuss the issues and then vote (instruction on voting during the call will be presented separately).

<p><b>Scientific Acceptability - Reliability:</b></p>	<ul style="list-style-type: none"> <li>• Are the measure specifications precise?</li> <li>• What is the data source(s)?</li> <li>• Are all data elements clearly defined?</li> <li>• Are eMeasures based on the QDM?</li> <li>• Are all appropriate codes included?</li> <li>• At what level was the measure tested? [data element, measure score or both?]</li> <li>• What type of reliability testing was performed?</li> <li>• Was an appropriate test population used?</li> <li>• Was testing performed using all the specified data source(s)?</li> <li>• Reliability results are typically reported on a scale of 0-1. What are the reliability test results for this measure?</li> <li>• What is your interpretation of these test results?</li> <li>• Are there any comments on the specifications or reliability of the measure?</li> </ul>
<p><b>Scientific Acceptability - Validity</b></p>	<ul style="list-style-type: none"> <li>• Are the specifications consistent with the evidence?</li> <li>• Are the exclusions supported by the evidence?</li> <li>• Are any patients or patient groups inappropriately excluded from the measure?</li> <li>• How is missing data handled?</li> <li>• At what level was the measure tested? [data element, measure score or both?]</li> </ul>



	<ul style="list-style-type: none"> <li>• What empiric testing of validity was performed?</li> <li>• Was face validity systematically assessed?</li> <li>• Is the measure stratified for disparities?</li> </ul> <p>For outcome measures:</p> <ul style="list-style-type: none"> <li>• Is an appropriate risk-adjustment strategy included in the measure?</li> <li>• Are the candidate and final variables included in the risk adjustment model adequately described for the measure to be implemented?</li> <li>• Describe the summary indicators of model fit, calibration and discrimination, and the adequacy of the model in terms of demonstrating adequate discrimination and calibration.</li> <li>• Are all of the risk adjustment variables present at the start of care? If not, describe the rationale provided.</li> <li>• Does the risk adjustment model include any factors related to disparities of care? If so, describe the rationale provided.</li> <li>• Are there any comments regarding the validity of the measure?</li> </ul>
<b>Feasibility</b>	<ul style="list-style-type: none"> <li>• Are the required data elements routinely generated and used during care delivery?</li> <li>• Are the required data elements available in electronic form, e.g., EHR or other electronic sources?</li> <li>• Is the data collection strategy ready to be put into operational use?</li> <li>• Are there any comments regarding feasibility of the measure?</li> </ul>
<b>Usability and Use</b>	<ul style="list-style-type: none"> <li>• For maintenance measures – is the measure used in at least one accountability application?</li> <li>• Is the measure publicly reported?</li> <li>• How can the performance results be used to further the goal of high-quality, efficient healthcare?</li> <li>• Do the benefits of the measure outweigh any potential unintended consequences?</li> <li>• Are there any comments regarding the use and usefulness of the measure?</li> </ul>

**GENERAL NQF MEMBER COMMENTS ON THE MEASURES**

**The following four comments were submitted as general comments:**

<b>GENERAL NQF MEMBER COMMENTS</b>
<p>We support the tri-society (ASGE/ACG/AGA) composite measure submitted the Centers for Medicare and Medicaid Services (CMS) for the 2014 Physician Quality Reporting System (PQRS) with the following elements, which have high impact relative to detecting adenomas or other colorectal cancer precursor or colorectal cancer during screening or surveillance colonoscopy:</p> <ul style="list-style-type: none"> <li>• Documentation of assessment of bowel preparation</li> </ul>



GENERAL NQF MEMBER COMMENTS	
<ul style="list-style-type: none"> <li>• Photodocumentation of completeness of colonoscopy including cecal intubation or ileocolonic anastomosis</li> </ul> <p>- Submitted by Dr. Michael P. Phelan, MD, FACEP, Cleveland Clinic</p>	
<p>We agree and strongly support the comments from the Consumer-Purchaser Disclosure Project especially on measures that are feasible and usable. We do not support measures #0098, #0622, #0635 or #0659 because they do not meet the usability and feasibility criteria. We do support measures #0658, #2065 and #2056 because these measures are both usable and feasible especially #2056 the Colonoscopy Quality Index. #2056 fills a need for a meaningful and useful indicator of whether the care was necessary and valuable. Consumers can use their purchasing power to indicate quality and value.</p> <p>- Submitted by Ms. Louise Y. Probst, MBA, RN, St. Louis Area Business Health Coalition</p>	
<p>We would strongly support the comments from the Consumer-Purchaser Disclosure Project, particularly around advancing only measures which are both feasible and useful. The proliferation of measures is not helpful for consumers and is clearly burdensome for providers. We would agree that #0098, #0622, #0635 and #0659 are probably not defensible if we apply the criteria of both feasible and USEFUL. We are also happy to see #2056 the Colonoscopy Quality Index back on the list as we think this is the type of measure most easily understandable and therefore most useful to consumers.</p> <p>-Submitted by Ms. Mary Lehman MacDonald, America's Health Insurance Plans</p>	
<p>We recommend integrating this measure set into registries in order to facilitate better data collection and longitudinal use across care settings for a larger population of patients.</p> <p>- Submitted by Ms. Carmella Bocchino, MBA, RN, American Federation of Teachers Healthcare</p>	

**MEASURE SPECIFIC COMMENTS: NQF MEMBERS AND STAFF**

In the following tables, NQF staff is providing assistance to the Committee in evaluating the measures, as well as including implementation comments for maintenance measures and general comments from our members:

<a href="#">0098 Urinary Incontinence: Assessment, Characterization, and Plan of Care for Urinary Incontinence in Women Aged 65 Years and Older – an administrative measure (NCQA)</a>	
Level of analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team Data source: Administrative claims, Paper Medical Records	
<b>Importance</b>	N/A
<b>Scientific Acceptability</b>	<ul style="list-style-type: none"> <li>• Reliability – tested at measure score only.</li> <li>• Validity – face validity only.</li> <li>• This measure has three numerators and two denominators</li> <li>• 2a1.25 says data source admin claims, but 1.1 testing says measure specified to use data from &amp; tested with data from abstracted paper record and abstracted EHR. Which is correct?</li> </ul>
<b>Usability</b>	<ul style="list-style-type: none"> <li>• When will PQRS data be publicly reported?</li> <li>• Do the multiple rates enhance usability?</li> <li>• Why is the measure titled “an administrative measure” when the data source that was tested is abstraction from medical records (EHR or paper)?</li> </ul>



<b>Feasibility</b>	Data requires chart abstraction.
<b>Competing Measures</b>	N/A
<b>NQF MEMBER COMMENTS</b>	
<i>Submitted by Dr. Amir Qaseem, MD, PhD, MHA, FACP, American College of Physicians</i>	The Performance Measurement Committee (PMC) of the American College of Physicians appreciates the opportunity to comment on the NQF Gastrointestinal & Genitourinary Measure Endorsement Project. The PMC is concerned that the threshold of “any” urinary incontinence is too strict, especially in women > 65. In addition, the PMC has concerns that it will be difficult to measure assessment, characterization, and a plan of care in place through administrative data alone in a paper based system. To collect this type of data it would most likely require a chart review or supplemental codes which would be administratively burdensome.
<i>Submitted by Dr. Michael P. Phelan, MD, FACEP, Cleveland Clinic</i>	Our organization agrees that this is clinically important topic but because the presence of incontinence is typically not documented in a discrete field, currently it would require significant manual chart review to identify, which is not practical or feasible. If an emeasure could be developed surrounding this measure with appropriate discrete fields it may be more feasible.
<i>Submitted by Dr. Matt Austin, PhD, Armstrong Institute for Patient Safety and Quality at Johns Hopkins University</i>	Our organization recommends changing the structure of this measure. It is unclear to us if the proposal is a single measure with three rates or three unique measures. If it is designed to be a single measure, it is unclear to us how the three rates would be used together to assess provider performance. A clarification on this point would be helpful. We would recommend breaking this proposal into two measures. The first measure would look at the adherence of assessing women for UI (rate A). The second measure would look at the adherence of characterizing AND creating care plans for women who have been diagnosed with UI (rates B and C). This proposed breakdown would align the denominator populations.
<i>Submitted by Dr. Carol Sakala, MSPH, PhD, Childbirth Connection</i>	Childbirth Connection does not support 0098 as it is a standard of care and check-the-box measure. Further, it is only collected through the Health Outcomes Survey and thus does not measure care for traditional Medicare beneficiaries.
<i>Submitted by Dr. David Hopkins, MS, PhD, Consumer-Purchaser Disclosure Project</i>	Do NOT Support This is a standard of care, check-the-box measure and only applies to the 65 years and older population. The measure developer provides an adequate explanation for the age limitation the data is collected through the Health Outcomes Survey that is only administered to Medicare Advantage beneficiaries but it does not obviate the problem and therefore the question of usability remains.



<b><u>2065 Gastrointestinal Hemorrhage Mortality Rate (AHRQ)</u></b>	
Level of analysis: Facility	
Data source: Administrative claims	
<b>Importance</b>	N/A
<b>Scientific Acceptability</b>	<ul style="list-style-type: none"> <li>• See additional testing review information posted on SharePoint in the 2065 measure folder</li> <li>• For more detailed information on the APR-DRG codes and descriptions, please go to <a href="http://www.aprdrgassign.com">www.aprdrgassign.com</a> and log in with UserID: NQFUser</li> <li>• Potential concerns around small numbers</li> </ul>
<b>Usability</b>	Potential concerns around small numbers
<b>Feasibility</b>	Potential concerns around small numbers
<b>Competing Measures</b>	N/A
<b>NQF MEMBER COMMENTS</b>	
Submitted by Dr. David Hopkins, MS, PhD, Consumer-Purchaser Disclosure Project	DO Support This measures a health outcome that is important to any patient with a principal diagnosis of gastrointestinal hemorrhage. In Stage 1, we agreed with the Committees recommendation that the developer capture information on patients with GI bleed who do not have it as a primary diagnosis. We are satisfied with the developers' response to this recommendation and are glad to see ongoing analysis of whether patients with secondary diagnosis may be included as well. Overall, this measure is both usable and feasible and we continue to support it in Stage 2.
Submitted by Dr. Carol Sakala, MSPH, PhD, Childbirth Connection	2065 is of special value as a health outcome measure that is important to patients with a primary diagnosis of gastrointestinal hemorrhage. We appreciate that the developer has responded to a request to capture information on patients with a GI bleed who do not have a primary diagnosis of gastrointestinal hemorrhage. Further, we appreciate the ongoing attention to whether patients with a secondary diagnosis of GI hemorrhage benefit from this measure. As structured, the measure is both feasible and usable, and we support it.

<b><u>0658 Endoscopy/Polyp Surveillance: Appropriate follow-up interval for normal colonoscopy in average risk patients (AMA PCPI)</u></b>	
Level of analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team	
Data source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Imaging/Diagnostic Study, Electronic Clinical Data : Registry	
<b>Importance</b>	NA
<b>Scientific Acceptability</b>	<ul style="list-style-type: none"> <li>• Reliability – tested at measure score only.</li> <li>• Validity – face validity only.</li> <li>• Has there been any comparison of the EHR measure results to the registry measure?</li> </ul> <p><b>eSpec review comments:</b>  <b>Issues identified:</b>            Inconsistency between exclusions and exceptions: the measure form lists exclusions, but the e-specs list it as exceptions. Can you clarify?</p>





	<p><b>AMA-PCPI Response:</b> Since the NQF measure form does not distinguish between exclusions and exceptions, we have listed our exceptions in field 2a1.8. Denominator Exclusions. In 2a1.9. Denominator Exclusion Details, you will note that we refer to “measure exception categories” and not to exclusions. By definition, exclusions are absolute and are applied to all patients, whereas exceptions would only apply if the patient does not meet the numerator of the measure. Please note that exclusions and exceptions are differentiated in our logic flow diagram included with the eSpecification.</p>
<b>Usability</b>	N/A
<b>Feasibility</b>	<ul style="list-style-type: none"> <li>• Both registry and EHR specifications are submitted. Are the results comparable?</li> <li>• Has any analysis been done on missing data?</li> </ul>
<b>Competing Measures</b>	N/A
<b>NQF MEMBER COMMENTS</b>	
Submitted by Dr. Amir Qaseem, MD, PhD, MHA, FACP, American College of Physicians	<p>The Performance Measurement Committee (PMC) of the American College of Physicians appreciates the opportunity to comment on the NQF Gastrointestinal &amp; Genitourinary Measure Endorsement Project. The PMC does not support this measure without the addition of an upper age limit of 75. A recent Guidance Statement published by the American College of Physicians recommends that clinicians stop screening for colorectal cancer in adults over the age of 75 years or in adults with a life expectancy of less than 10 years. (Citation: Amir Qaseem, Thomas D. Denberg, Robert H. Hopkins, Jr., Linda L. Humphrey, Joel Levine, Donna E. Sweet, Paul Shekelle, ; Screening for Colorectal Cancer: A Guidance Statement From the American College of Physicians. Annals of Internal Medicine. 2012 <a href="http://annals.org/article.aspx?articleid=1090701">http://annals.org/article.aspx?articleid=1090701</a>)</p>
<b>IMPLEMENTATION</b> Submitted by Ms. Louise Y. Probst, MBA, RN, St. Louis Area Business Health Coalition	<p>This measure is both usable and feasible and will improve safety and quality of care. We strongly agree with the comments from the Consumer-Purchaser Disclosure Project.</p>
Submitted by Dr. David Hopkins, MS, PhD, Consumer-Purchaser Disclosure Project	<p>DO Support We applaud the measure developer for providing two measures aimed at reducing unnecessary colonoscopies (this and measure 0659). This measure is both usable and feasible and will improve safety and quality of care.</p>
Submitted by Dr. Carol Sakala, MSPH, PhD, Childbirth Connection	<p>DO Support We applaud the measure developer for providing two measures aimed at reducing unnecessary colonoscopies (this and measure 0659). This measure is both usable and feasible and will improve safety and quality of care.</p>
Submitted by Dr. Michael P. Phelan, MD, FACEP, Cleveland Clinic	<p>This measure seeks to limit overuse of colonoscopy as a screening tool. Age limit is reasonable, and the exceptions appear appropriate. Reporting of an exception rate is also a reasonable component to this measure. Additionally, data provided demonstrates variation in data and a gap that can be addressed. At this time, we support this measure.</p>





<p><b><u>0659 Endoscopy/Polyp Surveillance: Colonoscopy Interval for Patients with a History of Adenomatous Polyps- Avoidance of Inappropriate Use (AMA-PCPI)</u></b></p> <p>Level of analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team            Data source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Imaging/Diagnostic Study, Electronic Clinical Data : Registry</p>	
<b>Importance</b>	Committee requested additional evidence in the checklist. Was this updated to your satisfaction?
<b>Scientific Acceptability</b>	<ul style="list-style-type: none"> <li>Reliability – tested at measure score only.</li> <li>Validity – face validity only.</li> <li>Has there been any comparison between the paper record and the eMeasure specifications of this measure?</li> </ul> <p><b>eSpec review comments</b>  <b>Issues identified:</b>            Inconsistency between exclusions and exceptions: the measure form lists exclusions, but the e-specs list it as exceptions. Can you clarify?  <b>AMA-PCPI Response:</b> Since the NQF measure form does not distinguish between exclusions and exceptions, we have listed our exceptions in field 2a1.8. Denominator Exclusions. In 2a1.9. Denominator Exclusion Details, you will note that we refer to “measure exception categories” and not to exclusions. Please note that exclusions and exceptions are differentiated in our measure logic included with the eMeasure.</p> <p><b>Logic Checks:</b>            Issues identified: It was found during the Meaningful Use Stage 2 clinical quality measure work that the use of certain operators and functions may prove difficult for CQM implementation. CURRENT and IMMEDIATE PRIOR were identified as possible confusion points as there definitions are not clear.            Suggest to the measure developer to use the ‘specific occurrence’ feature with the procedure datatype to specify the measure.  <b>AMA-PCPI Response:</b> We appreciate the feedback from NQF on the use of the logic operators “CURRENT” and “IMMEDIATE PRIOR”. At the time that the eMeasure was developed for NQF 0659, it was in coordination with NQF HIT staff, prior to the December, 2011 deliverable to HHS and, well in advance of the feedback received from the Certification team (ie, MITRE) on the use of these two logic operators. We will consider incorporating the use of the “specific occurrence” feature to revise the eMeasure at a later date.</p>
<b>Usability</b>	Are registry and EHR measure results comparable?
<b>Feasibility</b>	Has any analysis been done on missing data?
<b>Competing Measures</b>	N/A
<b>NQF MEMBER COMMENTS</b>	
Submitted by Dr. Amir Qaseem, MD, PhD, MHA, FACP, American	The Performance Measurement Committee (PMC) of the American College of Physicians appreciates the opportunity to comment on the NQF Gastrointestinal & Genitourinary Measure Endorsement Project. The PMC is



College of Physicians	concerned that the evidence for colorectal cancer screening at 3 year intervals is limited. According to the ACG cited, the recommended follow-up time can vary from 3-5 years depending on the number, type of adenomas, and grad of dysplasia. Requiring every patient to receive screening at 3 year intervals could induce overuse of screening tests.
Submitted by Dr. David Hopkins, MS, PhD, Consumer-Purchaser Disclosure Project	Do NOT Support This measure, as specified, suffers from two major failings: the exclusions provide a large loophole for providers to manipulate the results, and the lack of information about previous colonoscopies may hide evidence of poor care. The first failing could be addressed by using the list of specific exclusions to define medical reasons rather than citing them as examples and leaving it open for providers to declare that they had medical reasons to contradict evidence-based guidelines. As for the second, we cannot accept that a patient would be asked to undergo an unnecessary procedure because the provider is unable to track down their prior medical records.
Submitted by Dr. Carol Sakala, MSPH, PhD, Childbirth Connection	Childbirth Connection identifies two shortcomings of 0659. First, the exclusions of this measure enable providers to game the results. This could be rectified if the list of exclusions were changed into specific medical reasons, and providers were unable to merely claim a medical reason to deviate from best evidence. Second, lack of information about prior colonoscopies may make it difficult to ascertain the quality of care. It is unacceptable in 2013 to ask a patient to undergo and a payer to pay for such a procedure simply because medical records do not reveal past screenings.
Submitted by Dr. Matt Austin, PhD, Armstrong Institute for Patient Safety and Quality at Johns Hopkins University	Our organization seeks to understand the rationale for the interval of 3 or more years. The evidence provided in the proposal suggests an interval of 5-10 years. A minimum and maximum interval range could be useful for this clinical process measure.
Submitted by Dr. Michael P. Phelan, MD, FACEP, Cleveland Clinic	This measure seeks to limit overuse of colonoscopy as a surveillance tool. While there are mild variations in recommendation interval (3 versus 5 years), this issue has been adequately addressed by the developer. As above, data provided demonstrates variation in data and a gap that can be addressed. At this time, we support this measure.

<b>0622 GERD – Upper Gastrointestinal Study in Adults with Alarm Symptoms (ActiveHealth)</b>	
Level of analysis: Population : National	
Data source: Other	
<b>Importance</b>	Committee requested additional evidence and performance gap information in the checklist. CSAC was particularly interested in performance gap data. Was this updated to your satisfaction?
<b>Scientific Acceptability</b>	<ul style="list-style-type: none"> <li>• Are specifications aligned with the evidence?</li> <li>• Reliability – more information is need on the type of testing and explanation of the result “SNR=5”. Typically reliability testing using signal</li> </ul>



	<p>to noise ratio is calculated as “signal/ (signal + noise)” which generates values from 0-1.</p> <ul style="list-style-type: none"> <li>• Testing is not complete – testing submitted is for the measure before they implemented the Committee’s stage 1 checklist recommendations.</li> </ul>
<b>Usability</b>	<ul style="list-style-type: none"> <li>• The form notes that ActiveHealth does not plan to ever publically report measures – measures are in Aetna’s reporting system but results are not publically reported.</li> <li>• This measure is population level only. What does that mean for use as a performance measure?</li> </ul>
<b>Feasibility</b>	Given the data sources, can any other organization use or report on this measure?
<b>Competing Measures</b>	N/A
<b>NQF Member Comments</b>	
Submitted by Dr. Amir Qaseem, MD, PhD, MHA, FACP, American College of Physicians	<p>The Performance Measurement Committee (PMC) of the American College of Physicians appreciates the opportunity to comment on the NQF Gastrointestinal &amp; Genitourinary Measure Endorsement Project. There is currently a lack of evidence that a substantial quality gap exists. Therefore, this measure may create an unjustified measurement burden and will not improve quality of care. The measure specifications do not align with the clinical evidence presented in ACP latest clinical guideline (<a href="http://annals.org/article.aspx?articleid=147028">http://annals.org/article.aspx?articleid=147028</a>), which recommends the use of upper endoscopy in men and women with heartburn and alarm symptoms (dysphagia, bleeding, anemia, weight loss, and recurrent vomiting). The alarm symptoms (bleeding and recurrent vomiting) are not currently included in measure specifications and should be added to the denominator of the measure. The term “gastrointestinal study” in the numerator of the measure should be defined. For example, a barium study for diagnosis of GERD is not an evidence based standard of care. In addition, this measure should not be used without pairing with an overuse measure.</p> <p>There is a need to develop of an overuse measure to assess potential overuse of upper endoscopy in patients with GERD without alarm symptoms. Evidence shows that the use of upper endoscopy for GERD indications is rising, suggesting possible inappropriate and unnecessary use in patients with GERD without alarm symptoms</p>
Submitted by Dr. Matt Austin, PhD, Armstrong Institute for Patient Safety and Quality at Johns Hopkins University	<p>Our organization recommends initially using just the second denominator population for this measure: high risk patients (i.e., obese, male, or age 50) with a diagnosis of GERD with alarm symptoms (i.e., dysphagia or weight loss) in the past 12 months. Using one denominator simplifies the measure. As a first step, providers could focus their efforts on improving their rates in the high-risk patient population. Once they have reached a defined level of attainment in this subpopulation, the denominator can be expanded to all patients diagnosed with chronic GERD and who exhibited alarm symptoms in the past 12 months.</p>
Submitted by Dr. Carol Sakala, MSPH, PhD, Childbirth Connection	<p>As 0622 is a process measure that is the standard of care, Childbirth Connection does not support it. While it would be feasible, we question its usability and do not believe that the burden of collection is justified in this</p>



	case.
Submitted by Dr. David Hopkins, MS, PhD, Consumer-Purchaser Disclosure Project	Do NOT Support This is a measure of a process that is a standard of care. Therefore, although the measure is feasible, we question the usability and whether this data collection exercise is a valuable use of time.

<b>0635 Chronic Liver Disease - Hepatitis A Vaccination (ActiveHealth)</b>	
Level of analysis: Population : National	
Data source: Other	
<b>Importance</b>	N/A
<b>Scientific Acceptability</b>	Reliability – more information is need on the type of testing and explanation of the result “SNR=6”. Typically reliability testing using signal to noise ratio is calculated as “signal/ (signal + noise)” which generates values from 0-1.
<b>Usability</b>	<ul style="list-style-type: none"> <li>Form notes that ActiveHealth does not plan to ever publically report measures – measures are in Aetna’s reporting system but results are not publically reported</li> <li>This measure is population level only. What does that mean for use as a performance measure?</li> </ul>
<b>Feasibility</b>	Given the data sources, can any other organization use or report on this measure?
<b>Competing Measures</b>	Awaiting an update on harmonization with endorsed measure 0399, a related measure.
<b>NQF Member Comments</b>	
Submitted by Dr. David Hopkins, MS, PhD, Consumer-Purchaser Disclosure Project	Do NOT Support This is a measure of a process that is standard of care. We question whether and how results from this measure will improve health care and outcomes. Therefore, although it is certainly feasible, it is not particularly useful.
Submitted by Dr. Carol Sakala, MSPH, PhD, Childbirth Connection	As 0635 is a process measure and the standard of care, Childbirth Connection does not support it. Its potential contribution to improving health care and outcomes is extremely limited. It would be feasible to collect this measure, but its usability is in question.