Infectious Disease Endorsement Maintenance
Preliminary Workgroup Summary Evaluations

0058: Avoidance of antibiotic treatment in adults with acute bronchitis

Status: Maintenance, Original Endorsement: Aug 10, 2009

Description: The percentage of adults 18–64 years of age with a diagnosis of acute bronchitis who were not dispensed an antibiotic prescription.

Numerator Statement: Patients who were dispensed antibiotic medication (see Table 1) on or three days after an outpatient or ED encounter for acute bronchitis (a higher rate is better). The measure is reported as an inverted rate (i.e. 1- numerator/denominator) to reflect the number of people that were not dispensed an antibiotic.

Denominator Statement: All patients 18 years as of January 1 of the year prior to the measurement year to 64 years as of December 31 of the measurement year with a claim/encounter for a diagnosis of acute bronchitis (refer to Table 2) and an outpatient or ED visit code (refer to Table 3) during the Intake Period (January 1–December 24 of the measurement year).

Exclusions: N/A

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Health Plan, Integrated Delivery System

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Pharmacy

Measure Steward: National Committee for Quality Assurance

0058: Avoidance of antibiotic treatment in adults with acute bronchitis

IMPLEMENTATION COMMENTS:

- IDSA: The Quality Improvement Task Force of the Infectious Disease Society of America (IDSA) continues to support Measure #0058, Avoidance of antibiotic treatment in adults with acute bronchitis and Measure #0069, Appropriate treatment for children with upper respiratory infection (URI). Studies have shown that acute bronchitis and URI are virtually always of viral etiology, yet clinicians continue to prescribe antibiotics inappropriately for those conditions. The measures include only those patients without claims/encounters for a diagnosis of a comorbid condition for the prior 12 months, and without competing diagnoses or new medications in the prior thirty days. We would, however, like to draw attention to the fact that classification of URIs can be a very subjective process. That is, one physician might opt for a diagnosis of acute bronchitis, while another chooses "common cold." As a measure is used, there might be a shift in those subjective choices, either to avoid the evaluation or to choose a condition in which an antibiotic can at times be appropriate, such as acute sinusitis. It would be useful to consider a corollary to this measure that would look at all URI visits for the physician or practice during the evaluation time period to identify any shifts in coding of URIs. Although we submit our comments for consideration to improve the measure, we support endorsement of both measures for an additional 3 years.
  - o Developer response: Thank you for your support. We will bring your new measure suggestions to our measurement advisory panel for consideration.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the **Committee Reviewers:** Tiffany Osborn; Rekha Murthy; Curtis Collins; Mary Blank; Mohamad Fakih; Thomas File   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-5; N-1

1a. Impact: H-6; M-0; L-0; I-0; 1b. Performance Gap: H-5; M-0; L-1; I-0

Rationale:

- 1a. Very common diagnosis in URI and often inappropriate antibiotics given
- 1a. Very common diagnosis
- 1b. Still significant % patients receive unnecessary antimicrobials, although some reduction from 2009 (mean of 25.58 in 2009 to 22.03 in 2011 for 'commercial' patients)
- 1b. Data show a majority receive antibiotics inappropriately

1c. Evidence (*based on decision logic*): Y-6; N-0   IF a **Health Outcome**, rationale supports: Y-2; N-0; NA-4

Quantity: H-3; M-3; L-0; I-0; Quality: H-2; M-4; L-0; I-0; Consistency: H-3; M-3; L-0; I-0

0058: Avoidance of antibiotic treatment in adults with acute bronchitis

Rationale:
- The main issue with this measure is that it addresses one code 466. Shifting diagnosis to another code would miss many inappropriate Rx for acute bronchitis with antibiotics (Roth S, Am J Manag Care. 2012 Jun 1;18(6):e217-24).
- Note that developer gives an example of gap in 1b2 with no improvement with time. I am not clear how much this measure had an effect on improving antibiotic use for acute bronchitis.
- More recent Cochrane systematic review, 2012, with review of 15 trials, 2618 patients; limited evidence for marginal benefit of antimicrobials. "However, the magnitude of this benefit nees to be considered in the broader context of potential side effects, medicalisation for a self limiting condition, increased resistance to respiratory pathogens and cost of antibiotic treatment." ".....update provides clearer evidence on the lack of effectiveness of antibiotics for acute bronchits." From Up-To-Date (File T, author): • "We recommend NOT treating patients with presumed acute bronchitis with empiric antibiotic therapy (Grade 1A)"

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-6; N-0
2a. Reliability: H-3; M-3; L-0; I-0, 2b. Validity: H-4; M-1; L-1; I-0
Rationale:
- Consistent classification of acute bronchitis/URI would be challenging as it can be subjective.
- 2a. It is reliable, test is replicated.
- 2b. Validity is of great concern because of potential shift in diagnosis because it reflects one billing code, a simple change to bronchitis not specified will miss the cases (Am J Manag Care. 2012 Jun 1;18(6):e217-24). A change in coding of diagnosis may lead to significantly different results.
- Concern for "intended/unintended" consequence of this measure: A recent review using data from a large, integrated health system examined trends in antibiotic use for acute bronchitis from 2006 to 2009 and observed an unintended consequence of this measure. While there was a significant reduction of patients treated with antibiotics for diagnosis code 466.0 (acute bronchitis) there was a significant increase in the use of diagnosis code 490 (bronchitis, not otherwise specified) associated with antibiotic use. As a result, the odds of an antibiotic prescription for codes 466 and 490 combined decreased only slightly and suggested the measure influenced a change in diagnosis coding as an unintended consequence which resulted in continuing antibiotic use. [D. Roth S, Gonzales R, Harding-Anderer T, et al. Unintended Consequences of a Quality Measure for Acute Bronchitis. Am J Manag Care. 2012;18(6):e217-e224]

3. Usability: H-4; M-1; L-1; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:
- 3a. Used for public reporting
- 3b. Results are meaningful unless coding changes in the physician practices involved
- Conforms with CDC recommendations

4. Feasibility: H-2; M-4; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:
- 4a. Yes, available through EHR
- 4b. Yes EHR billing, and pharmacy databases
- 4c. Issue with coding may exclude a large number if there is a diagnosis shift
- 4d. Yes.
- See comment for validity above regarding "unintended consequences"

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-5; N-0
Rationale:
- Suitable for endorsement if the diagnosis code 466 is addressed. May consider adding other diagnosis that may be used by physicians interchangeably (eg 490 or other upper respiratory infections)
- Antimicrobial overuse remains high for this diagnosis fro which approx 90% of infections are viral in etiology and antimicrobials unwarranted. This leads to increase resistance and possible unnecessary adverse events (e.g, CDI) and cost.

Additional Comments/Questions:
- Any marginal benefit found on meta-analyses does not overrule downside

Workgroup Discussion

Importance to Measure and Report
- Antibiotics are overused for bronchitis and unnecessary antimicrobials are prescribed to patients. This measure encourages

0058: Avoidance of antibiotic treatment in adults with acute bronchitis

providers to not prescribe antibiotics.

- The performance gap data indicates the percentage of patients who had acute bronchitis but was not prescribed an antibiotic is quite low –22-25 percent. The 2011 data show that mean performance rate has declined since 2009 (2011: 22.03; 2009: 25.48 for Commercial) (2011: 23.57; 2009: 25.76 for Medicaid).

Scientific Acceptability of Measure Properties

- This is a health plan level measure.
- It was asked if the measure captures delayed prescriptions for patients with symptoms of bronchitis who were prescribed an antibiotic a week or so after having phone contact with their physician.
  - o The developer specified that they are not able to capture the encounter in claims data but may be able to do so in EHR. The EHR will be more flexible in the choices for dispensing medication.

Usability

- This measure is publicly reported through HEDIS, a part of PQRS and under consideration for Stage 2 meaningful use program. The e-measure will be available once Stage 2 meaningful use is published.

Feasibility

- The WG expressed concerns of unintended consequences due to a decline in using diagnosis code 466 but an increase in using diagnosis code 490, which indicates that patients are continuing to be prescribed antibiotics.
  - o The developer stated that the auditors look for notable swifts in diagnosis codes amongst plans to ensure plans are not 'gaming' the system.

0069: Appropriate treatment for children with upper respiratory infection (URI)

Status: Maintenance, Original Endorsement: Aug 10, 2009
Description: Percentage of children 3 months to 18 years of age with a diagnosis of URI who were not dispensed an antibiotic medication.
Numerator Statement: Patients who were dispensed antibiotic medication (Table 1) on or within 3 days after an outpatient or ED encounter for upper respiratory infection (URI) (a higher rate is better). The measure is reported as an inverted rate (i.e. 1- numerator/denominator) to reflect the number of children that were not dispensed an antibiotic.
Denominator Statement: All children age 3 months as of July 1 of the year prior to the measurement year to 18 years as of June 30 of the measurement year who had an ED or outpatient visit with only a diagnosis of nonspecific upper respiratory infection (URI) (Table 2) during the intake period (July 1st of the year prior to the measurement year to June 30th of the measurement year).
Exclusions: N/A
Adjustment/Stratification: No risk adjustment or risk stratification N/A N/A
Level of Analysis: Health Plan, Integrated Delivery System
Type of Measure: Process
Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Pharmacy
Measure Steward: National Committee for Quality Assurance

0069: Appropriate treatment for children with upper respiratory infection (URI)

IMPLEMENTATION COMMENTS:
- IDSA: Same comments as measure 0058.
  - Developer response: See above [response for measure 0058].

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Tiffany Osborn; Rekha Murthy; Curtis Collins; Mary Blank; Thomas File; Mohamad Fakih   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-5; N-1
1a. Impact: H-6; M-0; L-0; I-0; 1b. Performance Gap: H-4; M-1; L-1; I-0
Rationale:
- Common cause of ambulatory visits; If 84.49% not receiving ABX, not alot of opportunity for improvement?
- 1a. significant national problem
- 1b. less than optimal performance

1c. Evidence (*based on decision logic*): Y-6; N-0   IF a **Health Outcome**, rationale supports: Y-1; N-0; NA-5
Quantity: H-2; M-4; L-0; I-0 **Quality:** H-2; M-4; L-0; I-0 **Consistency:** H-3; M-3; L-0; I-0
Rationale:
- Most of evidence from studies of adults; In the Cochrance review, 4 studies included children
- Studies show inappropriate antibiotic use. Multiple risks for adverse events with inappropriate antibiotic use and increased risk for resistance.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-6; N-0
2a. Reliability: H-5; M-1; L-0; I-0; **2b. Validity:** H-3; M-3; L-0; I-0
Rationale:
- Diagnosis is subjective
- What if telephone interaction within 3 days, then visit for worsening 465 or "delayed Prescription"?
- 2a. The measure is reliable, can be extracted from EHR
- 2b. Face validity tested by a panel of experts

3. Usability: H-4; M-2; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:
- Useful for public reporting and quality improvement (reduction of inappropriate antibiotics. Results are useful for informing quality improvement

0069: Appropriate treatment for children with upper respiratory infection (URI)

4. Feasibility: <mark>H-3; M-3; L-0; I-0</mark>
(4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)
<u>Rationale</u>:
- ? of prior telephone encounters or 'delayed prescriptions'
- Data elements obtained from coding, pharmacy data and EMR.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-5; N-0</mark>
<u>Rationale</u>:
- Helps evaluating appropriate antibiotic use for a condition that rarely requires antibiotics. Its weakness is its dependency on coding.

<u>Additional Comments/Questions</u>:
- ? of magnitude of improvement based on performance gap, but important to continue to monitor since such a high volume condition.
- Note that the samples reported show no significant changes in practice or antibiotic use for the last 5 years (antibiotics used in about 15% as a median)

<mark>**Workgroup Discussion**</mark>

Importance to Measure and Report
- Antibiotics are overused for respiratory infections and unnecessary antimicrobials are prescribed to patients. This measure encourages providers to not prescribe antibiotics.
- Lots of evidence on the development of antibiotic resistance and overuse in URIs.
- Small improvement in performance rate (2011: 84.49; 2009: 83.61 for Commercial) (2011: 87.18; 2009: 85.49 for Medicaid); however, one WG member stated that even if the performance rate continues to improve this measure should not be retired due to the importance of the measure focus.
- Better performance of this overuse measure compared to 0058 – possible reasons include pediatricians are more selective in prescribing and there is a longer list of exclusions for this measure compared to 0058.
- A common practice is to give a patient a prescription to be filled only if they are not better in 3-5 days but may fill immediately.

Scientific Acceptability of Measure Properties
- This is a health plan level measure.
- It was asked if the measure captures delayed prescriptions for patients with symptoms of bronchitis who were prescribed an antibiotic a week or so after having phone contact with their physician.
  - The developer specified that they are not able to capture the encounter in claims data but may be able to do so in EHR. The EHR will be more flexible in the choices for dispensing medication.
  - What about low cost drugs from discount pharmacies?
    - Developer reported that these prescriptions are variably captured in the measure depending on whether the discount pharmacy shares the data.

Usability
- This measure is publicly reported through HEDIS, a part of PQRS and under consideration for Stage 2 meaningful use program. The e-measure will be available once Stage 2 meaningful use is published.

0500: Severe sepsis and septic shock: Management bundle
Status: Maintenance, Original Endorsement: Oct 24, 2008
Description: This measure will focus on patients aged 18 years and older who present with symptoms of severe sepsis or septic shock. These patients will be eligible for the 3 hour (severe sepsis) and/or 6 hour (septic shock) early management bundle.
Numerator Statement: Number of patients who meet criteria for severe sepsis and septic shock and sucessfully receive the following early management bundle as indicated.
WITHIN THREE HOURS OF SEVERE SEPSIS:
1) Measure lactate level
2) Obtain blood cultures prior to antibiotics
3) Administer broad spectrum antibiotics
4) Administer 30ml/kg crystalloid for hypotension or lactate ≥4mmol/L
WITHIN 6 HOURS OF INITIAL SYMPTOMS FOR SEPTIC SHOCK:
5) Apply vasopressors (for hypotension that does not respond to initial fluid resuscitation to maintain a mean arterial pressure ≥65mmHg)
6) In the event of persistent arterial hypotension despite volume resuscitation (septic shock) or initial lactate ≥4 mmol/L (36 mg/dl):
- Measure central venous pressure (CVP)
- Measure central venous oxygen saturation (ScvO2)
7)Remeasure lactate
Denominator Statement: Number of patients diagnosed or presenting with the symptoms of severe sepsis or septic shock.
Exclusions: Patients with advanced directives for comfort care or clinical conditions that preclude total measure completion should be excluded. Examples include but are not limited to mortality within the numerator time window (3 hrs for severe sepsis or 6 hrs for septic shock), patients who do not have the clinical evidence of an infection (severe sepsis or septic shock), patients for whom a central line is contraindicated, patients with coagulopathy, patients for whom central line placement was attempted but could not be inserted, or other medical, patient, or system reasons for exclusion.
Adjustment/Stratification: No risk adjustment or risk stratification None Henry Ford Hospital (HFH) encourages the results of this measure to be stratified by race, ethnicity, gender, and primary language, illness severity and have included these variables as recommended data elements to be collected.
Level of Analysis: Facility, Integrated Delivery System
Type of Measure: Composite
Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records, Electronic Clinical Data : Registry
Measure Steward: Henry Ford Hospital **Other organizations:** Henry Ford Hospital System(HFHS)
Society of Critical Care Medicine (SCCM)
Infectious Diseases Society of America (IDSA)
Institute for Healthcare Improvement (IHI)
Surviving Sepsis Campaign (SSC)
0500: Severe sepsis and septic shock: Management bundle
IMPLEMENTATION COMMENTS:
• IDSA: The IDSA supports endorsement of measure 0500. To be more consistent with the Surviving Sepsis Campaign guidelines, these items are referred to as part of a resuscitation rather than management bundle. We look forward to supporting future efforts to update this measure once updated guidelines are made public.
o Developer response: We would like to thank IDSA for their support of the measure, and look forward to collaborating with all stakeholders throughout the measure endorsement maintenance process.
Notes
Workgroup Preliminary Evaluations
The following evaluation ratings and comments are from the Committee Reviewers: Rekha Murthy; Tiffany Osborn; Mary Blank; Thomas File; Mohamad Fakih; Curtis Collins   *(comments separated by bullets)*

0500: Severe sepsis and septic shock: Management bundle

Importance to Measure and Report (*based on decision logic*): Y-5; N-0
1a. Impact: H-5; M-0; L-0; I-0; **1b. Performance Gap:** H-5; M-0; L-0; I-0
Rationale:

- High numbers of patients; data submitted showing some of the interventions have performance rates as low as 15% (ScvO2)
- 1a. Large number affected with poor outcomes
- 1b. Compliance with processes significantly varies

1c. Evidence (*based on decision logic*): Y-6; N-0   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-5
Quantity: H-5; M-1; L-0; I-0; Quality: H-4; M-2; L-0; I-0; Consistency: H-5; M-1; L-0; I-0
Rationale:

- Regarding total body of evidence, majority of the evaluations are bundle completion vs non-completion studies or observational studies. All either support a mortality benefit of EGDT or show no difference. None have shown increased mortality. Sufficient concern regarding CVP and ScvO2 exists that three government funded RCTs are currently on going.
- Multiple studies cited
- Studies and guidelines support the measures. Many aspects of the measure are supported by studies (eg, early goal directed therapy, early initial broad spectrum antibiotics).

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-5; N-1
2a. Reliability: H-2; M-3; L-1; I-0; **2b. Validity:** H-3; M-3; L-0; I-0
Rationale:

- Testing of reliability is at the data element level only. Validity testing is at the level of the measure score.
- Is Denominator based on clinical criteria (2a1.7) or ICD-9 codes? Will require expert manual review of records to assess appropriate interpretation--e.g. many patients will be admitted with serious infection but not sepsis and develop sepsis syndrome soon after admission; such patients will already be on antimicrobials so how will the measure of "blood cultures before antibiotics" be resolved? And who interprets "if obtaining blood cultures might delay administration of antibiotics?" Also who interprets if antimicrobials are appropriate "that have activity against all likely pathogens?"   In this day of increasing antimicrobial resistance, there really are none!!!
- 2a. Reliable, but needs EHR. Difficulty in identifying the denominator exclusion details and not clear how they would be accounted for unless charts are reviewed. Reliability in one center was high.
- 2b. Good external validity. Specific components of measure have been used in multiple studies involving a large number of patients.

3. Usability: H-5; M-1; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

- Based on studies and meta-analyses cited, but I have many questions of the logistics of medical record review
- 3a. Results meaningful and understandable
- 3b. Public reporting may help improve compliance and improve morbidity/ mortality

4. Feasibility: H-3; M-2; L-1; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:

- See prior comments about concerns I have for medical record interpretation
- 4b. Some of the data is abstracted by healthcare personnel, which will require large time commitment. The answer to near path to electronic collection was not addressed by the developer.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-4; N-1
Rationale:

- I strongly agree with this multi-measure proposal, but I need further clarification of how to acutally implement this outside of research study application. I believe this can be done but needs further discussion.
- Severe sepsis and septic shock are associated with poor outcomes. Improving the process will help reduce patient morbidity and mortality

Additional Comments/Questions:

- Of note a NQF steering committee recently discussed consideration to not renew a measure for severe pneumonia for the performing of blood cultures within 24 hours of admission to ICU. Since pneumonia is one of the most common causes of severe sepsis, there may be a potential conflict of measures. Personally I believe patients going to ICU for severe pneumonia should have blood cultures and there is level II data to support this and we can discuss.

Workgroup Discussion

0500: Severe sepsis and septic shock: Management bundle

Importance to Measure and Report

- The Sepsis Management bundle is aligned with the Surviving Sepsis Campaign – a global initiative to improve outcomes for sepsis and septic shock.
- Clinically there is not much controversy on the benefits of the bundle - most studies demonstrate a mortality benefit - more than 60 studies - but are mainly observational.
- There are differing opinions on some components of the bundle – particularly the use of central lines and CVP, especially in patients with sepsis and not septic shock.
- The measure requires coordination of care between the ED and ICU for optimal patient outcomes.

Scientific Acceptability of Measure Properties

- Some concerns with the specifications:
  o How are "timely fashion" "effective antibiotics" and "all likely pathogens" defined and who interprets it?
  o ICD-9 codes are not included in the specifications.
  o Is the diagnosis determined in the ED or at hospital discharge – potential lack of consistency?
  o What is the time window?
  o When to start the clock if a patient develops sepsis after initial presentation
  o Lack of risk-adjustment.

Usability

- Highmark is using the measure in its pay for performance program for past 2 years – initially had some issues with data collection, but soon resolved
- University of Kansas uses the measure in their EHR with real-time notifications

Feasibility

- Requires lots of data collection – feasibility concerns

0399: Paired Measure: Hepatitis C: Hepatitis A vaccination (paired with 0400)
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: Percentage of patients aged 18 years and older with a diagnosis of hepatitis C who have received at least one injection of hepatitis A vaccine, or who have documented immunity to hepatitis A
Numerator Statement: Patients who have received at least one injection of hepatitis A vaccine, or who have documented immunity to Hepatitis A
Denominator Statement: All patients aged 18 years and older with a diagnosis of hepatitis C
Exclusions: Documentation of medical reason(s) for not receiving at least one injection of hepatitis A vaccine
Documentation of patient reason(s) for not receiving at least one injection of hepatitis A vaccine
Adjustment/Stratification: No risk adjustment or risk stratification We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.
Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team
Type of Measure: Process
Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Registry
Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:** American Association for the Study of Liver Diseases, American Gastroenterological Association Institute

0399: Hepatitis C: Hepatitis A vaccination
IMPLEMENTATION COMMENTS:
• IDSA: IDSA supports endorsement of measure #0399.The current measure identifies patients who received at least one injection of hepatitis A vaccine or have documented immunity to hepatitis A. This is clinically important. This measure would be improved if it identified completion of the hepatitis A vaccine or documented immunity. Our members note the challenge in meeting this measure related to insufficient insurance coverage or no insurance coverage often encountered with some Hep C Patients.
o   Developer response: Thank you for your comment.

Notes

Workgroup Preliminary Evaluations
The following evaluation ratings and comments are from the Committee Reviewers: Tiffany Osborn; Rekha Murthy; Curtis Collins; Mary Blank; Thomas File; Mohamad Fakih   *(comments separated by bullets)*
Importance to Measure and Report (*based on decision logic*): <mark>Y-5; N-1</mark>
1a. Impact: H-5; M-0; L-0; I-1; 1b. Performance Gap: H-4; M-1; L-0; I-1
Rationale:
• Discusses Hep C but not co-infection with Hep A
• 67.47% mean performance of TIN/NPIs (what do initials stand for??)
• Vaccination rates still remain low in the country (Hepatology. 2011 Oct;54(4):1167-78) this measure may improve Hepatitis A vaccination rates and reduce risk of further liver damage if exposed to Hepatitis A.
1c. Evidence (*based on decision logic*): <mark>Y-5; N-1</mark>   IF a **Health Outcome**, rationale supports: <mark>Y-0; N-0; NA-6</mark>
Quantity: H-3; M-1; L-1; I-1; Quality: H-3; M-2; L-1; I-0; Consistency: H-3; M-2; L-0; I-1
Rationale:
• Additional research unlikely to change conclusion
• Many studies support Hep A vaccination for Hep C patients. Recent study shows gaps in vaccination in the VA population with chronic Hepatitis C infection. Although incidence of superinfection with acute HBV and HAV was low, but it was significantly lower in vaccinated patients. Hepatology. 2011 Jan;53(1):42-52
2. Scientific Acceptability of Measure Properties (*based on decision logic*): <mark>Y-5; N-0</mark>
2a. Reliability: H-4; M-1; L-0; I-0; 2b. Validity: H-3; M-2; L-0; I-0
Rationale:
• The measure includes at least one dose of hep A vaccine given. Different responses to vaccine occur with the number of doses

0399: Paired Measure: Hepatitis C: Hepatitis A vaccination (paired with 0400)
given (total of 2), although serologic response to one dose of Hep A vaccine is better than 1 dose of Hep B vaccine

3. Usability: H-3; M-2; L-0; I-0
(Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement)
Rationale:
- Yes. It will help physicians focus on improvements in the care of Hepatitis C patients. It will also provide to patients and purchasers tools to evaluate care.

4. Feasibility: H-3; M-2; L-0; I-0
(4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)
Rationale:
- Physician practices need to have EHR implemented.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-5; N-1
Rationale:
- Meets criteria for endorsement. required EHR for implementation

Additional Comments/Questions:

Workgroup Call Summary

Importance to Measure and Report
- Although not mandatory, CDC recommends that individuals with hepatitis C should get the hepatitis A and B vaccinations.
- The cited guideline rates the evidence as Class IIa – "Weight of evidence/opinion is in favor of usefulness/efficacy" and Level of Evidence C – "Only consensus opinion of experts, case studies, or standard of care."
- The WG noted that this is an important process of care but there may be no randomized clinical trials to support it. However, the WG asked the developer to provide data on immunogenicity as well as observational data on hepatitis vaccination.
  - The developer will provide the data prior to the in-person meeting.
- A WG member stated that measures being submitted to NQF for use as a quality measure and/ or for pay for performance should have evidence to support the measure focus.
- Immunization rates for Hepatitis A in children are rising.
- WG members asked the developers to provide any available evidence, even indirect, to support the measure.

Scientific Acceptability of Measure Properties
- The measure only specifies one injection – WG member thought this was a marker of vaccination – a single injection generally confers about 80 percent immunity.

Feasibility
- The WG was slightly concerned with the measure requiring only one injection of the hepatitis A vaccine. Members of the workgroup felt that if the patients had one injection, they will more than likely receive the remaining injections.

0400: Paired Measure: Hepatitis C: Hepatitis B vaccination (paired with 0399)

Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: Percentage of patients aged 18 years and older with a diagnosis of hepatitis C who have received at least one injection of hepatitis B vaccine, or who have documented immunity to hepatitis B
Numerator Statement: Patients who have received at least one injection of Hepatitis B vaccine, or who have documented immunity to Hepatitis B
Denominator Statement: All patients aged 18 years and older with a diagnosis of hepatitis C
Exclusions: Documentation of medical reason(s) for not receiving at least one injection of hepatitis B vaccine
Documentation of patient reason(s) for not receiving at least one injection of hepatitis B vaccine
Adjustment/Stratification: No risk adjustment or risk stratification We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.
Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team
Type of Measure: Process
Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Registry
Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:** American Association for the Study of Liver Diseases, American Gastroenterological Association Institute

0400: Hepatitis C: Hepatitis B vaccination

IMPLEMENTATION COMMENTS:
- IDSA: IDSA supports endorsement of measure #0400.This is clinically important. This measure would be improved if it identified completion of the hepatitis B vaccine or documented immunity. Our members note the challenge in meeting this measure related to insufficient insurance coverage or no insurance coverage often encountered with some Hep C Patients.
  - Developer response: Thank you for your comment.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Tiffany Osborn; Rekha Murthy; Curtis Collins; Mary Blank; Thomas File; Mohamad fakih   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-1
1a. Impact: H-4; M-0; L-0; I-1; 1b. Performance Gap: H-4; M-1; L-0; I-0
Rationale:
- Does not describe impact for co-infection with Hep C and Hep B
- 80.93% mean performance of TIN/NPIs
- Compliance with Hepatitis B vaccination is high. However, there is room for improvement with 11% of patients not being vaccinated. Gaps exist for minorities with Hepatitis C.

1c. Evidence (*based on decision logic*): Y-4; N-2   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-6
Quantity: H-2; M-3; L-0; I-1; Quality: H-1; M-4; L-1; I-0; Consistency: H-1; M-3; L-1; I-1
Rationale:
- No specific studies listed as evidence, but Class IIa level of evidence in Guideline
- Many studies support Hep B vaccination for Hep C patients. Recent study shows gaps in vaccination in the VA population with chronic Hepatitis C infection. In addition, incidence of superinfection with acute HBV and HAV was low, but it was significantly lower in vaccinated patients. Hepatology. 2011 Jan;53(1):42-52

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-6; N-0
2a. Reliability: H-4; M-2; L-0; I-0; 2b. Validity: H-3; M-3; L-0; I-0
Rationale:
- The measure includes at least one dose of hep B vaccine given. Different responses to vaccine occur with the number of doses given, leading to different % of patients with positive serology post-vaccination

3. Usability: H-3; M-3; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:
- 3a. Using EHR, this measure will provide meaningful data for physicians on their performance.
- 3b. Yes, but I am not sure that it will lead to further improvements in vaccination rates. We still the issue with having at least one dose documented (not three)

4. Feasibility: H-3; M-3; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:
- Unless EHR is present, it will be extremely difficult to obtain data accurately.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-5; N-0
Rationale:
- This measure will be very difficult if the physician practice lacks EHR. It provides important information to improve vaccination of Hep C patients if the data is extracted from EHR.

Additional Comments/Questions:
- Based on benefit/harm ratio, this should be done

Workgroup Discussion

0399: Paired Measure: Hepatitis C: Hepatitis A vaccination (paired with 0400)

Essentially the same issues as for measure 399.

Importance to Measure and Report

- Although not mandatory, CDC recommends that individuals with hepatitis C should get the hepatitis A and B vaccinations.
- The WG noted that this is an important process of care but there may be no randomized clinical trials to support it. However, the WG asked the developer to provide data on immunogenicity as well as observational data on hepatitis vaccination.
    - o The developer will provide the data prior to the in-person meeting.
- A WG member stated that measures being submitted to NQF for use as a quality measure and/ or for pay for performance should have data to support the measure focus.

Feasibility

- The WG was concerned with the measure requiring only one injection of the hepatitis B vaccine. The hepatitis B vaccine is usually given in 3 injections so the WG felt that one injection was a marker of likely full vaccination

0393: Hepatitis C: Testing for chronic hepatitis C – Confirmation of hepatitis C viremia
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: Percentage of patients aged 18 years and older with a diagnosis of hepatitis C seen for an initial evaluation who had HCV RNA testing ordered or previously performed
Numerator Statement: Patients for whom HCV RNA testing was ordered or previously performed
Denominator Statement: All patients aged 18 years and older with a diagnosis of hepatitis C seen for initial evaluation
Exclusions: Documentation of medical reason(s) for not ordering or performing HCV RNA testing
Documentation of patient reason(s) for not ordering or performing HCV RNA testing
Adjustment/Stratification: No risk adjustment or risk stratification None We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.
Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team
Type of Measure: Process
Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Registry
Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:** American Association for the Study of Liver Diseases, American Gastroenterological Association Institute
0393: Hepatitis C: Testing for chronic hepatitis C - Confirmation of hepatitis C viremia
IMPLEMENTATION COMMENTS:
• IDSA: The IDSA supports the endorsement of measure #0393. Our members report that, in their experience, a minority of referring physicians (~25%) are performing this testing. As well, having the measure call for reflexing positive HCV Ab screens to additional HCV RNA tests could be more useful.   o Developer response: Thank you for your comment.
Notes
Workgroup Preliminary Evaluations
The following evaluation ratings and comments are from the Committee Reviewers: Steven Brotman; Doug Campos-Outcalt; David Spach; Ray Chung   *(comments separated by bullets)*
Importance to Measure and Report (*based on decision logic*): Y-2; N-2   **1a. Impact:** H-4; M-0; L-0; I-0; **1b. Performance Gap:** H-0; M-3; L-1; I-0   **Rationale:**   • It is a high prevalent condition with large health impact. However, there was not evidence provided that this test is not being done.   • HCV infection common (approximately 4 million anti-HCV positive and approximately 2.7 million with chronic infection). Projections for next 15 years show huge burden of HCV disease in US. HCV RNA testing sorts out resolved/active infection which is critical since effective therapies available for HCV. Need some assistance in interpreting the gap data as reported in CMS PQRI.   **1c. Evidence** (*based on decision logic*): Y-2; N-2   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-4   **Quantity:** H-0; M-3; L-0; I-1; **Quality:** H-0; M-3; L-0; I-1; **Consistency:** H-0; M-2; L-1; I-1   **Rationale:**   • Studies on long term benefit or treatment, which results from the test, are all observational except one, and do not look at long term benefits / harms.   • Body of evidence does exist, but weakly addressed in measure. The measure defaults to AASLD Guidelines that were based on data and rated IB and 1A. Consistency not addressed.
2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-1   **2a. Reliability:** H-2; M-1; L-1; I-0; **2b. Validity:** H-1; M-2; L-1; I-0   **Rationale:**   • Test highly reliable and can easily be implemented consistently and tracked easily with EHR. Weakness is that EHR cannot easily capture exceptions for performing this test.
3. Usability: H-1; M-3; L-0; I-0   (*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

0393: Hepatitis C: Testing for chronic hepatitis C – Confirmation of hepatitis C viremia
Rationale: • This test is already used at high rates.   • 3a. Measure has already been in place in PQRS since 2008.   • 3b. Reporting data available on PCPI website and thus easy to monitor and evaluate for QI purposes
4. Feasibility: <mark>H-2; M-2; L-0; I-0</mark>   *(4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)*   **Rationale:**   • 4a. Any patient evaluated for positive anti-HCV should have HCV RNA generated as part of routine clinical care   • 4b. HCV RNA easily available in EHR   • 4c. Should have noted unintended consequence of false-positive HCV RNA?   • 4d. Data collection is straightforward
Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-2; N-2</mark>   **Rationale:**   • This test is not under used, very little room for improvement   • Measuring HCV RNA in newly diagnosed anti-HCV positive is critical step to determine who needs long-term HCV care and further testing (e.g. genotype, estimation of fibrosis, etc). This measure is a highly reliable test, obtained as part of routine care, and is easy to track with EHR. Only major issue is lack of EHR method of easily documenting exceptions to why HCV RNA not obtained.   **Additional Comments/Questions:**   • Need some more information/guidance on gap care statistics and methodology used in CMS PQRS.
Workgroup Call Summary
Importance to Measure and Report   • Hepatitis C affects a large portion of the baby boomer population. Just last week CDC recommended that all adults born from 1945 to 1965 receive hepatitis C screening. More patients with Chronic HCV will be identified.   • According the PQRS data submitted, the compliance rate is 95.86 percent which leaves little opportunity for improvement. The WG struggled with the idea of whether or not there is still a need for improvement with such a high performance rate. It was noted that the data presented was on a small population and a WG member reported that according to CDC the performance rate in 2010 was 38.4 percent.   • The WG indicated that there is little to no disparities data available for hepatitis C for the individual performance measures, though minorities are over-represented in the population of patients with HCV...   • The WG indicated that the information on evidence only referenced the practice guideline and did not provided sufficient information to evaluate the quality, quantity and consistency of the evidence. AMA-PCPI indicated that they additional information to the Committee.   **Scientific Acceptability of Measure Properties**   • This measure cannot receive a rating of high for reliability because it was only tested at the measure score; the measure can only be rated moderate.

0584: Hepatitis C: Viral load test

Status: Maintenance, Original Endorsement: Dec 04, 2009

Description: This measure identifies the percentage of patients with chronic Hepatitis C (HCV) who began HCV antiviral therapy during the measurement year and had HCV Viral Load testing 6 months prior to initiation of antiviral therapy.

Numerator Statement: Patients in the denominator who had an HCV Viral Load test 6 months prior to the initiation of antiviral therapy.

Denominator Statement: Our denominator is anyone with Hepatitis C diagnosed anytime in the past, based on historical claims on file, who have a new start of peginterferon in the last year, excluding people with documentation of a medical reason(s) for not performing quantitative HCV RNA testing within 6 months prior to initiation of treatment (CPT Category II code 3218F-1P).

Exclusions: Exclude anyone with a code which states the patient has a medical reason for not having the test done.

Adjustment/Stratification: No risk adjustment or risk stratification   No stratification.

Level of Analysis: Health Plan

Type of Measure: Process

Data Source: Administrative claims

Measure Steward: Resolution Health, Inc.

0584: Hepatitis C: Viral load test

IMPLEMENTATION COMMENTS:
- IDSA: IDSA supports endorsement of measure #0584. As noted earlier, this measure appears to be similar to measure #0395.
  - Developer response: While measure 0395 relies exclusively upon CPT4 category II codes for recognizing quantification of viral load, we consider category I billing claims as well laboratory test results tagged with LOINC codes. Further, our definition of chronic HCV infection includes an additional four ICD-9-CM codes, accepting a history of hepatic coma (070.44), unspecified disease acuity (070.70, 070.71), and a declaration of being an HCV carrier (V0262). Other differences between the two measures have been harmonized.
- HIV Medicine Association: We support continued NQF endorsement of this measure, as written, as it is still clinically relevant.
  - Developer response: The American Association for the Study of Liver Diseases rated the measure and the evidence supporting the measure Class I, Level A in 2009. This is based on the American College of Cardiology and American Heart Association Practice Guidelines.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Steven Brotman; Doug Campos-Outcalt; David Spach; Ray Chung   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-0

1a. Impact: H-3; M-1; L-0; I-0; **1b. Performance Gap:** H-0; M-4; L-0; I-0

Rationale:
- High rates of performance currently
- 1a. HCV has major disease burden in US. HCV RNA testing important prior to starting therapy for multiple reasons.
- 1b. Performance gap reported from 1.8 million administrative claims.

1c. Evidence (*based on decision logic*): Y-4; N-0   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-4

Quantity: H-3; M-1; L-0; I-0; **Quality:** H-1; M-3; L-0; I-0; **Consistency:** H-1; M-3; L-0; I-0

Rationale:
- Meta-analysis used had 12 studies. Extensive data from solid trials. But, do not agree with statement as listed--However, patients with rapid virologic response no matter the genotype, respond to short treatment (12-16 weeks). Even with DAA-based therapy and RGT, treatment for GT-1 is NOT 12-16 weeks. The statements in the measure are not entirely accurate--with RGT and GT-1 Rx, shortened duration of therapy not just based on Rapid Virologic Response--requires extended rapid virologic response.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-2; N-1

2a. Reliability: H-2; M-1; L-1; I-0; **2b. Validity:** H-1; M-1; L-1; I-0

Rationale:

0584: Hepatitis C: Viral load test

- 2a. Measure precise and easy to implement.
- 2b. Performance results reported. Validity testing did not include review of medical records. Uncertain how would identify if patient had recent HCV RNA level obtained by another medical provider (eg. referring physician obtains, sends to expert who then initiates therapy without repeating the HCV RNA level.

3. Usability: <mark>H-1; M-3; L-0; I-0</mark>
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

- 3a. Reporting internal to specific organization.
- 3b. Reporting 92% in 2011 for old measure (viral load any time prior to starting therapy)--with new measure (viral load within 6 months of starting therapy) only 68-84%.

4. Feasibility: <mark>H-1; M-3; L-0; I-0</mark>
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:

- 4a. Test easy to obtain and normally is obtained as part of routine care.
- 4b. Test easy to locate in EMR. Only issue is not having EMR documentation of recent HCV RNA obtained by another provider (with test in another EMR).
- 4c. If test was inaccurate, would have major consequences.
- 4d. Collection via electronic claims

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-3; N-1</mark>
Rationale:

- The measure is suitable but nearly identical to measure 0395. I do NOT understand why there are two nearly identical measures in place. It would seem to be a waste of resources and would create confusion.

Additional Comments/Questions:

- This is directly overlapping with 395, which has more evidence and the potential for corroboration with patient records. Would favor eschewing this one in favor of 395

<mark>**Workgroup Call Discussion**</mark>

Importance to Measure and Report

- This measure is similar to 0395 but it addresses a different level of analysis – health plan- and uses a different data source – administrative claims.

Scientific Acceptability of Measure Properties

- The measure will need to be harmonized (measure focus and target population) with the related measure, measure 0395.
- The developer noted that as a health plan level measure, administrative data can identify a test done by other providers within the plan.

0395: Paired Measure: Hepatitis C ribonucleic acid (RNA) testing before initiating treatment (paired with 0396)

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: Percentage of patients aged 18 years and older with a diagnosis of chronic hepatitis C who are receiving antiviral treatment for whom quantitative HCV RNA testing was performed within 6 months prior to initiation of antiviral treatment

Numerator Statement: Patients for whom quantitative HCV RNA testing was performed within 6 months prior to the intiation of antiviral treatment

Denominator Statement: All patients aged 18 years and older with a diagnosis of chronic hepatitis C who are receiving antiviral treatment

Exclusions: Documentation of medical reason(s) for not performing quantitative HCV RNA testing within 6 months prior to the initiation of treatment

Adjustment/Stratification: No risk adjustment or risk stratification None We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Registry

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:** American Association for the Study of Liver Diseases, American Gastroenterological Association Institute

0395: Paired Measure: Hepatitis C ribonucleic acid (RNA) testing before initiating treatment (paired with 0396)

IMPLEMENTATION COMMENTS:
- IDSA: IDSA does not support the endorsement of measure 0395 in the present form. Our physicians consider the standard of care to reflect updated treatment guidelines which are very specific concerning viral load measurements at various time points to determine continuation of therapy and/or length of response guided therapy (RGT) for GT1 patients on Protease Inhibitor therapy. In addition, measure 0395 appears to be similar to measure 0584, which suggests duplication/redundancy in measuring.
  - Developer response: While updated treatment guidelines are specific regarding viral load measurements at various time points in order to determine continuation of therapy and/or length of response guided therapy, guidelines also support HCV RNA testing prior to initiation of treatment in order to identify the best course of treatment for the patient. According to NIH guidelines, determination of the HCV level provides important information on the likelihood of response to treatment in patients undergoing antiviral therapy. Our measure differs from measure 0584 in that it includes the 6 month time window, for HCV RNA level measurement. The time window was added to ensure that there is a recent HCV RNA level recorded to maximize the likelihood that treatment is appropriate for the patient's current viral load.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the **Committee Reviewers:** Steven Brotman; Doug Campos-Outcalt; David Spach; Ray Chung   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-2

1a. Impact: H-3; M-1; L-0; I-0; **1b. Performance Gap:** H-0; M-4; L-0; I-0

Rationale:
- 1a. High burden of HCV. Baseline HCV RNA has important implications for treatment. Also need to document patient still has chronic HCV prior to starting therapy.
- 1b. Gap in care for this measure shown based on 2008-2010 CMS PQRI reporting.

1c. Evidence (*based on decision logic*): Y-2; N-2   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-4

Quantity: H-0; M-2; L-1; I-1; **Quality:** H-0; M-2; L-1; I-1; **Consistency:** H-0; M-2; L-1; I-1

Rationale:
- This is a test that is used based on concensus rather than proof of effectiveness
- Evidence from practice guideline. Direct data not provided.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-1

0395: Paired Measure: Hepatitis C ribonucleic acid (RNA) testing before initiating treatment (paired with 0396)

2a. Reliability: H-2; M-1; L-1; I-0; 2b. Validity: H-1; M-2; L-1; I-0

Rationale:

- 2a. The measure is precisely identified and timeframe identified. Test used is highly reliable. I don't think there really is any valid medical reason why NOT to perform quantitative HCV RNA testing within 6 months of starting therapy, unless patient had it obtained via another medical provider or via a research study.
- 2b. False negatives can occur, but in less than 4%. Results reported from CMS PQRI.

3. Usability: H-2; M-2; L-0; I-0

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale:

- 3a. In use in PQRS since 2008
- 3b. Measure easy to understand and easy to implement.

4. Feasibility: H-2; M-1; L-0; I-0

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

Rationale:

- 4a. Quantitative HCV RNA should be obtained as part of routine care prior to treatment.
- 4b. Available in EMR.
- 4c. Unintended consequences could occur with false-negative result (patient would not get HCV treated)
- 4d. Measure in use. Found to be feasible.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-3; N-1

Rationale:

- Obtaining quantitative HCV RNA within 6 months prior to treatment is critical step in management of HCV.
- A strength of this measure (compared with 584) is its confirmation of findings by chart review.

Additional Comments/Questions:

- This measure is nearly identical to 0584 (HCV Viral Load Test). These measures compete with each other.

Workgroup Call Discussion

Importance to Measure and Report

- The WG indicated that the information on evidence only referenced the practice guideline and did not provided sufficient information to evaluate the quality, quantity and consistency of the evidence. AMA-PCPI indicated that they additional information to the Committee.
- The WG noted that the issues identified for measure 0393 applied to 0395.

Scientific Acceptability of Measure Properties

- The WG determined that measures 0395 and 0393 were not overlapping; each measurement usually occurs years apart.
- It was noted that testing before beginning treatment is important because occasional (though rare) patents may spontaneously clear the virus.

0396: Paired Measure: HCV genotype testing prior to treatment (paired with 0395)

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: Percentage of patients aged 18 years and older with a diagnosis of chronic hepatitis C who are receiving antiviral treatment for whom HCV genotype testing was performed prior to initiation of antiviral treatment

Numerator Statement: Patients for whom HCV genotype testing was performed prior to initiation of antiviral treatment

Denominator Statement: All patients aged 18 years and older with a diagnosis of chronic hepatitis C who are receiving antiviral treatment

Exclusions: None

Adjustment/Stratification: No risk adjustment or risk stratification None We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Registry

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:** American Association for the Study of Liver Diseases, American Gastroenterological Association Institute

0396: HCV genotype testing prior to treatment

IMPLEMENTATION COMMENTS:

- IDSA: IDSA supports the endorsement of measure 0396 with modifications. This is standard of care and is typically required by health plans as part of the prior authorization process, indicating GT1a vs GT1b. The 6 month timeframe is not appropriate. Once chronic infection is established, the genotype is unchanged. Therefore, a genotype obtained prior to 6 months before the start of treatment is acceptable; retesting is unnecessary and might represent overuse.
  - o Developer response: There is no 6 month time window associated with this measure. The measure description is as follows: Percentage of patients aged 18 years and older with a diagnosis of chronic hepatitis C who are receiving antiviral treatment for whom HCV genotype testing was performed prior to initiation of antiviral treatment. The numerator time window is "once prior to initiation of antiviral treatment." The 6 month time window is associated with measure 0395 only.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the **Committee Reviewers:** Steven Brotman; Doug Campos-Outcalt; David Spach; Ray Chung   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-2

1a. Impact: H-3; M-1; L-0; I-0 **1b. Performance Gap:** H-0; M-3; L-1; I-0

Rationale:

- Seems to be high compliance now
- 1a. HCV disease burden substantial and projected to have even bigger impact. HCV genotype critical for determining appropriate therapy. Current state-of-art therapy completely different for GT1 than GT 2 or 3.
- 1b. Gap in care shown in CMS PQRI data given, but I find it incredibly hard to believe any medical provider in 2012 would treat HCV without knowing baseline genotype. Disparity in treatment response well documented. Not clear that medical providers less frequently test HCV genotype in minorities. Statement on genotype 1b is misleading--"...most favorable response." This should be clarified most favorable among GT1, not more favorable than GT2 or GT3.

1c. Evidence (*based on decision logic*): Y-3; N-1   **IF** a **Health Outcome**, rationale supports: Y-0; N-0; NA-4

Quantity: H-1; M-1; L-1; I-1 Quality: H-1; M-2; L-0; I-1 Consistency: H-1; M-2; L-0; I-1

Rationale:

- The use of genotyping for treatment decisions seems well acceted, but not much documentation was presented for it.
- Evidence for importance of relationship of GT and response is extensive and consistent. The information provided here just cites guidelines and states consistency not addressed by guidelines. Thus, I am rating these High based on my knowledge of subject

0396: Paired Measure: HCV genotype testing prior to treatment (paired with 0395)
matter, not information provided.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-1
2a. Reliability: H-2; M-1; L-1; I-0; **2b. Validity:** H-2; M-1; L-1; I-0
Rationale:
- 2a. Measure clearly stated; measure is reliable and can easily and consistently be implemented.
- 2b. Measure has high validity. Performance measures given.

3. Usability: H-1; M-2; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:
- 3a. Measure has been in place in PQRS since 2008
- 3b. Results easy to understand and interpret.

4. Feasibility: H-2; M-2; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:
- 4a. Routinely obtained during care.
- 4b. Genotype data will appear in lab data in EHR.
- 4c. Potential error in GT would lead to wrong treatment regimen.
- 4d. Already in use and found to be reliable.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-2; N-2
Rationale:
- Did not provide evidence of the usefulness of the test. I think this evidence exists but the authors did not present it. Instead they rely or common practice and expert opinion.
- Performing HCV genotype is mandatory for any patient who will receive therapy for HCV. Lab test is easy to obtain (as long as patient has circulating HCV), reliable, and easy to understand. Results have profound implications for treatment regimen chosen.

Additional Comments/Questions:
- Performing HCV Genotype is essential, but as noted earlier, I find it hard to believe that there are medical providers out there who will treat HCV without knowledge of the patient's HCV genotype. Concerned that gap identified may not represent true lack of knowledge of patient's HCV genotype, but may more likely reflect pateint had prior GT performed and info not easily accessible in EHR.

Workgroup Call Summary

Importance to Measure and Report
- The WG indicated that the information on evidence only referenced the practice guideline and did not provided sufficient information to evaluate the quality, quantity and consistency of the evidence. AMA-PCPI indicated that they additional information to the Committee.

Scientific Acceptability of Measure Properties
- A WG member was concerned with the kappa score of 0.56.

0397: Hepatitis C: Antiviral treatment prescribed
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: Percentage of patients aged 18 years and older with a diagnosis of chronic hepatitis C who were prescribed at a minimum peginterferon and ribavirin therapy within the 12 month reporting period
Numerator Statement: Patients who were prescribed at a minimum peginterferon and ribavirin therapy within the 12 month reporting period
Denominator Statement: All patients aged 18 years and older with a diagnosis of chronic hepatitis C
Exclusions: Documentation of medical reason(s) why a patient was not prescribed at a minimum peginterferon and ribavirin therapy (eg, patient was not a candidate for therapy, could not tolerate)
Documentation of patient reason(s) why a patient was not prescribed at a minimum peginterferon and ribavirin therapy (eg, patient declined)
Documentation of system reason(s) why a patient was not prescribed at a minimum peginterferon and ribavirin therapy (eg, patient has no insurance coverage, therapy not covered)
Adjustment/Stratification: No risk adjustment or risk stratification None We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.
Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team
Type of Measure: Process
Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry
Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:** American Association for the Study of Liver Diseases, American Gastroenterological Association Institute

0397: Hepatitis C: Antiviral treatment prescribed
IMPLEMENTATION COMMENTS:

- IDSA: IDSA does not support the endorsement of measure #0397. Although all patients should be considered potential treatment candidates, the measure logic does not provide a feasible, usable way to identify denominator exclusions. The denominator exclusion data sources used in this measure are rarely submitted (e.g., CPT-II codes). Given this major limitation, this is a measure of resource use, not quality of care. Endorsement is not recommended. Also, the measure description does not address protease inhibitor use in the genotype 1 HCV infected person.
  - Developer response: For clarification, this measure does not include exclusions, but includes medical, patient, and system exceptions. In the AMA-PCPI methodology, exclusions are absolute and apply to all patients and therefore are not part of clinical judgment within a measure. Exceptions are used to remove patients from the denominator of a performance measure when a patient does not receive a therapy or service AND that therapy or service would not be appropriate for the patient due to specific reasons. Exceptions are not absolute, and are based on clinical judgment, individual patient characteristics, or patient preferences. This measure was included in the PQRS program in 2008, 2009, and 2010 and we have not received feedback regarding difficulty in reporting the exceptions through the use of CPT II codes and modifiers. EHR specifications have also been submitted for electronic reporting of this measure. The updated evidence-based guideline from AASLD, published in 2011, describes treatment with protease inhibitors for genotype 1 patients, in addition to peginterferon and ribavirin therapy. As such, the measure language has been updated to capture this treatment. The measure description is written as follows: "Percentage of patients aged 18 years and older with a diagnosis of chronic hepatitis C who were prescribed at a minimum peginterferon and ribavirin therapy within the 12 month reporting period," with "at a minimum" intended to allow for the additional treatment recommended for genotype 1 patients.

Notes

Workgroup Preliminary Evaluations
The following evaluation ratings and comments are from the Committee Reviewers: Steven Brotman; Doug Campos-Outcalt; David Spach; Ray Chung *(comments separated by bullets)*
Importance to Measure and Report (*based on decision logic*): Y-2; N-2
1a. Impact: H-4; M-0; L-0; I-0; 1b. Performance Gap: H-0; M-4; L-0; I-0
Rationale:

0397: Hepatitis C: Antiviral treatment prescribed

- 1a. HCV common problem in US with projected very high disease burden. Effective treatment of HCV can lead to permanent eradication/cure.
- 1b.Gap in care shown in CMS PQRI data. African-American are patients treated less frequently than Caucasian and with lower response rates.
- 1a. large population numbers, with finite progression to ESLD and death from liver failure.
- 1b. Important caveat, a very large number of pts are untreated because of perceived intolerability or prior treatment experience. This is a treatment area in flux, as gt 1 SOC is now PEG/RBV + TVR or BOC, but soon will give way to IFN sparing all oral regimens in the next 2-4 years.

1c. Evidence (*based on decision logic*): Y-2; N-2   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-4
Quantity: H-2; M-0; L-1; I-1; Quality: H-1; M-1; L-1; I-1; Consistency: H-1; M-1; L-0; I-2
Rationale:

- Evidence for benefits/harms of treatment was not presented.
- The body of evidence is extensive regarding treatment response rates. Measure cites AASLD 2011 Guidelines. Multiple new studies have shown effectiveness of GT1 with PegINF + RBV + DAA. Results have been consistent and excellent RCT have been performed.
- Many studies to support high rates of sustained response.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-2; N-2
2a. Reliability: H-0; M-2; L-1; I-1; 2b. Validity: H-0; M-2; L-1; I-1
Rationale:

- It is not clear that if patients decline treatment or have no viral load that they can be excluded from the deniminator.
- 2a. The measure accounts for new DAA therapy for GT1, but by stating "minimum of PEGINF + RBV" it is implying PEGINF + RBV alone is acceptable therapy for GT1 (which is not what most experts would recommend).
- 2b. Problem with denominator is that some experts are deferring therapy in stable patients to wait for INF-free regimens likely available in future. Also, denominator does not account for previously-treated patients. Information why patient NOT treated may be more difficult to obtain from EMR.

3. Usability: H-1; M-3; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

- 3a. Currently in use in PQRS and has been since 2008.
- 3b. Measures suitable for QI. Measures suitable for public reporting and QI, but denominator (fail to treat) reasons may be difficult to discern--some providers may be unfairly rated.

4. Feasibility: H-0; M-4; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:

- 4a. Information related to treatment would be available as part of routine care.
- 4b. Required elements in EHR, except for possible rationale related to deferring therapy.
- 4c. None noted
- 4d. Measure feasible to implement--this is more complicated in that measure already in place, but has been modified.
- 4c. The potential for misreporting exists with regard to prior treatment experience (nonresponse, intolerability)

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-2; N-2
Rationale:

- Evidence for short and long term benefits/harms of treatment were not provided.
- Treatment of chronic HCV is critical in reducing morbidity, eradicating HCV, and preventing transmission. Essential that all patients with chronic HCV undergo consideration for treatment.
- Overall meets criteria for endorsement.

Additional Comments/Questions:

- The evidence for short term benefit exists and the authors could add this to the document.
- This measure has several nuances--eg. modification of prior regimen based on new DAA agents. Measure complicated by fact that some experts will want to defer HCV therapy 1-3 years for INF-free regimens for certain patients.

Workgroup Call Summary
Importance to Measure and Report

0397: Hepatitis C: Antiviral treatment prescribed

- The WG indicated that the information on evidence only referenced the practice guideline and did not provided sufficient information to evaluate the quality, quantity and consistency of the evidence. AMA-PCPI indicated that they additional information to the Committee.
- WG members noted that treatment for HCV is rapidly evolving. Some clinicians are waiting until newer, oral, non-interferon drugs are available in 1-2 year before treating.
- The PQRS data presented identified a mean value of 68%. The developer was asked what the exception rate was – they will bring the data to the meeting. A WG member suggested that the number of patients that will be excluded by the measure through the exceptions, e.g., choosing to delay treatment, significant co-morbidities, intolerance to medications, previous poor response to medications, patient refusal, cost, will be greater than those that are captured in the measure.

Scientific Acceptability of Measure Properties

- The Committee noted that the availability of enhanced drugs is in the near future. Physicians may be deferring treatment for their patients until the new drugs/treatment are available. The developer acknowledged that the measure is currently flexible to allow for modifications when new drugs become available.
- The Committee discussed the need for exclusion for physician or patient deferral of treatment due to new upcoming treatments.
  - The developer said this would be a medical reason.

0398: Hepatitis C: HCV RNA testing at week 12 of treatment
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: Percentage of patients aged 18 years and older with a diagnosis of chronic hepatitis C who are receiving antiviral treatment for whom quantitative HCV RNA testing was performed at no greater than 12 weeks from initiation of antiviral treatment
Numerator Statement: Patients for whom quantitative HCV RNA testing was performed at no greater than 12 weeks from the initiation of antiviral treatment
Denominator Statement: All patients aged 18 years and older with a diagnosis of chronic hepatitis C who are receiving antiviral treatment
Exclusions: Documentation of medical reason(s) for not performing quantitative HCV RNA testing at no greater than 12 weeks from the initiation of antiviral treatment Documentation of patient reason(s) for not performing quantitative HCV RNA testing at no greater than 12 weeks from the initiation of antiviral treatment
Adjustment/Stratification: No risk adjustment or risk stratification None We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.
Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team
Type of Measure: Process
Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Registry
Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI)

0398: Hepatitis C: HCV RNA testing at week 12 of treatment
IMPLEMENTATION COMMENTS: • IDSA: IDSA supports the endorsement of measure #0398 with modification. RNA testing at 12 weeks is a reasonable measurement criterion for patients with genotype 2/ 3 infection. This measure should be modified to reflect management of patients with genotype 2/3 only. This measure does not address new RNA testing guidelines for genotype 1 patients on protease inhibitor treatment. A future measure should address this area.   o Developer response: The AMA-PCPI uses evidence-based guidelines to support the development of AMA PCPI measures. The updated evidence-based guideline from AASLD, published in 2011, indicates that treatment modifications may be necessary for genotype 1 patients at weeks 4, 8, or 12, based on HCV RNA levels and dependent upon the type of treatment the patient is receiving (including the use of protease inhibitors). Therefore, consistent with the guideline, the numerator language was updated to capture "Patients for whom quantitative HCV RNA testing was performed at **no greater than 12 weeks** from the initiation of antiviral treatment." The numerator definition is as follows: **12 Weeks from Initiation** – Patients for whom testing was performed between 4-12 weeks from the initiation of antiviral treatment will meet the numerator for this measure (depending upon the specific antiviral therapy used).

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Steven Brotman; Doug Campos-Outcalt; David Spach; Ray Chung   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-2

1a. Impact: H-3; M-1; L-0; I-0 1b. Performance Gap: H-1; M-3; L-0; I-0

Rationale:

- 1a. HCV has high disease burden and treatments are now available for all HCV GTs. Measurement of HCV RNA levels early in treatment has become extremely important component of modern therapy. Early week 4 and 12 responses (RGT) can determine required duration of therapy for patients with GT1. With GT1-detectable virus at week 12 strong predictor of treatment failure; undetectable HCV RNA at weeks 4 and 12 strong predicts good response with shortened duration therapy.
- 1b. Gap in "care" reported in measure--uncertain if gap in measurement of HCV RNA < 12 weeks. Gap in care and treatment responses well documented.

1c. Evidence (*based on decision logic*): Y-2; N-2   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-4

0398: Hepatitis C: HCV RNA testing at week 12 of treatment

Quantity: H-0; M-2; L-1; I-1; **Quality:** H-0; M-2; L-0; I-2; **Consistency:** H-0; M-2; L-0; I-2

Rationale:
- A guideline as used as the citation. It was based on assessment of studies but it is hard to assess without actually seeing it.
- Body of evidence showing HCV RNA levels obtained early in treatment (week 12) have strong negative predictive value (eg. with older PEGINF + RBV regimens, failure to obtain Early Virologic Response (> 2 log reduction in HCV RNA at week 12)) is strong predictor of treatment failure. This concept has been modified in modern treatment era.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-2; N-2

2a. Reliability: H-1; M-1; L-2; I-0; **2b. Validity:** H-1; M-1; L-2; I-0

Rationale:
- 2a. This measure is NOT precise. States obtain prior to week 12--this could be week 1, week 4, etc. The measure should be much more precise based on GT. With Telaprevir-based therapy, HCV RNA key measurements are at baseline, weeks 4, 12, and 24. With Boceprevir-based therapy, HCV RNA key measurements are baseline weeks, 8,12, and 24. The week 8 with boceprevir corresponds to 4 weeks after boceprevir started.
- 2b. Because the measure is imprecise, it is not valid.

3. Usability: H-1; M-3; L-0; I-0

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale:
- 3a. Has been in use in PRQS since 2008

4. Feasibility: H-2; M-2; L-0; I-0

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

Rationale:
- 4a. Drawing HCV RNA correspond with appropriate scheduled clinic visits after starting therapy.
- 4b. Information easily obtained via EMR.
- 4c. Test has high accuracy
- 4d. Measure is feasible. More precise recommendation for HCV RNA Testing after initiating therapy would be more difficult to implement since more complicated.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-2; N-2

Rationale:
- Assuming that the gudielne was based on solid evidence. If we should not make this assumption then I would check no.
- Checking HCV RNA responses early after starting therapy plays a much more important role in GT 1 than with GT2 and 3. Since treatment of GT1 with DAA + PEGING + RBV has moved to response-guided therapy, this measurement is outdated and will NOT have the same impact as a more precise measure. Further, the wording of the measurement is too vague (essentially check HCV RNA sometime before 12 week)--that is not a measure that will have significant impact on clinical practice and SVR rates with GT1 HCV.

Additional Comments/Questions:

Workgroup Call Summary

Importance to Measure and Report
- Early response to hepatitis C will aid in avoiding patients receiving toxic drugs for a long period of time.
- The WG indicated that the information on evidence only referenced the practice guideline and did not provided sufficient information to evaluate the quality, quantity and consistency of the evidence. AMA-PCPI indicated that they additional information to the Committee.
- The monitoring of viral load for various treatment regimens has become very specific. This measure is not specific enough as it counts any test from Week 1-12 and this would not be adequate monitoring.
    - The developer responded that they created a rather simple measure that would "get at the general concept" because a more specific measure would be too difficult due to the varying monitoring schedules of different drugs.

Scientific Acceptability of Measure Properties
- The measure is lacking precision; the timing of the test performed should be more specific. In order to drive quality improvement, the measure should be more defined. The developer noted that the focus of the measure is to ensure the patient receives treatment.
- Children and adolescents are usually not treated due to the years require to develop fibrosis.

0394: Hepatitis C: Counseling regarding use of contraception prior to antiviral treatment
Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: Percentage of female patients aged 18 to 44 years and all men aged 18 years and older with a diagnosis chronic hepatitis C who are receiving antiviral treatment who were counseled regarding contraception prior to the initiation of antiviral treatment

Numerator Statement: Patients who were counseled regarding contraception prior to the initiation of treatment

Denominator Statement: All female patients aged 18 to 44 years and all male patients aged 18 years and older with a diagnosis of chronic hepatitis C who are receiving antiviral treatment

Exclusions: Documentation of medical reason(s) for not counseling patient regarding contraception

Adjustment/Stratification: No risk adjustment or risk stratification None We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:** American Association for the Study of Liver Diseases, American Gastroenterological Association Institute

0394: Hepatitis C: Counseling regarding use of contraception prior to antiviral treatment

IMPLEMENTATION COMMENTS:

- IDSA: IDSA supports endorsement of measure #0394. Our physicians consider it to be the standard of care. They report the need for better standardization of information covered during counseling and standard documentation of methods of contraception (often 2).
    - Developer response: Thank you for your comment.
- HIV Medicine Association: We recommend that NQF carefully review the feasibility of this and other behavioral health measures, as performance of such screenings is not readily captured by most electronic health record (EHR) systems without a manual chart review.
    - Developer response: Based on the testing results, this measure was found to be feasible for implementation. We have provided testing data, which includes a comparison of E.H.R. automated reports to visual inspection of the medical record and had a kappa score of 0.54. This score shows that the measure is reliable and shows that the information can be accurately collected in both an electronic health record and a paper medical record.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Steven Brotman; Doug Campos-Outcalt; David Spach; Ray Chung *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-1; N-2

1a. Impact: H-1; M-3; L-0; I-0; 1b. Performance Gap: H-0; M-3; L-0; I-0

Rationale:

- It is not clear how well contraceptive counseling actually reduces pregnancies while on ribaviron. It is not clear why men with hep c need to be counseled.
- 1a. Burden of HCV disease in US substantial; ribavirin part of all treatment regimens; ribavirin pregnancy category X. Comment-- High impact summary on measure just defaults to high impact of HCV disease--does not specifically address high impact of this exact measure.
- 1b. Performance gap exists (seems small by CMS/PQRI data) and African-Americans have poorer response, Not clear if minority groups receive lower rates of counseling for use of contraception prior to starting ribavirin.

1c. Evidence (*based on decision logic*): Y-1; N-3   IF a **Health Outcome**, rationale supports: Y-0; N-1; NA-3

Quantity: H-0; M-1; L-2; I-1; Quality: H-0; M-1; L-2; I-1; Consistency: H-0; M-1; L-1; I-2

Rationale:

0394: Hepatitis C: Counseling regarding use of contraception prior to antiviral treatment
• No evidence presented that counseling on contraception actually reduces pregnancies. • Little data on toxicity. Rate yes based on potential benefit to patients outweigh potential harms.
2. Scientific Acceptability of Measure Properties (*based on decision logic*): <mark>Y-3; N-1</mark> **2a. Reliability: H-1; M-2; L-1; I-0; 2b. Validity: H-0; M-3; L-1; I-0** <u>Rationale:</u> • Some counseling not captured by E H R . • 2a. Wording of measure is ambiguous. Could be misinterpreted as counseling the patient to take contraceptives prior to treatment (versus counseling the patient prior to starting treatment that they will need to use contraception (during treatment and for 6 months after taking ribavirin)). • 2b. Main issue is that counseling is not a standard item in EMR. Analysis showed expert panel review (visual inspection) correlated with automated EHR report 2b.
3. Usability: <mark>H-2; M-2; L-0; I-0</mark> (*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*) <u>Rationale:</u> • 3a. Has been in use in PQRS since 2008. Easy to understand yes/no, did person receive counseling. • 3b. Measures are available on PCPI web site.
4. Feasibility: <mark>H-1; M-2; L-1; I-0</mark> (*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*) <u>Rationale:</u> • 4a. Data should routinely be generated in all patients prior to starting therapy as part of good clinical care. • 4b. Data regarding counseling likely to be more difficult to find in EMR as compared with lab data. • 4c. Only unintended consequence would be misinterpretation of measure based on ambiguous wording. • 4d. No difficultly and measure easy to implement.
Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-2; N-2</mark> <u>Rationale:</u> • Relationship between counseling and pregnancy avoidance not presented. • Critical that treatment does not cause permanent severe side effect as could occur with ribavirin use during pregnancy (or within 6 months of becoming pregnant). As noted earlier, wording of measure is NOT ideal and is ambiguous. From the measure, it is also not entirely clear exactly what wording should be in the counseling--that is also extremely important. Eg. Females: Use effective contraception so that you do not get pregnant while taking ribavirin and for the following 6 months after finishing ribavirin treatment. Males: Use effective contraception so that you do not get your female partner pregnant while you are taking ribavirin and for the 6 months after finishing ribavirin treatment. <u>Additional Comments/Questions</u>: • Not clear why men should be included.
<mark>**Workgroup Discussion**</mark>
Importance to Measure and Report • WG members indicated that this measure does not seem to be a measure of quality but more of a 'check the box' measure. There is no additional information captured in the measure other than the patient did or did not receive counseling. The Committee suggested that the developer create an outcome measure instead. • The WG indicated that the information on evidence only referenced the practice guideline and did not provided sufficient information to evaluate the quality, quantity and consistency of the evidence. AMA-PCPI indicated that they additional information to the Committee. • It is unclear why men are included. The developer suggested it might be a theoretical caution by the pharmaceutical company and the product warnings as they could not find any data to support it.

0401: Hepatitis C: Counseling regarding risk of alcohol consumption
Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: Percentage of patients aged 18 years and older with a diagnosis of hepatitis C who were counseled regarding the risks of alcohol consumption at least once within the 12 month reporting period

Numerator Statement: Patients who were counseled regarding the risks of alcohol consumption at least once within the 12 month reporting period

Denominator Statement: All patients aged 18 years and older with a diagnosis of hepatitis C

Exclusions: None

Adjustment/Stratification: No risk adjustment or risk stratification We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:** American Association for the Study of Liver Diseases, American Gastroenterological Association Institute

0401: Hepatitis C: Counseling regarding risk of alcohol consumption

IMPLEMENTATION COMMENTS:

- IDSA: IDSA supports endorsement of measure #0401. This is a critical measure, and when possible, counseling of patients should include spouse, significant other or partner to improve outcome.
  - Developer response: Thank you for your comment.
- HIV Medicine Association: We recommend that NQF carefully review the feasibility of this and other behavioral health measures, as performance of such screenings is not readily captured by most electronic health record (EHR) systems without a manual chart review.
  - Developer response: Based on the testing results, this measure was found to be feasible for implementation. We have provided testing data, which includes a comparison of E.H.R. automated reports to visual inspection of the medical record and had a kappa score of 0.47. This score shows that the measure is reliable and shows that the information can be accurately collected in both an electronic health record and a paper medical record.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the **Committee Reviewers:** Rekha Murthy; Curtis Collins; Tiffany Osborn; Mary Blank; Thomas File; Mohamad Fakih   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-2

1a. Impact: H-4; M-2; L-0; I-0; 1b. Performance Gap: H-4; M-2; L-0; I-0

Rationale:

- 78.1% mean performance rate of TIN/NPIs
- 1a. Hep C has a high impact of patient health, but the measure addresses counseling for alcohol consumption- does not equate to cessation
- 1b. Gap in performance and disparities in care

1c. Evidence (*based on decision logic*): Y-4; N-2   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-6

Quantity: H-1; M-4; L-0; I-1; Quality: H-1; M-4; L-1; I-0; Consistency: H-2; M-2; L-1; I-1

Rationale:

- Application cites 3 references (one guideline) showing effect of high alcohol intacke in HCV-infected patients; Guideline recommendation
- Clinical practice evidence. Recommendations based on guidelines. consistency not addressed.

0401: Hepatitis C: Counseling regarding risk of alcohol consumption
2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-6; N-0 **2a. Reliability:** H-3; M-3; L-0; I-0 **2b. Validity:** H-3; M-3; L-0; I-0 <u>Rationale:</u> • Reliability: moderate agreement when comparing to EHR • Validity: face validity was done using expert opinion consensus with 11/13 strongly agreed with validity
3. Usability: H-2; M-4; L-0; I-0 (*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*) <u>Rationale:</u> • Required manual inspection of medical record • 3a. Used as 1st step for public reporting • 3b. Gather data on physician performance: my concern that documentation of counseling does not equate change in patient behavior
4. Feasibility: H-2; M-4; L-0; I-0 (*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*) <u>Rationale:</u> • Required manual inspection of medical record • 4a. Yes, found in EHR • 4b. Yes, needs to be in EMR • 4c. No data given • 4d. Yes, unless no EMR available
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-4; N-1 <u>Rationale:</u> • My concern is the impact of this measure on patient care. Documentation of counseling may not equate counseling of the risk with alcohol consumption. In addition, the measure may not be feasible unless EHR is implemented in all practices <u>Additional Comments/Questions</u>: • Based on benefit/harm, this process of care should be performed
Workgroup Call Summary
Importance to Measure and Report • The WG indicated that the information on evidence only referenced the practice guideline and did not provided sufficient information to evaluate the quality, quantity and consistency of the evidence. • The cited guideline indicated the guideline was graded IIB– "Usefulness/efficacy is less well established by evidence/opinion" and Level of Evidence = C – "Only consensus opinions of experts, case studies, or standard of care • Members of the WG specified that this documentation measure does not seem to be a measure of quality. Counseling is not the same as cessation of alcohol use. The WG was struggling to identify how this measure would improve the overall impact on the patient.   o The developer indicated that there are a small number of studies that suggest counseling does aid in the reduction of alcohol consumption. • Clinicians must evaluate for alcohol use before prescribing anti-viral therapy. **Scientific Acceptability of Measure Properties** • A WG member was concerned with the kappa score of 0.47. **Feasibility** • The measure requires manual chart abstraction.

0298: Central line bundle compliance

Status: Maintenance, Original Endorsement: Nov 15, 2007

Description: Percentage of intensive care patients with central lines for whom all
elements of the central line bundle are documented and in place.
The central line bundle elements include:
•Hand hygiene
•Maximal barrier precautions upon insertion
•Chlorhexidine skin antisepsis
•Optimal catheter site selection, with avoidance of the femoral vein for central venous access in patients 18 years and older
•Daily review of line necessity with prompt removal of unnecessary lines

Numerator Statement: Number of intensive care patients with central lines for whom all elements of the central line bundle are documented and in place.
The central line bundle elements include:
• Hand hygiene
• Maximal barrier precautions upon insertion
• Chlorhexidine skin antisepsis
• Optimal catheter site selection, with avoidance of the femoral vein for central venous access in patients 18 years and older
• Daily review of line necessity with prompt removal of unnecessary lines

Denominator Statement: Total number of intensive care patients with central lines on the day of sample.

Exclusions: Exclude patients less than 18 years of age at the date of ICU admission and patients outside the intensive care unit and patients whose lines were not placed in the intensive care unit

Adjustment/Stratification: No risk adjustment or risk stratification NA NA

Level of Analysis: Facility

Type of Measure: Composite

Data Source: Paper Medical Records

Measure Steward: Institute for Healthcare Improvement

0298: Central line bundle compliance

IMPLEMENTATION COMMENTS:

- IDSA: The IDSA supports endorsement of measure #0298. We recognize the importance of bundling activities together to improve infection rates. If total compliance is not 100%, it would be useful to know if certain components are consistently omitted, and if so, infection rates associated with each type of omission could be determined. Hand hygiene, maximal barrier precautions upon central line insertion, and chlorhexidine skin antisepsis are essential and there should be no exceptions to compliance. Optimal catheter site selection is less conclusive. Some studies have not demonstrated higher infection rates with jugular as compared with subclavian insertion sites. There are noninfectious risks associated with subclavian placement that are less common with jugular sites. It is unclear that a requirement for documentation of the reason for not using the subclavian vein improves care. While it is important to complete daily review of necessity of the line, after years of this measure, it is prudent to evaluate if there is any impact- to see if hospitals with less than 100% compliance have higher BSI rates, or longer catheter-days. Grading 100% compliance and individual compliance would enrich the data. In addition, other measures require daily assessment of necessity, (foley urinary catheters). If bundled, it would be interesting to see if a dedicated discussion of all lines would result in earlier removal.
  - Developer response: Thank you for your comments. The central line bundle was developed as an all/none measure, with the recommendation that hospitals assess compliance with individual bundle elements to identify opportunities for improving compliance. IHI has found that hospitals begin to demonstrate improvement in outcomes (central line-associated BSI) when they reliably provide all five components of the bundle. I also agree with your comments re: daily review of necessity of the lines and promoting a "dedicated discussion" of all lines, urinary catherter. Many hospitals have demonstrated success in removal of unnecessary central lines, urinary catheters by incorporating into overall review during daily rounding.

- University of Texas-MD Anderson Cancer Center: On behalf of The University of Texas MD Anderson Cancer Center, we appreciate the opportunity to provide feedback on this measure. Have you used this measure for any of the following purposes? 1. Quality improvement (internal to your specific organization) **Yes** 2. Quality improvement with benchmarking (external benchmarking to multiple organizations) **No** 3. Professional certification or recognition program **No** 4. Regulatory and accreditation programs **Yes – TJC NPSG.07.04.01** 5. Payment program **No** 6. Public reporting **No** 7. Other **N/A** Have you encountered challenges while implementing this measure? 1. Problems with measure specifications **No** 2. Challenges in obtaining the necessary data **No** 3. Lack of harmonization with related measures (same measure focus or same target population) **No** 4. Difficulties interpreting or explaining the performance results **No** 5. Difficulties obtaining reliable and valid comparisons of performance **We use an internal benchmark** 6. Unintended consequences **No** Do you have suggestions for how this measure could be improved? **No** Is there a better measure that should be considered in place of this measure? **No** Should this measure receive endorsement for another three years? **Yes**
  - o Developer response: Thank you for your comments. The central line bundle was developed as an all/none measure, with the recommendation that hospitals assess compliance with individual bundle elements to achieve high reliability with the bundle. The focus of the central line bundle as a process (vs. outcome) measure was for internal improvement, as you reflect you have used the measure.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Tiffany Osborn; Rekha Murthy; Curtis Collins; Mary Blank; Thomas File; Mohamad Fakih   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): **Y-3; N-3**
1a. Impact: H-5; M-1; L-0; I-0; **1b. Performance Gap:** H-4; M-0; L-1; I-1
Rationale:

- Common use of central lines; high morbidity/mortality of infection
- 1a. CLABSI is associated with significant morbidity and mortality. It has been addressed through multiple national efforts and public reporting of the final outcome CLABSI. The developer suggests using the elements of the IHI central line bundle to evaluate compliance with proper insertion method of central line. The measure addresses placement in the intensive care units at the hospital only. Note that there are already mechanisms to evaluate final outcomes for CLABSI.
- 1b. Performance gap exists between hospitals, but I am not sure that "reporting the use of the checklist for central line placement and extracting medical record documentation of evaluation for need" is associated with reduced infection rates. The central line checklist is based on team work between the one placing the line and the assistant (ICU nurse). The ICU nurse documents while with the operator whether s/he was compliant with the proper insertion steps (hand hygiene, complete barrier, chlohexidine use). It is essential to make sure that documentation in records is a true reflection of the procedure elements. Obtaining information from records regarding choosing catheter site and daily evaluation for need are more difficult to extract.

1c. Evidence (*based on decision logic*): **Y-5; N-1**   IF a **Health Outcome**, rationale supports: **Y-1; N-0; NA-5**
Quantity: H-4; M-0; L-0; I-1; **Quality:** H-2; M-3; L-0; I-0; **Consistency:** H-3; M-1; L-0; I-1
Rationale:

- Not answered
- Significant data and guidelines. Actually more recent than just listed in the worksheet
- 1c. The IDSA guidelines support "bundling" the steps to reduce risk for infection and incorporating them in the process. The SHEA compendium supports using a checklist with a healthcare worker observing and is able to stop the operator if there is a breach is aseptic technique. Documentation is addressed by the suggested measure, but not the process of directly observing the process of placement. The quality of the different components of the measures is high, although many of the quality improvement efforts were implemented with the help of changes in culture of the healthcare workers (eg, CUSP or other high reliability efforts).

2. Scientific Acceptability of Measure Properties (*based on decision logic*): **Y-3; N-3**
2a. Reliability: H-0; M-4; L-2; I-0; **2b. Validity:** H-1; M-1; L-3; I-1
Rationale:

0298: Central line bundle compliance

- Validity tested in study design
- 2a. The reliability of the measure: described as all or none indicator. For example, if documentation is not present for evaluation of need for one day out of the whole duration of use, then the result would be not compliant. I am not aware of studies that show a relation between not documenting daily need and increased risk of CLABSI. The reliability depends on the chart review process. This is a review of paper records and not based on EHR. Developer states that no reliability and validity testing has been done.

3. Usability: H-3; M-2; L-1; I-0

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale:

- Studies cited
- 3a. The measure is understandable but difficult to collect the individual items because of paper record review. With no data on reliability and validity, it would be hard to explain usability.

4. Feasibility: H-0; M-5; L-1; I-0

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c. Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

Rationale:

- Relies on voluntary reporting, may or may not be in electronic medical record, may or may not need to be abstracted from the chart.
- The measure is important, but reliability of data collection for the composite make it difficult. Inter-rater reliability is not great
- Muliple element measure
- 4. Most of the elements are routinely documented in paper records. No EHR plans. The data is susceptible to errors and documentation bias. It would be interesting to review the compliance with the checklist correct steps as far as documentation. The value of obtaining the results of compliance with the bundle would be based on the baseline compliance which is not available.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-4; N-1

Rationale:

- Infections are of significant consequence and prevention is paramount. Elements based on study data and graded guideline recommendations
- The developer has not provided adequate information on reliability, validity, feasibility and usability.

Additional Comments/Questions:

Workgroup Discussion

Importance to Measure and Report

- Great concept
- Just a documentation measure – not necessarily reflecting what is happening at the bed side
- Quite a bit of evidence for the components but less evidence for the entire bundle
- The CLABSI outcome measure already exists – why do we need the process measure?

Scientific Acceptability of Measure Properties

- No exclusions for emergent central line access
- "Avoidance of femoral site" needs definition; what if femoral is the only site left?
  - Developer reports that avoidance does not mean never – there needs to be a rationale for using the site
  - No information on validity

Usability

- Not much information provided
- A WG member thought that this measure had been dropped by CMS from the HHS Action Plan

Feasibility

- Daily checks and documentation is very hard

0412: HIV/AIDS: Hepatitis B Vaccination

Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: Percentage of patients aged six months and older with a diagnosis of HIV/AIDS, who have received at least one hepatitis B vaccination, or who have documented immunity
Numerator Statement: Patients who have received at least one injection of hepatitis B vaccination, or who have documented immunity
Denominator Statement: All patients aged six months and older with a diagnosis of HIV/AIDS, with at least two visits in the measurement year, with at least 90 days in between each visit
Definition of "Medical Visit" - any visit with a health care professional who provides routine primary care for the patient with HIV/AIDS (may be but is not limited to a primary care clinician, ob/gyn, pediatrician, infectious diseases specialist)
Exclusions: None.
Adjustment/Stratification: No risk adjustment or risk stratification N/A N/A
Level of Analysis: Clinician : Group/Practice, Clinician : Individual
Type of Measure: Process
Data Source: Administrative claims, Electronic Clinical Data
Measure Steward: National Committee for Quality Assurance **Other organizations:** Physician Performance Measures (Measures) and related data specifications have been developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement™ (the Consortium) and the National Committee for Quality Assurance (NCQA). The Health Resources and Services Administration (HRSA) and the Infectious Diseases Society of America also participated in the development of this measure.

0412: HIV/AIDS: Hepatitis B vaccination

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Aaron Milstone; Jeffrey Beal; Kathleen Brady; Sue Elam; Tom Giordano   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-3; N-2
1a. Impact: H-3; M-2; L-0; I-0; 1b. Performance Gap: H-1; M-2; L-1; I-1
Rationale:
- There are no data that there is room for improvement.
- 1b.2 No data regarding total # in survey. Stats still show opportunity for improvement.
- 1b.3 Source cited for data lacks detailed information
- 1b.4 No disparities noted
- Impact data are mix of US and international. Performance gap data not restricted to HIV-infected. No disparities data.

1c. Evidence (*based on decision logic*): Y-5; N-0   IF a **Health Outcome**, rationale supports: Y-2; N-0; NA-3
Quantity: H-3; M-2; L-0; I-0; Quality: H-0; M-5; L-0; I-0; Consistency: H-0; M-5; L-0; I-0
Rationale:
- This is based on need for Hep B vaccine, not supporting 1 dose of vaccine as a measure
- Numerous studies have been conducted on this topic in both adults,adolescents and children. The data support the recommendation primarily come from a well-designed clinical trial without randomization and from cohort and case-controlled studies. The findings, however, are consistent in favoring Hep B vaccination in this population.
- >5 studies cited
- Vaccine protects against HBV acquisition, not transmission. Data on risk of acquisition after HIV recognized are not presented.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-2
2a. Reliability: H-0; M-2; L-2; I-1; 2b. Validity: H-0; M-2; L-3; I-0
Rationale:
- Low correlation of EMR and chart review. 1 vaccine is poor measure of protection
- Separation between automated and manual performance too great on validity testing. Not sure I agree with one HBV meeting the need for those with immune suppression. No control for CD4 at time of vaccination.

0412: HIV/AIDS: Hepatitis B Vaccination
2a and 2b. Percentage difference between automated and manual results was 59% which is outside of acceptable norms.   2b1. The evidence is for completion of Hep B vaccination and not just one dose of vaccine.   2b2.3 Concern re: % agreement b/w automated vs manual and coding criteria   - Reads like should be addressed every 12 mos, but really only need vaccine once. Very low reliability and validity from EHR.

3. Usability: H-0; M-3; L-2; I-0

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale:
- 1 vaccine is poor indicator of protection
- A similar Hep vaccination measure is currently in use by HAB. Unclear,however, why the HAB measure is not being considered for use instead of this one.
- 3a.1 No expected date is noted for disclosure

4. Feasibility: H-1; M-3; L-1; I-0

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

Rationale:
- The required data elementa are routinely generated but not always consistently reported/retrievable from EHR based on the data provided.
- 4c Coding errors?
- Problems accessing old paper records and interpreting lab results not addressed.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-2; N-3

Rationale:
- Poor validity.
- Reliability and validity are low.

Additional Comments/Questions:

Workgroup Call Discussion

Importance to Measure and Report
- The WG had questions about the number of people measured and whether there was opportunity for improvement.
- No information on disparities was provided.
- The data provided does not support the measure focus (i.e. how many patients with HIV received only one hepatitis B injection). Is one injection enough?

Scientific Acceptability of Measure Properties
- The WG noted a significant difference in results between electronic health records versus manual calculation.
- The WG questioned why the developers are looking to see if only one vaccination was administered.
  - The developers explained that in order to capture if all three doses were administered, they would have to change the denominator to allow for a longer look back period, since there is timing between injections to factor. The developers were concerned they would lose patients in follow up.
  - The developer also noted they were trying to harmonize with the hepatitis C measure (AMA-PCPI).
  - Looking at the first injection provides data that the provider is starting to initiate proper care.
- The data was not available for the WG to see how often providers are using the "patient has documented immunity to hepatitis B" using category II codes.
  - The developer explained the CPT procedure code being captured is whether the vaccine was administered or previously administered. There's another code for patient has documented immunity to hepatitis B and another code for hepatitis B vaccine series previously received. Documented immunity might also be found in a paper medical record as a note.
- The WG was concerned with the numerator specifications where this is performed every 12 months.
  - One member pointed out that the vaccination is performed once. If a person has been an HIV carrier for a decade, they may have received HBV series 10 years ago; would a physician keep indicating every 12 months that the person is immune?

0404: HIV/AIDS: CD4 cell count or percentage performed
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: Percentage of patients aged six months and older with a diagnosis of HIV/AIDS, with a CD4 cell count or percentage performed at least once every 6 months
Numerator Statement: Patients with a CD4 cell count or percentage performed at least once every 6 months
Denominator Statement: All patients aged 6 months and older with a diagnosis of HIV/AIDS, who had at least two medical visits during the measurement year, with at least 90 days between each visit
Exclusions: None
Adjustment/Stratification: No risk adjustment or risk stratification N/A N/A
Level of Analysis: Clinician : Group/Practice, Clinician : Individual
Type of Measure: Process
Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Laboratory
Measure Steward: National Committee for Quality Assurance **Other organizations:** Physician Performance Measures (Measures) and related data specifications have been developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement™ (the Consortium) and the National Committee for Quality Assurance (NCQA). The Health Resources and Services Administration (HRSA) and the Infectious Diseases Society of America also participated in the development of this measure.

0404: HIV/AIDS: CD4 cell count or percentage performed
Implementation Comments:
• HIV Medicine Association: We support continued NQF endorsement of this measure, but urge harmonization and alignment of the NQF-endorsed measure across and among federal agencies.
o Developer response: Thank you for your support.

Notes

Workgroup Preliminary Evaluations
The following evaluation ratings and comments are from the Committee Reviewers: Aaron Milstone; Jeffrey Beal; Kathleen A. Brady, MD; Sue Elam; Tom Giordano *(comments separated by bullets)*
Importance to Measure and Report (*based on decision logic*): Y-4; N-1
1a. Impact: H-5; M-0; L-0; I-0; **1b. Performance Gap:** H-2; M-3; L-0; I-0
Rationale:
• 1.b. Not high because no information on disparities in care by populaton groups can be completely captured by current reporting mechanisms.
• 1b.2 Less than optimal performance with stats noted
• 1b.4 No disparities noted?
1c. Evidence (*based on decision logic*): Y-4; N-1   IF a **Health Outcome**, rationale supports: Y-2; N-0; NA-2
Quantity: H-4; M-1; L-0; I-0; **Quality:** H-2; M-3; L-0; I-0; **Consistency:** H-3; M-1; L-0; I-1
Rationale:
• Moderate on Quantity as most not RCTs.
• CD4 measurement is critical step in process body of evidence >5 studies with large # of pts
2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-5; N-0
2a. Reliability: H-1; M-4; L-0; I-0; **2b. Validity:** H-0; M-5; L-0; I-0
Rationale:
• Unclear how this performs using other health records
• 2a1 Well defined; implementation straight forward
• 2b2 Validity testing and face validity discrepancies?
3. Usability: H-4; M-1; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

0404: HIV/AIDS: CD4 cell count or percentage performed

Rationale:
- 3a.1 CMS PQRS use presently and previous 3 years
- 3a.2 7% improvement
- 3b.1 QI tool that is being used

4. Feasibility: <mark>H-1; M-4; L-0; I-0</mark>
(4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)
Rationale:
- 4c Coding errors?

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-5; N-0</mark>
Rationale:
Additional Comments/Questions:

<mark>**Workgroup Discussion**</mark>

Importance to Measure and Report
- For this measure, the average performance rate in PQRS for eligible professional was 76.8 percent in 2009 and 83.9 percent in 2010, and the developer felt that those numbers indicate there was a gap in care with room for improvement.
- There was no information on disparities.

Scientific Acceptability of Measure Properties
- This measure was tested at the measure score level; as such the highest rating can be moderate.
- The WG was concerned that the testing for this measure used a small sample of clinics in the same geographic area, which used the same EHR. Geographic variation of testing cites would have made the results more valid.
- The WG noted a significant difference between electronic health records versus manual calculation; noting there was confusion about the numerator criteria (i.e. which codes to use, timing of the CD4 count)
  - The developer explained that when the measure was tested, they were using a CD4/CD8 ratio code that was included in the list of codes. They have since removed the CD8 ratio because it is not an appropriate CD4 test to do.
  - The developer also noted that the confusion regarding timing may be because the measure specifies 'within six months' whether that means within each six-month's period of the year or if it meant every six months is not clear. The developer is open to making that more clear.
- The measure is looking for either a claim for the CD4 test or a note in the medical record. The measure does not require the results be present.
  - The developer noted that if this measure was ever developed as an eMeasure, there would be more opportunity to look for results.
  - One WG member noted that because this is a process measure, the initial process was completed for this measure. While receiving the testing results and acting on the testing results are clinically important, verifying that the test results are in the medical record is much more difficult.
  - The title of the measure says performed, and therefore the data should support that it's the result that the measure is capturing not just whether or not that lab was ordered.
    - The Developer will check in with their expert panel to modify the measure title to reflect looking for documentation of the results.
- The WG noted the ambiguity in the numerator statement, and suggested that it should read like the denominator statement, saying **CD4 cell count or percentage performed at least 90 days apart**. Doing this would also address the inconsistency in manual versus electronic health records calculation.
  - The developer stated that the intent would be every six months in between because that's more aligned with the guideline. However, they are willing to take that back to their expert panel and work on some clarifying language.

0405: HIV/AIDS: Pneumocystis jiroveci pneumonia (PCP) prophylaxis
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: Percentage of patients aged 6 weeks or older with a diagnosis of HIV/AIDS, who were prescribed Pneumocystis jiroveci pneumonia (PCP) prophylaxis

Numerator Statement: Numerator 1: Patients who were prescribed *Pneumocystis jiroveci* pneumonia (PCP) prophylaxis within 3 months of CD4 count below 200 cells/mm3

Numerator 2: Patients who were prescribed *Pneumocystis jiroveci* pneumonia (PCP) prophylaxis within 3 months of CD4 count below 500 cells/mm3 or a CD4 percentage below 15%

Numerator 3: Patients who were prescribed *Pneumocystis jiroveci* pneumonia (PCP) prophylaxis at the time of HIV diagnosis

Report a rate for each numerator (e.g., Numerator 1/Denominator 1, etc.) and a total rate (Total Numerator/Total Denominator)

Denominator Statement: Denominator 1. All patients aged 6 years and older with a diagnosis of HIV/AIDS and a CD4 count below 200 cells/mm3, who had at least two visits during the measurement year, with at least 90 days in between each visit; and,

Denominator 2. All patients aged 1 through 5 years of age with a diagnosis of HIV/AIDS and a CD4 count below 500 cells/mm3 or a CD4 percentage below 15%, who had at least two visits during the measurement year, with at least 90 days in between each visit; and,

Denominator 3. All patients aged 6 weeks through 12 months with a diagnosis of HIV, who had at least two visits during the measurement year, with at least 90 days in between each visit

Total denominator: The sum of the three denominators

Exclusions: Denominator 1 Exclusion: Patient did not receive PCP prophylaxis because there was a CD4 count above 200 cells/mm3 during the three months after a CD4 count below 200 cells/mm3

Denominator 2 Exclusion: Patient did not receive PCP prophylaxis because there was a CD4 count above 500 cells/mm3 or CD4 percentage above 15% during the three months after a CD4 count below 500 cells/mm3 or CD4 percentage below 15%

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Pharmacy

Measure Steward: National Committee for Quality Assurance **Other organizations:** Physician Performance Measures (Measures) and related data specifications have been developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement™ (the Consortium) and the National Committee for Quality Assurance (NCQA). The Health Resources and Services Administration (HRSA) and the Infectious Diseases Society of America also participated in the development of this measure.

0405: HIV/AIDS: *Pneumocystis jiroveci* pneumonia (PCP) prophylaxis

IMPLEMENTATION COMMENTS:
- HIV Medicine Association: Change to percentage of patients (regardless of age) with CD4+ counts fewer than 200. We support continued NQF endorsement with this change.
  - Developer response: We have convened an expert panel to provide us with guidance about aligning this measure with current guidelines.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the **Committee Reviewers:** Michael Farber; Kalpana Ramiah; Adam Thompson; Ed Septimus   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-1

1a. Impact: H-3; M-0; L-0; I-0; **1b. Performance Gap:** H-2; M-1; L-0; I-0

<u>Rationale:</u>
- 1a.3 - Evidence cited indicates that PCP prophylaxis is one of the most cost effective treatments for persons with HIV. Without prophylaxis PLWH are at increased risk for PCP which pre-treatement was a significant cause of mortality.
- 1b.2 - CMS PQRS evidence demonstrates a performance gap is present for the prescription of PCP to PLWH with specified CD4

0405: HIV/AIDS: Pneumocystis jiroveci pneumonia (PCP) prophylaxis

counts
- 1b.4 - Disparity data is not presented though studies have shown that there is disparity is prescription of PCP by race/ethnicity…though this evidence is not cited it should be considered.

1c. Evidence (*based on decision logic*): Y-2; N-1   **IF** a **Health Outcome**, rationale supports: Y-1; N-1; NA-1
Quantity: H-2; M-0; L-0; I-1; **Quality:** H-1; M-2; L-0; I-0; **Consistency:** H-2; M-1; L-0; I-0
Rationale:
- 1c.5 - While the previous evidence cites the specific CDC guidelines for PCP, the section does not specify that the studies are specific to PCP ... only that it comes from CDC guidelines.
- 1c.6 - Quality of the studies differs depending on the age of the patient - adults is moderate with randomized controlled trial, adolescents, less so with non RCT.
- 1c.7-1c.8 - Consistency statement is that consistency is high yet net benfit is not specified to PCP but rather OIs in general - yet with the body of evidence being insufficient to rate, consistency statement and lack of specificity to PCP earned a moderate rating

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-2; N-1
2a. Reliability: H-2; M-0; L-0; I-1; **2b. Validity:** H-1; M-1; L-0; I-1
Rationale:
- 2a.1.25 Measure steward cites validity testing as reliability testing yet the data source does not cite that EHRs were used ... which according to interpretation of guidance would require separate reliability testing.   Rating will be consistent with validity testing since submitted for both.
- 2b.2.1 - Data source is EHR yet not cited in the data source for the measure; is the data sample representative?
- 2b.2.2 - Measure validity score is good yet the face validity includes no description of whether it was sytematically assessed. Validity is thus rated insufficient.
- 2b.3.1 - Threat to validity is discussed and addressed as new exclusion (patients who had a blip in CD4 only)
- 2b.5 - Performance results from CMS PQRS reported and shown that there is still a performance gap

3. Usability: H-2; M-1; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:
- 3a.1 - Measure used in CMS PQRS
- 3a.2 - Measure outcome showed substantial (14%) improvement in clinical care meaning it was meaningful for public reporting (similar measure used in HIVQual program)
- 3a.3 - Being considered for CMS EHR incentive program
- 3b.1 - 3b.2 - Similar measure used in HIVQual Program - CMS PQRS showed 14% improvement in site using measure.

4. Feasibility: H-1; M-2; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:
- 4a.1 - Data generate during care delivery
- 4b.2 - All elements are in electronic sources
- 4c - Prescription does not ensure patient actually recieved the medication or took it…some may think simple prescrition meets care and not follow-up on whether the medication was actually acquired or taken .
- 4d. Data collection statement is not specific to this measure, rated as insufficient.
- Some difficulties with capturing data elements in the EMR

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-2; N-1
Rationale:
- With the quantity of the body of evidence and both reliability/validity rated as insufficient, currently the measure cannot meet the criteria for endorsement.

Additional Comments/Questions:

Workgroup Discussion

Importance to Measure and Report
- WG members indicated the process of care utilizing PCP prophylaxis is perhaps closer to a real patient level outcome in terms of its potential impact and so it scores highly on importance and evidence.

Scientific Acceptability of Measure Properties
- WG members noted that whether measure is being captured using electronic medical record or by visual of inspection of the

0405: HIV/AIDS: Pneumocystis jiroveci pneumonia (PCP) prophylaxis
medical records there is a high degree of correlation with automated calculation as well as the manual calculation of performance. • The specifications for this measure include denominator exclusions which make the measure more accurate by aligning with current guidelines. • This measure is stratified by age. **Usability** • Some WG members thought that this measure would be difficult to capture as it requires extensive programming not available to all providers. • Requires hand counting the denominator, due to the number of exclusions. • When available, the eMeasure the measure implementer will be able to review the measure as the XML, however, to convert from EHR to XML data will require local programming capabilities and commitment to be able to make this into a feasible electronic measure which is an important resource consideration for many programs. **Feasibility** • Difficult to calculate this measure, if doing manual chart review. • The developer is under contract with the ONC to develop eMeasures for consideration of use and the meaningful use program

2083: Prescription of HIV antiretroviral therapy
Status: New Submission
Description: Percentage of patients, regardless of age, with a diagnosis of HIV prescribed antiretroviral therapy for the treatment of HIV infection during the measurement year
Numerator Statement: Number of patients from the denominator prescribed HIV antiretroviral therapy during the measurement year.
Denominator Statement: Number of patients, regardless of age, with a diagnosis of HIV with at least one medical visit in the measurement year
Exclusions: There are no patient exclusions.
Adjustment/Stratification: No risk adjustment or risk stratification Not applicable
Level of Analysis: Population : Community, Population : County or City, Facility, Clinician : Group/Practice, Population : National, Population : Regional, Population : State
Type of Measure: Process
Data Source: Electronic Clinical Data : Electronic Health Record, Paper Medical Records, Electronic Clinical Data : Pharmacy
Measure Steward: Health Resources and Services Administration - HIV/AIDS Bureau **Other organizations:** The Centers for Disease Control

2083: Prescription of HIV antiretroviral therapy
Notes

The following evaluation ratings and comments are from the Committee Reviewers: Aaron Milstone; Jeffrey Beal; Kathleen Brady; Sue Elam; Tom Giordano   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-3; N-2
1a. Impact: H-5; M-0; L-0; I-0; **1b. Performance Gap:** H-3; M-1; L-0; I-1
Rationale:
- Increased treatment=decrease VL=decrease transmission,M&M
- 1b. Considerable variation/less than optimal performance across providers and populations + disparities discussion
- Didn't show deficiencies in ART prescription

1c. Evidence (*based on decision logic*): Y-4; N-1   **IF a Health Outcome**, rationale supports: Y-1; N-0; NA-4
Quantity: H-3; M-1; L-0; I-1; **Quality:** H-1; M-4; L-0; I-0; **Consistency:** H-3; M-2; L-0; I-0
Rationale:
- Insufficient data to require treatment of all patients with HIV.This does not provide exclusions for people that refuse treatment or are not given treatment for various reasons.
- Data somewhat limited for persons with CD4 counts over 500.
- > 5 studies; RCT,meta analysis, observational. Body of evidence supports guidelines
- 1c. Consistency ?>500 CD4=treatment
- Data for CD4>500 is less strong.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-5; N-0
2a. Reliability: H-1; M-4; L-0; I-0; **2b. Validity:** H-0; M-4; L-1; I-0
Rationale:
- What about exceptions that are not accounted for
- 2a. Good sampling-testing details explained
- 2b. Validity--face threats to validity??

3. Usability: H-3; M-1; L-0; I-1
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:
- Process for reporting is not outlined inclusion is pending for CMS

4. Feasibility: H-2; M-3; L-0; I-0
(4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified

2083: Prescription of HIV antiretroviral therapy

4d. Data collection strategy can be implemented)
Rationale: • List of ARVs has some potential for difficulties in data collection • 4c Not addressed
Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-4; N-1</mark> **Rationale:** • Not enough data to support in the population presented. concern about face validity given some gray areas that will not be accounted for in data capture • Rx of ARV associated with reduced morbidity/mortality/transmission Rx of ARV has chance for development of viral resistance;short and longterm toxicities/affects/medication SEs Cost associated with Rx
Additional Comments/Questions:
<mark>**Workgroup Call Discussion**</mark>

Importance to Measure and Report
• One WG member noted that while we recognize the importance of this clinically, the current guidelines that are presented for pediatric population, in children less than 5 years of age, state for those that are asymptomatic with a CD4 percentage rate of 25 percent or an HIV RNA less than a hundred thousand copies, a physician should consider treatment. If you have a group in which the recommendation nationally is to consider treatment. This does not in any way taken as a new account pediatric population.
   o The developer noted that this measure is not a clinic or facility measure, but the measure should also look at the population, at the state or regional level.
   o This performance measure is not supposed to be zero or 100 percent. For quality improvement purposes, the developer is aware that some patients will not make it into the numerator. HRSA has a very stringent view on pediatric population as there is always consistent monitoring, and more in-depth monitoring on that population to see where they are in their HIV disease and seeing if they're eligible to be put on to antiretroviral therapy.
• The WG noted that the evidence is somewhat limited for persons whose CD4 count is greater than 500.
   o This measure is limited to persons whose CD4 cell count is greater than 500 and the intent of the for treating over 500 is not so much to say that you can't do it but rather the measure is intended to be interpreted as one "may treat;" that everyone is permitted to treat over that CD4 cell count. In large jurisdictions including San Francisco and New York City health officials are implementing policy that all patients diagnosed with HIV infection regardless of CD4 cell count are being treated.
   o The developer states that if a limitation is applied, the utility of the measure would be limited because the measure would exclude large fractions of people. It's true that most patients who are identified with HIV infection in the United States, they are identified at a CD4 cell count under 500; however it is the intent of HRSA to identify people very early in the course of their disease.
   o The measure could be used in an early intervention clinic.
      ▪ The WG noted the developer's response; however members are still uncomfortable being held to a standard backed by limited evidence.
Feasibility
• The list of ARVs has some potential for difficulties in data collection. The WG would prefer outlining the medications that should not be used together, rather than the approach of an abstractor trying to review regiments to see if they are consistent with the guideline.

0406: HIV/AIDS: Adolescent and adult patients who are prescribed potent antiretroviral therapy

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: Percentage of patients with a diagnosis of HIV/AIDS, with at least two visits during the measurement year, with at least 90 days between each visit: aged 13 years and older who have a history of a CD4 count =500 cells/mm3; aged 13 years and older who have a history of an AIDS-defining illness, regardless of CD4 count; or who are pregnant, regardless of CD4 count or age, who were prescribed potent antiretroviral therapy

Numerator Statement: Patients who were prescribed potent antiretroviral* therapy

*Potent antiretroviral therapy is described as any antiretroviral therapy that has demonstrated optimal efficacy and results in durable suppression of HIV as shown by prior clinical trials

Denominator Statement: A. All patients aged 13 years and older with a diagnosis of HIV/AIDS, with at least two medical visits during the measurement year, with at least 90 days between each visit, who have a history of a CD4 count less than or equal to 500 cells/mm3; and

B. B. All patients aged 13 years and older with a diagnosis of HIV/AIDS, with at least two medical visits during the measurement year, with at least 90 days between each visit, who have a history an AIDS-defining illness**, regardless of CD4 count; and

C. All patients with a diagnosis of HIV/AIDS, with at least two medical visits during the measurement year, with at least 90 days between each visit, who are pregnant, regardless of CD4 count or age.

**The most commonly used case definition for AIDS is the 1993 Revised Surveillance Case Definition from the CDC. It includes: Candidiasis of bronchi, trachea, or lungs; candidiasis, esophageal; cervical cancer, invasive; coccidiodomycosis, disseminated or extrapulmonary; cryptococcosis, extrapulmonary; crytosporidiosis, chronic intestinal (greater than 1 month's duration); cytomegalovirus disease (other than liver, spleen, or nodes); cytomegalovirus retinitis (with loss of vision); encephalopathy, HIV-related; herpes simplex: chronic ulcer(s) (greater than 1 month's duration); or bronchitis, pneumonitis, or esophagitis; histoplasmosis, disseminated or extrapulmonary; isosporiasis, chronic intestinal (greater than 1 month's duration); Kaposi's sarcoma; lymphoma, Burkitt's (or equivalent term); lymphoma, immunoblastic (or equivalent term); lymphoma, primary, of brain; mycobacterium avium complex or M. kansasii, disseminated or extrapulmonary; mycobacterium tuberculosis, any site (pulmonary or extrapulmonary); mycobacterium, other species or unidentified species, disseminated or extrapulmonary; pneumocystis carinii pneumonia; pneumonia, recurrent; progressive multifocal leukoencephalopathy; salmonella septicemia, recurrent; toxoplasmosis of brain; wasting syndrome due to HIV. (Aberg, 2009; National Center for Infectious Diseases Division of HIV/AIDS)

Definition of "Medical Visit" - any visit with a health care professional who provides routine primary care for the patient with HIV/AIDS (may be but is not limited to a primary care clinician, ob/gyn, pediatrician, infectious diseases specialist)

Note: For potent antiretroviral therapy recommendations refer to current DHHS guidelines available at www.aids.gov

Aberg JA, Kaplan JE, Libman H, Emmanuel P, Anderson JR, Stone VE, Oleske JM, Currier JS, Gallant JE; HIV Medicine Association of the Infectious Diseases Society of America. Primary care guidelines for the management of persons infected with human immunodeficiency virus: 2009 update by the HIV medicine Association of the Infectious Diseases Society of America. Clin Infect Dis. 2009 Sep 1;49(5):651-81. Available at http://www.uphs.upenn.edu/bugdrug/antibiotic_manual/idsahivprimarycare2009.pdf. Accessed May 25, 2012.

National Center for Infectious Diseases Division of HIV/AIDS. 1993 Revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. MMRW Recomm Rep. 1992;41(RR-17):1-19.

Exclusions: None

Adjustment/Stratification: No risk adjustment or risk stratification N/A N/A

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Pharmacy

Measure Steward: National Committee for Quality Assurance **Other organizations:** Physician Performance Measures (Measures) and related data specifications have been developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement™ (the Consortium) and the National Committee for Quality Assurance (NCQA). The Health Resources and Services Administration (HRSA) and the Infectious Diseases Society of America also participated in the development of this measure.

0406: HIV/AIDS: Adolescent and adult patients who are prescribed potent antiretroviral therapy

Implementation Comments:

- HIV Medicine Association: In keeping with current clinical practice guidelines, we recommend deletion of qualifications to measure percentage of all patients prescribed antiretroviral therapy, such that the measure should read as follows: Percentage of patients with a diagnosis of HIV/AIDS with at least two visits during the measurement year, with at least 60 days – or whichever interval is selected for the medical visit measure -- between each visit who were prescribed potent antiretroviral therapy.
  - Developer response: We convened an expert panel to provide us with guidance about aligning this measure with current guidelines. The expert panel did not support deleting qualifications from the denominator of this measure. Based on the current treatment guidelines and evidence, we will be measuring whether the following populations received potent ART: patients 13 and older with at least two visits, at least 90 days apart, who have a history of a CD4 count below or equal to 500 cells/mm3; patients 13 and older with at least two visits, at least 90 days apart, who have a history of an AIDS-defining illness; and patients, regardless of age, who are pregnant.

Notes

The following evaluation ratings and comments are from the Committee Reviewers: Aaron Milstone; Jeffrey Beal; Kathleen Brady; Sue Elam; Tom Giordano   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-1
1a. Impact: H-5; M-0; L-0; I-0; 1b. Performance Gap: H-2; M-2; L-1; I-0
Rationale:
- 1b.2 Variability in stats from 2 data sources
- 1b.4 No disparities data
- Gardner is a review (not systematic); 97% meet standard in small sample; HIVQUAL data are on viral suppression, not prescription of ART; no disparities data.

1c. Evidence (*based on decision logic*): Y-5; N-0   IF a **Health Outcome**, rationale supports: Y-1; N-0; NA-4
Quantity: H-5; M-0; L-0; I-0; Quality: H-4; M-1; L-0; I-0; Consistency: H-5; M-0; L-0; I-0
Rationale:
- Quantity of studies of studies included is over 45 with 20 RCTs.
- Quality of evidence is moderate given that RCTs available for persons with CD4 counts <350 and only cohort data available for 350-500.
- Consistency of evidence is consistent in direction and magnitude.
- >5 studies with large patient population;rct

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-5; N-0
2a. Reliability: H-1; M-4; L-0; I-0; 2b. Validity: H-1; M-4; L-0; I-0
Rationale:
- Unclear on how well "potent" is defined. Unclear how this will perform using other EMRs outside of test set
- 2a is moderate because both data elements and measure score were not assessed.
- No disparities noted

3. Usability: H-3; M-2; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:
- CMS PQRS--small # providers cited over 2 yr period? HIVQUAL-US--larger gap in care

4. Feasibility: H-2; M-2; L-1; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:
- Coding errors

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-5; N-0
Rationale:
- Meets all criteria for endorsement.
Additional Comments/Questions:
- No disparities information.

Importance to Measure and Report
- This applies to patients 13 years and older with a CD4 count, less than or equal to 500.
- The measure has a performance of 97 percent, and the WG questioned whether there is a performance gap.
  - Additionally the data presented on performance gap just were on viral suppression not on and did not show a performance gap in prescription of ART.

0406: HIV/AIDS: Adolescent and adult patients who are prescribed potent antiretroviral therapy

- No disparities data was presented.

Scientific Acceptability of Measure Properties
- The WG noted that the numerator statement was vague and questioned how "potent ART" is determined by the measure.
  - o The developer explained that potent ART is captured using a Category II Code that the provider uses to test that the patient is on potent ART. The developers refer providers to the HRSA guidelines about what treatment patients with HIV should be on. The developer uses this approach because treatment guidelines frequently change and it is not feasible to maintain the list of acceptable drugs and acceptable combinations, and be able to have a Category II Code that ties to those acceptable and recommended combinations.
    - ▪ The measure uses Category II Code for whether potent antiretroviral therapy is prescribed.
  - o One member inquired as to whether this could only apply to treatment-naïve patients, but the developer stated that because we are referring patients to the treatment guidelines, they can look to the guideline to explain what treatments should be used for experienced as well as naïve patients.
- This measure is looking at whether or not providers are documenting that a patient is on a potent antiretroviral therapy.
- The WG questioned how well the Category II code is documented, however those data were not available to the developer, as the measure is used for PQRS.

Usability
- It is time consuming for the provider to figure out if the therapy qualifies for this measure. A physician could have a patient who is highly resistant to a highly potent regimen that maybe in this group of OK drugs to prescribe. Yet, that is not the ultimate goal of what we want. We want the patients to be undetectable for a large number of reasons.
  - o This measure seems to capture a code that a patient is on a specific drug.
    - ▪ The developer explained that the intent of the measure is to capture patients who are not being prescribed ART when they should be.
- The WG was concerned, citing the difficulty of manual abstracting and programming for this measure because of the vast number of options for drug combinations. One member cited that if the key is to know if the patient has undetectable viral load or not, then why waste the effort on measurement of the therapy. If they're not undetectable on viral load, it doesn't necessarily mean that they're not on an antiviral regimen. It is more than likely, if they are undetectable on viral load, they are on an antiretroviral regimen.
  - o The developer explained there are differences in who was prescribed the antiretroviral therapy or who's not prescribed the antiretroviral therapy, and who doesn't become undetectable. There are two different types of intervention: (1) that someone may not be undetectable because their doctor didn't prescribe it or (2) it may be undetectable because the patient didn't take their medicine.
  - o One member stated, that it's two different steps in the process, and noted there is value to measuring the steps independently as well as a global performance.
- There might be difficulty calculated the numerator and denominator because it is difficult to figure out who's got AIDS and who's got HIV just based on their ICD-9 codes There are ICD-10 codes for AIDS versus HIV.

0407: HIV/AIDS: HIV RNA control after six months of potent antiretroviral therapy

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: Percentage of patients aged 13 years and older with a diagnosis of HIV/AIDS, who had at least two medical visits during the measurement year, with at least 90 days between each visit, who are receiving potent antiretroviral therapy*, who have a viral load <200 copies/mL after at least 6 months of potent antiretroviral therapy*

*Potent antiretroviral therapy is described as any antiretroviral therapy that has demonstrated optimal efficacy and results in durable suppression of HIV as shown by prior clinical trials

Numerator Statement: Patients with an HIV viral load <200 copies/mL

Denominator Statement: All patients aged 13 years or older with a diagnosis of HIV/AIDS, with at least two visits in the measurement year, with at least 90 days between each visit, who received potent antiretroviral therapy* for at least 6 months

Definition of "Medical Visit" - any visit with a health care professional who provides routine primary care for the patient with HIV/AIDS (may be but is not limited to a primary care clinician, ob/gyn, pediatrician, infectious diseases specialist)

*Potent antiretroviral therapy is described as any antiretroviral therapy that has demonstrated optimal efficacy and results in durable suppression of HIV as shown by prior clinical trials

Exclusions: None

Adjustment/Stratification: No risk adjustment or risk stratification N/A N/A

Level of Analysis: Population : County or City, Clinician : Group/Practice, Clinician : Individual

Type of Measure: Outcome

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Pharmacy

Measure Steward: National Committee for Quality Assurance **Other organizations:** Physician Performance Measures (Measures) and related data specifications have been developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement™ (the Consortium) and the National Committee for Quality Assurance (NCQA). The Health Resources and Services Administration (HRSA) and the Infectious Diseases Society of America also participated in the development of this measure.

0407: HIV/AIDS: HIV RNA control after six months of potent antiretroviral therapy

Implementation Comments:

- HIV Medicine Association: We recommend that this measure be updated as follows: Percentage of patients aged 13 years and older with a diagnosis of HIV/AIDS who had at least two medical visits during the measurement year, with at least 60 days – or whichever interval is selected for the medical visit measure -- between each visit, who are receiving potent antiretroviral therapy**, who have a viral load below limits of quantification* after at least 6 months of potent antiretroviral therapy.

  *Using laboratory cutoff level for reference laboratory used by that clinic.
  ** Potent antiretroviral therapy is described as any antiretroviral therapy that has demonstrated optimal efficacy and results in durable suppression of HIV as shown by prior clinical trials.

  Rationale: There are now sufficient medications to achieve viral control that this measure should strive to simply capture the percentage of patients in care and on ART who are virally suppressed.
  We support continued NQF endorsement of this measure, with the above update.
  - Developer response: We agree with removing the plan of care component from this measure. After convening an expert panel to review this measure, we will submit the following measure to NQF: Percentage of patients aged 13 years and older with a diagnosis of HIV/AIDS, who had at least two medical visits during the measurement year, with at least 90 days between each visit, who are receiving potent antiretroviral therapy, who have a viral load <200 copies/mL after at least 6 months of potent antiretroviral therapy.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the **Committee Reviewers:** Aaron Milstone; Jeffrey Beal; Kathleen Brady; Sue Elam; Tom Giordano *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-5; N-0
1a. Impact: H-5; M-0; L-0; I-0; 1b. Performance Gap: H-2; M-3; L-0; I-0

0407: HIV/AIDS: HIV RNA control after six months of potent antiretroviral therapy

Rationale:
- Gaps in data peformance are provided. No disparities data available.
- 1b.2 Stats indicate significant gap
- 1b.4 No discussion of disparities
- Single source of data for performance gap; no disparities data.

1c. Evidence (*based on decision logic*): <mark>Y-5; N-0</mark>   IF a **Health Outcome**, rationale supports: <mark>Y-1; N-0; NA-4</mark>
Quantity: H-5; M-0; L-0; I-0; **Quality:** H-3; M-2; L-0; I-0; **Consistency:** H-3; M-2; L-0; I-0
Rationale:
- Unclear why it does not include kids >13 years old
- 1c.5 >5 studies with large pt population
- 1c.6 Quality reflected in summary statement only
- No consistent or clear data for VL<200 cut point.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): <mark>Y-4; N-1</mark>
2a. Reliability: H-1; M-3; L-1; I-0; **2b. Validity:** H-2; M-3; L-0; I-0
Rationale:
- Unclear how this will perform in other sites using different EMRs. also unclear how   "potent" is defined and how hard this will be to pull out of EMRs.
- 2a and 2b. EHR testing at the measure level only and not at data element level.
- Threats to validity not addressed
- Reliability low-measure does not specify which VL to assess in measurment period.

3. Usability: <mark>H-4; M-1; L-0; I-0</mark>
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:
- 3a Current use CMS
- 3b Current use shows need for improvement

4. Feasibility: <mark>H-2; M-3; L-0; I-0</mark>
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:
- Measure already in operational use.
- ART for 6 months may be challenging to measure; blips can have VL>200.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-5; N-0</mark>
Rationale:
- Meets all criteria for endorsement.
- With modification and clarification

Additional Comments/Questions:
- Do not agree with HIVMA's recommendation to use the laboratory cutoff for quantification of viral loads instead of <200. This would add considerable burden and complexity to the measure. The ability to obtain this information from EHR would need to be tested first.

<mark>**Workgroup Call Discussion**</mark>

Importance to Measure and Report
- It is unclear why children under 13 are excluded.
  - The developer stated they would discuss that with their expert panel about lowering the age limit. The developer suspected that one reason the measure might be specified this way is that it is linked to another potent ART measure and the denominator for that population is greater than 13.

Scientific Acceptability of Measure Properties
- There was ambiguity regarding which viral load you're supposed to use. After six months of potent ART, could have more than one viral load in the 12-month measurement period. If so, do you use the last viral load? Do you use any viral load? Do you use the lowest viral load?
- The WG noted that the numerator statement was vague and questioned how potent ART is captured by the measure.
  - The developer explained that potent ART is captured using a Category II Code that the provider uses to test that the patient is on potent ART. The developers refer providers to the HRSA guidelines about what treatment patients with HIV

should be on. The developer uses this approach because treatment guidelines frequently change and it is not feasible to maintain the list of acceptable drugs and acceptable combinations, and be able to have a Category II Code that ties to those acceptable and recommended combinations.

- They use Category II Code for whether potent antiretroviral therapy is prescribed.

Usability

- This measure is currently in use by CMS.

Feasibility

- Moderate feasibility due to an unclear definition of viral suppression and how to deal with a "blip."
- The measure handles missing data (i.e. not having the lab value) by counting it at a numerator miss if the patient record does not have that their RNA results.

2082: HIV viral load suppression
Status: New Submission
Description: Percentage of patients, regardless of age, with a diagnosis of HIV with a HIV viral load less than 200 copies/mL at last HIV viral load test during the measurement year
Numerator Statement: Number of patients in the denominator with a HIV viral load less than 200 copies/mL at last HIV viral load test during the measurement year
Denominator Statement: Number of patients, regardless of age, with a diagnosis of HIV with at least one medical visit in the measurement year
Exclusions: There are no patient exclusions.
Adjustment/Stratification: No risk adjustment or risk stratification Not applicable Not applicable
Level of Analysis: Facility, Clinician : Group/Practice
Type of Measure: Outcome
Data Source: Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory, Paper Medical Records
Measure Steward: Health Resources and Services Administration - HIV/AIDS Bureau **Other organizations:** The Centers for Disease Control

2082: HIV viral load suppression
Notes

Workgroup Preliminary Evaluations
The following evaluation ratings and comments are from the Committee Reviewers: Aaron Milstone; Jeffrey Beal; Kathleen Brady; Sue Elam; Tom Giordano   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-1
1a. Impact: H-5; M-0; L-0; I-0; **1b. Performance Gap:** H-4; M-1; L-0; I-0
Rationale:
- 1b. Considerable variation/less than optimal performance across providers and population discussion on disparities

1c. Evidence (*based on decision logic*): Y-4; N-1   IF a **Health Outcome**, rationale supports: Y-1; N-0; NA-3
Quantity: H-3; M-1; L-0; I-1; **Quality:** H-3; M-2; L-0; I-0; **Consistency:** H-3; M-2; L-0; I-0
Rationale:
- No data to support recommendations in all children
- Signifcant relationship b/w viral load suppression and reduced morbidity, mortality and HIV transmission >5 studies, RCT,meta analysis, observatonal high quality of evidence sited for recommendations on treatment to reduce disease and death moderate quality re reduction in transmission
- Moderate because data weaker for CD4>500.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-4; N-0
2a. Reliability: H-1; M-4; L-0; I-0; **2b. Validity:** H-1; M-3; L-0; I-0
Rationale:
- 2a and 2b. Reliability and validity assessed only at the measure level.
- 2a. Good sampling, well defined testing data
- 2b. Validity--face potential threats???

3. Usability: H-3; M-1; L-0; I-1
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:
- No data to suipport in all children
- 3b. Meaningful; easy to access for QI significant relationship b/w viral load suppression and reduced M&M; transmission

4. Feasibility: H-2; M-3; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:
- Did not address 4c.

2082: HIV viral load suppression

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-4; N-1

Rationale:
- This will systematically ding providers that accept higher viral loads in kids that are asymptomatic and have higher viral loads
- Collection and analysis of HIV viral load suppression has clear cut implications for survival and improved quality of life. The measure is a routine standard of care. Suppression of viral load has tremendous implications for prevention of HIV infections. Identifying disparities found in this measure can assist with strategies to lessen the disparity that will have meaningful outcome for survival.
- Meets all endorsement critieria.

Additional Comments/Questions:

Workgroup Call Discussion

Importance to Measure and Report
- While there is a movement towards treating all children with HIV, there are providers who do not treat asymptomatic high viral loads and high CD4 counts. This measure does not account for this.
- This measure is a snapshot of a treatment process whose end results is suppressing viral loads. The WG noted however, that when you do not examine each process step you can miss important reason why someone might not be suppressed (i.e. they refused ART or are not adherent). These reasons maybe important from a clinical standpoint, and are important for measuring performance at the individual level.
    - The developer responded by noting that this performance measure is not supposed to be zero or 100 percent. For quality improvement purposes, the developer is aware that some patients will not make it into the numerator.
        - Additionally, the measure developer, in preparation for eMeasures, has drawn back from including a large of exclusion criteria in order to simplify the measures and address the lack of structured data elements, for these reasons, within electronic health records.

2081: Newly enrolled in medical care

Status: New Submission

Description: Percentage of patients, regardless of age, with a diagnosis of HIV who were newly enrolled and had a medical visit in each of the 4-month periods in the measurement year

Numerator Statement: Number of patients in the denominator who had at least one medical visit in each 4-month period of the measurement year (Measurement year is a consecutive 12-month period of time.).

Denominator Statement: Number of patients, regardless of age, with a diagnosis of HIV who was newly enrolled with a medical provider and had at least one medical visit in the first 4 months of the measurement year. "Newly enrolled" patients are those who are: newly diagnosed with HIV and new to medical care; patients new to medical care (previously diagnosed with HIV and never received HIV medical care); patients who transferred their medical care to your organization; or patients returning to medical care after a 2-year absence (patients re-engaged by the same organization).

Exclusions: Patients who died at any time during the measurement year.

Adjustment/Stratification: No risk adjustment or risk stratification   Not applicable

Level of Analysis: Facility, Clinician : Group/Practice

Type of Measure: Process

Data Source: Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: Health Resources and Services Administration - HIV/AIDS Bureau **Other organizations:** The Centers for Disease Control

2081: Newly enrolled in medical care

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Kalpana Ramiah; Adam Thompson; Ed Septimus *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-1

1a. Impact: H-2; M-1; L-0; I-0; **1b. Performance Gap:** H-2; M-1; L-0; I-0

Rationale:

- 1a.3 - Evidence cited indicates "each 'no show' clinic visited conveyed a 17% increased risk of delayed viral load suppression." Evidence showed consistency of visits in the first year showed a relationship to survival. Evidence showed CD4 counts significantly greater "among those with optimal retention." *compared to the the medical visit measure, the evidence for this measure was specific to the medical visit and not simply to the measurement of the clinical indicators - this showed the relationship between the two ... lack of retention showed higher rate of virologic failure.
- 1b.1 - Suboptimal retention suggested higher mortality rates and retention showed greater mean increase in baseline CD4
- 1b.2 - Evidence submitted shows that there is a decline in medical visit frequency over time with a higher rate having 2 or more visits during a 6 month interval but less over a 18-24 month and a larger drop when evaluated over 3-5 years showing a steady decline in retention over time. Evidence showed almost a 1/3 of patients experience a gap in care and less than half met the HRSA requirement for all years in outpatient care.
- 1b.4 - Disparity data summarized showing females, racial minorities, and patients lacking private health insurance were "significantly more likely to fail to establish care." Patients in care differed by demographic and risk group.   Evidence is same as other submitted measures by this steward - how does it differ - could the same information be gleaned from the frequency measure?

1c. Evidence (*based on decision logic*): Y-2; N-1   IF a **Health Outcome**, rationale supports: Y-2; N-0; NA-1

Quantity: H-1; M-2; L-0; I-0; **Quality:** H-0; M-2; L-0; I-1; **Consistency:** H-2; M-1; L-0; I-0

Rationale:

- 1c.2-3 Type of evidence was on selected studies rather than the entire body of evidence, used clinical practice guidelines as well and a systematic review of evidence (other than within guideline development)
- 1c.5 - Evidence for Quantity is specific to retention and entry to care - though could also be anti-retroviral adherence. Cites that 2 studies were specific to monitoring retention to care. DHHS citation were specific to frequency of lab values (which would be a "medical visit" according to measure specifications).

2081: Newly enrolled in medical care

- 1c.6 - Measure steward does not discuss in detail the quality of evidence - rated insufficient
- 1c.7 - Consistency statement is supported with the data from the studies including confidence intervals and increased risk for death data
- 1c.8 - Medical visit benefits seems to include counseling which could be with a non-medical provider (risk reduction/case management/adherence) - does the evidence support that this is considered a medical visit? No harms identified; cost of visits could be high depending on the frequency.
- 1c.12 - Evidence graded using modified GRADE system - grade AI – AIII
- 1c.24 - No other guidelines addressing retention to care (DHHS guideline graded A using non specified grading)

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-0
2a. Reliability: H-1; M-2; L-0; I-0; **2b. Validity:** H-0; M-3; L-0; I-0
<u>Rationale:</u>

- 2a1.1 - Measure specifications are for a one year period with a visit in each of the quarters - shows retention to care in the first year through increased frequency of the visit(s) - responds to critique of how does this measure differ from frequency - newly enrolled would need greater frequency to establish the linkage with the new care site - makes sense.
- 2a1.7 - Denominator details require a date of death to ensure exclusion - means clinicians can't simply guess if patient disappeared that death was the cause, must document
- 2a.1.25 - Data Source - EHR indicated as data source (and paper records) no reliability testing needed yet supplied.
- 2a.2.1 - Data sample is presented as representative of facility types, geographic divisions, demographics, and insurance status and coverage types. Three calendar years of data are presented with increasing number of patients over the three year sample. Patient characteristics are presented with breakdowns indicated by the Importance to Measure disparity statement. Data presented are representative of 2009 CDC surveillance data for PLWH
- 2a.2.2 - Analytic method is presented and discussed in detail - signal to noise ratio is calculated with ICC of <1 (.0080776) with 13 sites reporting and pediatric sites combined, associated CI are calculated and presented and appear acceptable
- 2b.1.1 - Measure specifications are aligned with studies indicating poor retention, increased rates of missed medical visits and gaps lead to poorer health outcomes.
- 2b.2.1 - Data sample is presented and is sample presented also for reliability (same comments apply for representative statements) 2b.2.2 - Analytic method is described in detail as face validity established and systematically assessed by technical work group using modified Delphi Process as well as feedback from providers using the measure related to feasibility and usability of the measures.
- 2b.2.3 - Results show that the technical work group deemed the measure important, usable, and feasible.
- 2b.3 - No discussion of threats to validity
- 2b5 - Data sample shown with ability to capture disparity data and performance results presented including min, max, mean and quartiles
- 2b6 - Measure not tested on multiple data sources - highest rating is moderate
- 2c.1 - Data is stratified for disparities

3. Usability: H-1; M-2; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
<u>Rationale:</u>

- 3.1 - Measure is currently being used for quality improvement with benchmarking external to reporting organization
- 3a.1 - Measure is used in national quality improvement project, technical work group saw utility in public reporting, measure (upon endorsement) will seek inclusion in Stage 3 of CMS EHR incentive program and PQRS
- 3a.2 - Measure is able to distinguish difference in performance and data presented with consistency and top/bottom performers - QI project participants reported that measure is meaningful to management of *their* patient populations and is understandable by patients and providers – reporting sites showed a sharp decline in number of sites reporting during the fourth measurement period – why? Are the providers finding this measure less useful?
- 3b.1 - Used currently for QI in national project
- 3b.2 - Measure is unique in that it addresses recent literature pertaining to retention to care and established relationship (in evidence presented) between frequency and risk of death and other health outcomes   Frequency measure rated high, this measure is being used less broadly - why? Usability here rated moderate in light of the comparison.

4. Feasibility: H-2; M-1; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

2081: Newly enrolled in medical care

Rationale:

- 4a. Data generated during delivery of care
- 4b. Data available in electronic claims
- 4c. No susceptibility to errors discussed - would paper charts match with EHR data? does the removal of incarcerated and transferred to care skew the performance? While cited in 4d as difficult to capture - might it be worth identifying how this affects the performance outcome and interpretation given the high rates of incarceration in persons of color who also show high incidence and disease burden?
- 4d - Collection strategy changed based on feedback from providers to eliminate certain exclusions - cited that data variance is a result of performance and not differences in data availability.
- Developers indicate no known inaccuracies!

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-3; N-0</mark>
Rationale:

- Measure differs from frequency in that it measures a different time period (1 year vs. 24 months) as well a different visit frequency (4 per year versus 2 per year). This would align with newly enrolled to care needing greater frequency in the first year to establish the linkage to improve retention.   Ability to stratify for disparities is good; allows for providers to respond to beliefs about demographic disparities in newly enrolled demographics and identify any persons who are newly enrolled and may not have a good initial linkage.
- Overlaps with 2079

Additional Comments/Questions:

<mark>Workgroup Discussion</mark>

Importance to Measure and Report

- This measure overall did not have the same degree of evidence as the other visit measures.
- WG members asked the developer about the every four-month's visit -- one of the points that some papers make is that early on, more frequent visits might be needed and later, visits as infrequently as every six months are OK.
    - o HRSA responded it was the paper by Mike Mugavero in 2009 where he looked at missed visits among establishing initial outpatient ambulatory medical care, but wanted to verify for the in-person meeting.
- Another question focused on inclusion in the denominator individuals who also transferred care. And if location has been stable, receiving care for 10 years and they switch care, is the expectation of this measure with that individual would then have four medical hits over the next year if they had transferred it and what is the justification for that? If someone was regularly in care and they switched doctors, and they're used to going once every six months because they're virally suppressed and adherent, that this measure would either low perform for the providers because the individual doesn't want to be seen that frequently or the patient might feel that there's a burden on their medical visits by having to go more frequently than they had thereby causing what would be like a part of a negative relationship that's formed, what their expectations as they come to see their doctor more? Also, individuals who may have high co-pays with their insurance– if the providers really push the performance measure, could be an undue burden not on the system or the data but on the patients themselves.
    - o The developer noted that a person who has transferred would be included and it would be a visit every four months or three in a year. "*We came from the perspective of this is a person who's going to be mentoring potentially a new healthcare system, not just a new provider. So, there was the potential risk that the client, may have difficulty navigating a new system*".

Scientific Acceptability of Measure Properties

- The group as a whole thought this measure was less reliable and less valid by the numbers than the other two measures regarding retention - not enough studies to support this and also is the reliability of what you're actually measuring that it would reflect what you're trying to determine and that is that people are being seen early in the course of their identification. However, the reliability and validity was sufficient.

Feasibility

- Easy to measure visits with various data systems.
- Impact of incarcerated patients not being able to follow-up in care not accounted for in the measure.

2079: Medical visit frequency

Status: New Submission

Description: Percentage of patients, regardless of age, with a diagnosis of HIV who had at least one medical visit in each 6-month period of the 24-month measurement period with a minimum of 60 days between medical visits

Numerator Statement: Number of patients in the denominator who had at least one medical visit in each 6-month period of the 24-month measurement period with a minimum of 60 days between first medical visit in the prior 6-month period and the last medical visit in the subsequent 6-month period. (Measurement period is a consecutive 24-month period of time.)

Denominator Statement: Number of patients, regardless of age, with a diagnosis of HIV with at least one medical visit in the first 6 months of the 24-month measurement period.

Exclusions: Patients who died at any time during the 24-month measurement period.

Adjustment/Stratification: No risk adjustment or risk stratification Not applicable Not applicable

Level of Analysis: Facility, Clinician : Group/Practice

Type of Measure: Process

Data Source: Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: Health Resources and Services Administration - HIV/AIDS Bureau **Other organizations:** The Center For Disease Control and Prevention

2079: Medical visit frequency

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the **Committee Reviewers:** Kalpana Ramiah; Adam Thompson; Ed Septimus *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-3; N-1
1a. Impact: H-4; M-0; L-0; I-0; 1b. Performance Gap: H-3; M-1; L-0; I-0
Rationale:

- 1a.3 - Evidence cited indicates "each 'no show' clinic visited conveyed a 17% increased risk of delayed viral load suppression." Evidence showed consistency of visits in the first year showed a relationship to survival. Evidence showed CD4 counts significantly greater "among those with optimal retention." *compared to the the medical visit measure, the evidence for this measure was specific to the medical visit and not simply to the measurement of the clinical indicators - this showed the relationship between the two ... lack of retention showed higher rate of virologic failure.
- 1b.1 - Suboptimal retention suggested higher mortality rates and retention showed greater mean increase in baseline CD4
- 1b.2 - Evidence submitted shows that there is a decline in medical visit frequency over time with a higher rate having 2 or more visits during a 6 month interval but less over a 18-24 month and a larger drop when evaluated over 3-5 years showing a steady decline in retention over time. Evidence showed almost a 1/3 of patients experience a gap in care and less than half met the HRSA requirement for all years in outpatient care.
- 1b.4 - Disparity data summarized showing females, racial minorities, and patients lacking private health insurance were "significantly more likely to fail to establish care." Patients in care differed by demographic and risk group.

1c. Evidence (*based on decision logic*): Y-3; N-1   IF a **Health Outcome**, rationale supports: Y-3; N-0; NA-1
Quantity: H-0; M-4; L-0; I-0; **Quality:** H-0; M-3; L-0; I-1; **Consistency:** H-3; M-1; L-0; I-0
Rationale:

- 1c.2-3 Type of evidence was on selected studies rather than the entire body of evidence, used clinical practice guidelines as well and a systematic review of evidence (other than within guideline development)
- 1c.5 - Evidence for Quantity is specific to retention and entry to care - though could also be anti-retroviral adherence. Cites that 2 studies were specific to monitoring retention to care. DHHS citation were specific to frequency of lab values (which would be a "medical visit" according to measure specifications).
- 1c.6 - Measure steward does not discuss in detail the quality of evidence - rated insufficient
- 1c.7 - Consistency statement is supported with the data from the studies including confidence intervals and increased risk for death

2079: Medical visit frequency

data

- 1c.8 - medical visit benefits seems to include counseling which could be with a non-medical provider (risk reduction/case management/adherence) - does the evidence support that this is considered a medical visit? No harms identified; cost of visits could be high depending on the frequency.
- 1c.12 - evidence graded using modified GRADE system - grade AI – AIII 1.2.24 - no other guidelines addressing retention to care (DHHS guideline graded A using non specified grading)

2. Scientific Acceptability of Measure Properties (*based on decision logic*): <mark>Y-4; N-0</mark>
2a. Reliability: H-2; M-2; L-0; I-0; **2b. Validity: H-0**; M-4; L-0; I-0
<u>Rationale</u>:

- 2a.1 - Numerator time window specifies a 24 month period of time measuring actual retention - aligns with evidence from Importance to Measure
- 2a1.7 - Denominator details require a date of death to ensure exclusion - means clinicians can't simply guess if patient disappeared that death was the cause, must document
- 2a.1.25 - Data Source - EHR indicated as data source (and paper records) no reliability testing needed yet supplied.
- 2a.2.1 - Data sample is presented as representative of facility types, geographic divisions, demographics, and insurance status and coverage types. Three calendar years of data are presented with increasing number of patients over the three year sample. Patient characteristics are presented with breakdowns indicated by the Importance to Measure disparity statement. Data presented are representative of 2009 CDC surveillance data for PLWH
- 2a.2.2 - Analytic method is presented and discussed in detail - signal to noise ratio is calculated with ICC of <1 (.0365188) with 13 sites reporting and pediatric sites combined, associated CI are calculated and presented and appear acceptable
- 2b.1.1 - Measure specifications are aligned with studies indicating poor retention, increased rates of missed medical visits and gaps lead to poorer health outcomes.
- 2b.2.1 - Data sample is presented and is sample presented also for reliability (same comments apply for representative statements)
- 2b.2.2 - Analytic method is described in detail as face validity established and systematically assessed by technical work group using modified Delphi Process as well as feedback from providers using the measure related to feasibility and usability of the measures.
- 2b.2.3 - Results show that the technical work group deemed the measure important, usable, and feasible.
- 2b.3 - No discussion of threats to validity
- 2b5 - Data sample shown with ability to capture disparity data and performance results presented including min, max, mean and quartiles
- 2b6 - Measure not tested on multiple data sources - highest rating is moderate
- 2c.1 - data is stratified for disparities
- Testing not performed for excluded patients-also incarceration is not listed as exclusion

3. Usability: <mark>H-2; M-1; L-0; I-0</mark>
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
<u>Rationale</u>:

- 3.1 - Measure is currently being used for quality improvement with benchmarking external to reporting organization
- 3a.1 - Measure is used in national quality improvement project, technical work group saw utility in public reporting, measure put forward to fill measurement gap for retention - implementation date set for later in 2012
- 3a.2 - Measure is able to distinguish difference in performance and data presented with consistency in site reporting and top/bottom performers - QI project participants reported that measure is meaningful to management of *their* patient populations and is understandable by patients and providers
- 3b.1 - Used currently for QI in national project as well as intended use in several large govt agencies, OHA, VA by end of 2012
- 3b.2 - Measure is unique in that it addresses recent literature pertaining to retention to care and established relationship (in evidence presented) between frequency and risk of death and other health outcomes

4. Feasibility: <mark>H-2; M-1; L-0; I-0</mark>
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c. Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
<u>Rationale</u>:

- 4a. Data generated during delivery of care
- 4b. Data available in electronic claims
- 4c. No susceptibility to errors discussed - would paper charts match with EHR data? does the removal of incarcerated and

2079: Medical visit frequency
transferred to care skew the performance? While cited in 4d as difficult to capture - might it be worth identifying how this affects the performance outcome and interpretation given the high rates of incarceration in persons of color who also show high incidence and disease burden? • 4d - Collection strategy changed based on feedback from providers to eliminate certain exclusions - cited that data variance is a result of performance and not differences in data availability.
Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-3; N-0</mark> Rationale: • This measure seems to be more aligned with recent focus on retention as the process rather than adherence to medical visit. The 24 month window allows for a broader analysis of the patient retention. Definition of medical visit is not given and the measure could be strengthened with a more specific definition. Quality of the body of evidence is insufficient for rating however with reliability and validity testing presented, rated as passing criteria for endorsement. Additional Comments/Questions: • Similar to 2080-? combine
<mark>Workgroup Discussion</mark>
Importance to Measure and Report • There is a significant opportunity for improvement, HRSA cites that overall under 50 percent of the patients meet the criteria for medical visit frequency. • The WG questioned the reasoning behind lack of exclusions regarding reasons for not coming back to visit i.e. for persons who are incarcerated, have relocated, or transferred service.   o The developers explained that it is difficult to provide structured data related to those three elements. HRSA is also responding to direction from ONC requesting simplified measure specifications that are readily used in EHRs. Exclusions such as these are difficult to find in structured data fields. The developers hope to use the results in order to set benchmarks (re: CMS meaningful use project). HRSA will clarify this point on their measure submission form. **Scientific Acceptability of Measure Properties** • This measure was tested at the measure score level; as such the highest rating can be moderate. **Usability** • These data are being used across several major government agencies, private healthcare providers, as well as the 12 cities project. • This measure has been put forward to fill the measurement gap for retention of HIV care. **Feasibility** • The data are available in electronic health records

2080: Gap in medical visits
Status: New Submission
Description: Percentage of patients, regardless of age, with a diagnosis of HIV who did not have a medical visit in the last 6 months of the measurement year
Numerator Statement: Number of patients in the denominator who did not have a medical visit in the last 6 months of the measurement year (Measurement year is a consecutive 12-month period of time).
Denominator Statement: Number of patients, regardless of age, with a diagnosis of HIV who had at least one medical visit in the first 6 months of the measurement year. (The measurement year can be any consecutive 12-month period.)
Exclusions: Patients who died at any time during the measurement year.
Adjustment/Stratification: No risk adjustment or risk stratification Not applicable Not applicable
Level of Analysis: Facility, Clinician : Group/Practice
Type of Measure: Process
Data Source: Electronic Clinical Data : Electronic Health Record, Paper Medical Records
Measure Steward: Health Resources and Services Administration-HIV/AIDS Bureau **Other organizations:** The Centers For Disease Control

2080: Gap in medical visits
Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Michael Farber; Kalpana Ramiah; Adam Thompson; Ed Septimus   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): <mark>Y-3; N-1</mark>
1a. Impact: H-4; M-0; L-0; I-0; **1b. Performance Gap:** H-4; M-0; L-0; I-0
Rationale:
- 1a.3 - Evidence cited indicates "each 'no show' clinic visited conveyed a 17% increased risk of delayed viral load suppression." Evidence showed consistency of visits in the first year showed a relationship to survival. Evidence showed CD4 counts significantly greater "among those with optimal retention."
- 1b.1 – This measure differs from the other measures in that it identifies in the denominator who is in need of a medical visit in the second 6 months – patients lists can be generated to determine who needs follow-up
- 1b.2 - Evidence submitted shows that there is a decline in medical visit frequency over time with a higher rate having 2 or more visits during a 6 month interval but less over a 18-24 month and a larger drop when evaluated over 3-5 years showing a steady decline in retention over time. Evidence showed almost a 1/3 of patients experience a gap in care and less than half met the HRSA requirement for all years in outpatient care.
- 1b.4 - Disparity data summarized showing females, racial minorities, and patients lacking private health insurance were "significantly more likely to fail to establish care."   Patients in care differed by demographic and risk group.

1c. Evidence (*based on decision logic*): <mark>Y-3; N-1</mark>   IF a **Health Outcome**, rationale supports: <mark>Y-3; N-0; NA-1</mark>
Quantity: H-2; M-2; L-0; I-0; **Quality:** H-1; M-2; L-0; I-1; **Consistency:** H-4; M-0; L-0; I-0
Rationale:
- 1c.1 – Sufficient evidence to suggest that lower rates of significant gaps between medical visits is associated with patient quality of care outcomes
- 1c.2-3 Type of evidence was on selected studies rather than the entire body of evidence, used clinical practice guidelines as well and a systematic review of evidence (other than within guideline development)
- 1c.5 - Evidence for Quantity is specific to retention and entry to care - though could also be anti-retroviral adherence. Cites that 2 studies were specific to monitoring retention to care. DHHS citation were specific to frequency of lab values (which would be a "medical visit" according to measure specifications).
- 1c.6 - Measure steward does not discuss in detail the quality of evidence - rated insufficient
- 1c.7 - Consistency statement is supported with the data from the studies including confidence intervals and increased risk for death data
- 1c.8 - Medical visit benefits seems to include counseling which could be with a non-medical provider (risk reduction/case

2080: Gap in medical visits

management/adherence) - does the evidence support that this is considered a medical visit? No harms identified; cost of visits could be high depending on the frequency.

- 1c.12 - Evidence graded using modified GRADE system - grade AI – AIII
- 1.c.24 - No other guidelines addressing retention to care (DHHS guideline graded A using non specified grading)

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-4; N-0
2a. Reliability: H-1; M-3; L-0; I-0; 2b. Validity: H-1; M-3; L-0; I-0
Rationale:

- 2a1.1 - Measure specifications are for a one year period with a visit in the first and second 6 months – this measure to me is the most comparable to the currently endorsed measure but the measure specifications do not define medical visit – this may be a strength of this measure to identify when and where the patients are not showing a hit across the entire medical system – which medical visit is not defined – perhaps the totality of data showing exposure to any point of care in the system (prescriber or not) is useful in determining who is really "lost to care" – interesting.
- 2a1.7 - Denominator details require a date of death to ensure exclusion - means clinicians can't simply guess if patient disappeared that death was the cause, must document
- 2a.1.25 - Data Source - EHR indicated as data source (and paper records) no reliability testing needed yet supplied.
- 2a.2.1 - Data sample is presented as representative of facility types, geographic divisions, demographics, and insurance status and coverage types. Three calendar years of data are presented with increasing number of patients over the three year sample. Patient characteristics are presented with breakdowns indicated by the Importance to Measure disparity statement. Data presented are representative of 2009 CDC surveillance data for PLWH
- 2a.2.2 - Analytic method is presented and discussed in detail - signal to noise ratio is calculated with ICC of <1 (.032194) with 13 sites reporting and pediatric sites combined, associated CI are calculated and presented and appear acceptable
- 2b.1.1 - Measure specifications are aligned with studies indicating poor retention, increased rates of missed medical visits and gaps lead to poorer health outcomes.
- 2b.2.1 - Data sample is presented and is sample presented also for reliability (same comments apply for representative statements)
- 2b.2.2 - Analytic method is described in detail as face validity established and systematically assessed by technical work group using modified Delphi Process as well as feedback from providers using the measure related to feasibility and usability of the measures.
- 2b.2.3 - Results show that the technical work group deemed the measure important, usable, and feasible.
- 2b.3 - No discussion of threats to validity
- 2b5 - Data sample shown with ability to capture disparity data and performance results presented including min, max, mean and quartiles
- 2b6 - Measure not tested on multiple data sources - highest rating is moderate
- 2c.1 - Data is stratified for disparities
- Did not run data for excluded patients

3. Usability: H-3; M-1; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

- 3.1 - Measure is currently being used for quality improvement with benchmarking external to reporting organization
- 3a.1 - Measure is used in national quality improvement project, technical work group saw utility in public reporting, measure (upon endorsement) will seek inclusion in Stage 3 of CMS EHR incentive program and PQRS
- 3a.2 - Measure is able to distinguish difference in performance and data presented with consistency and top/bottom performers - QI project participants reported that measure is meaningful to management of *their* patient populations and is understandable by patients and providers – reporting sites showed a sharp decline in number of sites reporting during the fourth measurement period – why? Are the providers finding this measure less useful?
- 3b.1 - Used currently for QI in national project
- 3b.2 – Measure identifies which patients are in need of follow-up; measure brings together all disciplines to engage in decreasing gaps of medical visits – does not leave burden on the prescriber – data can be stratified to determine risk factors associated with lost to follow up (evidence provided that this has been done by at least one provider in the reporting group)
- For this measure, how can one adjust for non compliant patients which may be beyond control of physician

4. Feasibility: H-3; M-1; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

2080: Gap in medical visits

Rationale:
- 4a. Data generated during delivery of care
- 4b. Data available in electronic claims
- 4c. No susceptibility to errors discussed - would paper charts match with EHR data? does the removal of incarcerated and transferred to care skew the performance? While cited in 4d as difficult to capture - might it be worth identifying how this affects the performance outcome and interpretation given the high rates of incarceration in persons of color who also show high incidence and disease burden?
- 4d - Collection strategy changed based on feedback from providers to eliminate certain exclusions - cited that data variance is a result of performance and not differences in data availability.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-4; N-0</mark>

Rationale:
- Compared to the medical visit endorsed measure this measure seems to be associated more with follow-up rather than identification of a adherence to visits – its semantics but seems an important difference that during implementation could change how the providers utilize the measure and interpret performance
- Would like to see discussion about exclusions and non compliant patients

Additional Comments/Questions:

Workgroup Discussion

Importance to Measure and Report
- There is a significant opportunity for improvement.
- One WG member questioned whether the evidence is robust enough to say that retention in care measured over a two-year timeframe is more closely linked with long-term patient outcome than retention in care over a one-year timeframe which is the central difference between measure 2079 and 2080.
  - o The developers cited an increase in baseline CD4 count was significantly higher with optimal retention over 24 months as opposed to patients who did not have optimal retention.
  - o From a consumer perspective, at the one-year mark – there is the scare of the initial diagnosis that keeps the patient in care. However, at the two to three-year mark patients tend to drop off because they get comfortable with their care.

Scientific Acceptability of Measure Properties
- This measure was tested at the measure score level; as such the highest rating can be moderate.
- The measure does not provide a definition of a 'medical visit'. WG members questioned HRSA as to what they envisioned as far as a 'medical visit' (i.e. could it be a visit with a clinical social worker -- would that count as a hit in their system?)
  - o HRSA replied that the original source of the data is in the EHR or a paper chart and when abstracted from either of those sources it could be defined as a face-to-face visit with a physician, a nurse practitioner, or a physician's assistant, or with someone else who is licensed to prescribe in their jurisdiction. In the event this measure is endorsed the developers will specify the measures for use in electronic health records and will use CPT codes to define what would characterize a medical visit.

0403: HIV/AIDS: Medical visit
Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: Percentage of patients, regardless of age, with a diagnosis of HIV/AIDS, with at least two medical visits during the measurement year, with a minimum of 90 and 180 days between each visit

Numerator Statement: Numerator 1: Patients with at least two medical visits during the measurement year, with a minimum 0f 90 days between each visit

Numerator 2: Patients with at least two medical visits during the measurement year, with a minimum of 180 days between each visit

Definition of "Medical Visit" - any visit with a health care professional who provides routine primary care for the patient with HIV/AIDS (may be a primary care physician, ob/gyn, pediatrician or infectious diseases specialist)

Denominator Statement: All patients, regardless of age, with a diagnosis of HIV/AIDS

Exclusions: None.

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data

Measure Steward: National Committee for Quality Assurance **Other organizations:** Physician Performance Measures (Measures) and related data specifications have been developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement™ (the Consortium) and the National Committee for Quality Assurance (NCQA). The Health Resources and Services Administration (HRSA) and the Infectious Diseases Society of America also participated in the development of this measure.

0403: HIV/AIDS: Medical visit

IMPLEMENTATION COMMENTS:
- HIV Medicine Association: We understand that different time periods for medical visit intervals are under consideration by various measures development groups (60, 90 or 180 days). We note that the National HIV/AIDS Strategy (NHAS) performance metrics reflect a 90-day time period between medical visits. We strongly urge that whichever interval is chosen for this measure, it should be uniform across payers and health care platforms. We support continued NQF endorsement with uniformity of medical visit intervals across measures in which medical visit frequency is a factor.
  - Developer response: NCQA has included HRSA, the CDC, and a representative from HIVQUAL in discussions about the time interval for the HIV/AIDS: Medical Visit measure. We have attempted to align with other HIV measurement programs.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the Committee Reviewers: Michael Farber; Kalpana Ramiah; Adam Thompson; Ed Septimus (*comments separated by bullets*)

Importance to Measure and Report (*based on decision logic*): Y-4; N-2

1a. Impact: H-4; M-0; L-0; I-0; **1b. Performance Gap:** H-1; M-3; L-0; I-0

Rationale:
- Data were submitted that demonstrate variations in this process measure but no data were submitted on health disparities
- 1a.3 - The NHAS seeks to improve the percentage of patients receiving continous medical care which means the measure of medical visits is aligned with a national priority. Not receiving regular medical care prohibits patients from accessing life saving medications.
- 1b.1 Improvements in medical visits result in improvement in health outcomes.
- 1b.2 NHAS indicates only 73% of HIV patients recieved the standard of care which shows room for improvement
- 1b.4 - Known disparities exist yet the measure is not stratefied by disparties - weakness to the measure - rated moderate.
- Great

1c. Evidence (*based on decision logic*): Y-3; N-1   **IF a Health Outcome**, rationale supports: Y-2; N-1; NA-1

Quantity: H-2; M-1; L-0; I-1; **Quality:** H-0; M-3; L-0; I-1; **Consistency:** H-2; M-2; L-0; I-0

Rationale:
- 5+ studies identified in the report (1c.5)

0403: HIV/AIDS: Medical visit

- There is question if visits every 90 days are needed in stable patients (1c.14), but general agreement on the frequency of CD4 and viral load monitoring every 3-4 months (note that this might not require a "visit" as defined by the measure). Financial harms to patients are not measured and benefits to stable patients are not identified (re consistency). Note the evidence is graded level II by HIVMA and level III by NYS.
- 1c.4 - The evidence for HIVMA is cited as to the importance of measuring CD4 and Viral Load which is not always an indication that the patient actually had a "medical visit." The evidence for the NYSDOH is specific to medical visits
- 1c.5 - The evidence cited for the HIVMA guideline - is it specific to "medical visit" - seems only the provision of no the need for - only to CD4 count and viral load monitoring which can occur without a medical visit actually occurring. The NYSDOH evidence cited is a literature review not an actual study itself
- 1c.6 No discussion of the evidence other than a brief descriptive statement.
- 1c.7 - Consistency statement that the studies demonstrate regular medical visits decrease mortality - yet without the studies specifically cited, rated moderate.
- Great

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-1
2a. Reliability: H-1; M-2; L-0; I-1; 2b. Validity: H-1; M-2; L-0; I-1
Rationale:

- Reliably collected electronically compared to manually.
- Face validity has been measured and is high.
- No risk-adjustment strategy is identified
- It is suggested not to use this as a measure of disparities of care (2c)
- While performance results have been reported (2b.5) the impact on patient outcome is not clearly identified.
- 2a - Measure steward indicates that reliability testing was not needed yet in the data source they do not specify EHRs - only electronic data which would necessitate separate reliability testing.
- 2b.2.1 - Tested using the EHR yet this is not specified as a data source for the measure. Was the sample representative? Face validity was cited yet there was no indication if face validity was systematically assessed. According to the guidance this would earn an insufficient to rate. Thus both validity and reliability are rated insufficient to rate.
- 2b.2.3 - Validity testing indicated a low difference betwen the paper and electronic sources.
- 2b.3 - No patient exclusions - how do they handle persons who are incarcerated or transfer care to another provider? This could show that the site is performing worse that it is; high rates of incarceration are prevalent in HIV infected inviduals, particularly persons of color in some regions
- 2b5.1 - The measure is specified at 90 and 180 days in the numerator and yet the meaning difference in performance is shown at a 60 day interval.
- 2c - No disparity data presented.
- No exclusion in this measure such as deaths or incarceration
- Great

3. Usability: H-1; M-3; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

- The document states that this measure has or may be used by many organizations as a quality measure. there are no data presented that instituting such a measure leads to better process or outcome of care. So it may be easy to use but proof of patient benefit is not shown
- 3a.1 - Measure currently being used in Medicaid Eligible Adults - but only used in 2012 - why was measure not picked up sooner? - considered for inclusion in CMS EHR Incentive program
- 3a.2 - HRSA uses similar measure for public reporting and accontability
- 3b.2 - Measure aligned with current clinical guildelines however has a different interval than the data cited in the NHAS - cites retention measure from HRSA as similar but they appear different, not sure I would confidently compare the two.
- Problems accurately capturing exception per developer
- Great

4. Feasibility: H-1; M-3; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified*
4d. Data collection strategy can be implemented)
Rationale:

0403: HIV/AIDS: Medical visit
4a - Routinely generated and used during care delivery   4b - All data in electronic sources   4c - No identified susceptibilities to errors - yet how sites define a missed visit could change the data .   4d. Descriptive statement no specific to this measure but rather all measures from this particular measure steward; insufficient to rate.   - Again per developer, they admit to difficulties with accurately capturing certain aspect on the health record   - Great

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-3; N-1</mark>

Rationale:
- It is likely accurately collected electronically and offers a marker of the process of care. However, whether modifications in performance on this measure will lead to better outcome is unclear. The range of providers who are allowed in the numerator (see "definition of medical visit 2a1.1), while easier to collect electronically, may dilute the impact of the measure. Laboratory evaluations for CD4 and viral load that may not be performed in the context of a medical visit are perhaps more impoortant to consider than the visit itself, if communication of results can be documented (the visit may not be important: perhaps just getting the labs with an email that says "doing well--keep going" is enough). This concept comes out in the data and recommendations on "2 visits per year in stable patients".
- Insuffient ratings the quantity of evidence and both reliability and validity do not currently allow for the measure to meet the criteria for endorsement. Measure reflects attendance to visits but not linkage to care or retention to care which is more currently where the PLWH community is more interested . . . not did we make a medical visit but rather were we consistently in care over extended periods of time.
- Great

Additional Comments/Questions:
- some overlap with 2081, 2080
- great

<mark>Workgroup Discussion</mark>

Importance to Measure and Report
- This measure is used in the PQRS program.
- This measure reports two rates, one for 90 days and one for 180 days.
- There was some concern that the data cited in the evidence section does not support the measure focus, rather they are based on CD4, and viral load monitoring frequency instead of physician visit frequency.
    - There was significant evidence presented for this measure around the importance of having lab values yet from a patient's standpoint, labs can be drawn and results can be received without ever having to see a clinician.
- This measure does not collect information on disparities. The developer explained that this is because CMS is implementing the measure in PQRS and CMS does not report using disparity. This measure as currently specified would not prohibit a practice or groups to look at disparities if they were using this measure.

Scientific Acceptability of Measure Properties
- Definition of a medical visit is not clear; not all providers listed will be HIV providers (i.e. a visit to the OB/GYN being counted as a marker of retention for HIV care would overestimate the reported rate)
    - Data on staying in care aren't necessarily focused on seeing a provider as much as getting CD4 virus count.
    - There are many things that occur in a visit. Some of them are not easily measurable. The one that is very easily measurable is how often you get a CD4 count but that is not the sole issue for being in the visit.
    - The developer explained that the reasoning behind the 'medical visit' definition is to compensate for areas in the country where they may not be HIV specialists.
    - Workgroup members indicated there should be a reason for the visit pertaining to HIV that should be documented. The developer has provided some guidance in the definition of the medical visit.
- One WG member felt that laboratory studies may be a better indicator to capture a 'visit,' stating that often times physicians' measure things not by a visit. Looking at lab studies would allow the measure to discern the nature of the visit of a patient who is seeing an OB/GYN.
- The validity testing was performed using electronic health records, yet the data sources specified do not include EHRs.
    - The developer responded:
        - *The measures were developed for the PQRS program which uses administrative claim and CPT category 2 codes. The measures haven't been implemented in PQRS when they were tested back in 2009. So given that, the tested in the EHR should see whether that type of information is available even though the measures are*

0403: HIV/AIDS: Medical visit

> *using claims data and (CPT 2) code. And one reason why we think that works is that category Q codes actually – because they are quality claims or quality administrative data– they're not like other claims data. The provider could be using a paper medical record or an electronic medical record or some other types of claim to then report their category Q codes. So we do think that the EHR testing data does support the category Q code specifications for the measure.*

- The developer used face validity where they convened expert panel to reassess these measures in 2012. The panel was surveyed and given a 5-point record scale to grade the measures according to face validity.

Feasibility

- The developer is under contract with the ONC to develop eMeasures for consideration in the meaningful use program.
- The CPT codes would be available in an EHR based data collection system so that enhances its feasibility as well.

0408: HIV/AIDS: Tuberculosis (TB) screening

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: Percentage of patients aged 3 months and older with a diagnosis of HIV/AIDS, for whom there was documentation that a tuberculosis (TB) screening test was performed and results interpreted (for tuberculin skin tests) at least once since the diagnosis of HI V infection.

Numerator Statement: Patients for whom there was documentation that a tuberculosis (TB) screening test was performed and results interpreted (for tuberculin skin tests) at least once since the diagnosis of HIV infection.

Denominator Statement: All patients aged 3 months and older with a diagnosis of HIV/AIDS, who had at least two visits during the measurement year, with at least 90 days in between each visit

Definition of "Medical Visit" - any visit with a health care professional who provides routine primary care for the patient with HIV/AIDS (may be but is not limited to a primary care clinician, ob/gyn, pediatrician, infectious diseases specialist)

Exclusions: Documentation of Medical Reason for not performing a tuberculosis (TB) screening test (e.g., patients with a history of positive PPD or treatment for TB, patient declined)

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data

Measure Steward: National Committee for Quality Assurance **Other organizations:** Physician Performance Measures (Measures) and related data specifications have been developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement™ (the Consortium) and the National Committee for Quality Assurance (NCQA). The Health Resources and Services Administration (HRSA) and the Infectious Diseases Society of America also participated in the development of this measure.

0408: HIV/AIDS: Tuberculosis (TB) screening

IMPLEMENTATION COMMENTS:
- HIV Medicine Association: We support continued NQF endorsement of this measure, as written. It is still clinically relevant.
  - Developer response: Thank you for your support.

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the **Committee Reviewers:** Michael Farber; Kalpana Ramiah; Adam Thompson; Ed Septimus   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-1; N-2
1a. Impact: H-2; M-0; L-1; I-0; **1b. Performance Gap:** H-1; M-0; L-1; I-1
Rationale:
- 1a.3 - Evidence cited for high impact is incidence of undiagnosed HIV in patients with diagnosed TB. There is evidence cited that 30% of PLWH who have latent TB will develop active TB but no evidence given of what percentage of PLWH are diagnosed with TB.
- 1b.2 Measure steward indicates the measure is not publicly reported and uses evidence from the HIVQual measure however the specifications of the two measures are different; this measure specifies that results are interpreted but does not indicate that results are documented - the HIVQual measure does. This would seem to make the measures and thus the performance scores uncomparable for the purpose of demonstrating performance gap.
- 1b.4 Measure stewards indicate a lack of consistent standards related to capture and documentation of the data as the barrier to reporting disparties - later they will cite their own evidence support as contrary evidence which clearly states annual screening to be recommended for persons at high risk . . . without stratifying the data for disparities, high risk cannot be determined. Again a problem with the measure specifications in relationship to the evidence.

1c. Evidence (*based on decision logic*): Y-2; N-1   IF a **Health Outcome**, rationale supports: Y-2; N-0; NA-1
Quantity: H-2; M-0; L-0; I-1; **Quality:** H-1; M-1; L-0; I-1; **Consistency:** H-2; M-1; L-0; I-0
Rationale:

0408: HIV/AIDS: Tuberculosis (TB) screening

- 1c.5 - Measure steward cites CDC guidelines on treatment of OIs but not TB specifically. The studies used to support quantity cannot from the submission be associated with TB without more information. Quantity is therefore insufficient and would need more specific information on which studies were TB related.
- 1c.6 - The quality of the evidence is also insufficient for the same reasons as the rating of quantity. Based on the measure submission from the steward, one cannot determine that these studies were TB specific.
- 1c.7 - While the studies may be consistent without the quantity and quality described more specifically, I would hesistant to agree that they are consistent without knowing more - rated as moderate because the measure has previous endorsement and the HIVMA recommended.
- 1c.14 - Contraversy/Contrary Evidence is the same guidelines used as supporting evidence. This guideline recommends annual testing for high risk individuals yet the measure avoids this specification. It would appear that the argument of the data burden (made later in scientific acceptability section) is inconsistent with the guidelines used to support the measure itself. The guideline even specifies what high risk is and while difficult perhaps to capture, should not be ignored. Overall - while a process, without ensuring that the results are documented rather than simply interpreted, it would appear that this process measure may not lead to the desired health outcome. Without specifying who can interpret those results, the process could fail (particularly considering that some providers allow patients to interpret screening results and report via phone). The desired health outcome is not ensured by this measure as written.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-2; N-1
2a. Reliability: H-1; M-1; L-1; I-0; **2b. Validity:** H-0; M-1; L-2; I-0
Rationale:

- 2a1.1 - Numerator Statement - measure as written specifies that the results are interpreted but does not indicate by whom and how (if at all) the results are documented. The interpretation, as written, could be done by anyone and the variability among provider perceptions of acceptable interpretors is too vast (as cited by the measure stewards) to have the numerator remain so unspecific. Numerator also indicates screening at least once; the guideline cited however recommends annual screening for high risk populations, the evidence does not support the numerator statement.
- 2a.1.4 - Denominator Statement - the specification of medical visit is good - it is confusing as to why the measure steward would be specific on the definition of medical visit and not on the interpretor of the skin test result.
- 2a.1.25 Reliability testing is waived according to the measure stewards because the data source is electronic health records - while the validity testing is done with the EHR the data source is listed as electronic clinical data which may or may not be the EHR itself - more specificity is needed in the data source.
- 2b2.1 - The sample is not specific as described - "multiple, complex needs in the Midwest region." How was this sample determined to be representative of a measure seeking national use? Face validity is cited yet there is no explanation of whether face validity was systematically assessed and if so, using what method.
- 2b2.3 - The results show a a disparity in the results between the EHR and the actual paper record of 20%. The measure steward indicates this is due to a lack of standardized fields in the site EHRs. Why then not be more specific in the measure to force this standardization? The measure as written seems to support the continued lack of standardization rather than acting a method to help support this standardization thus weakening its usefulness to address the identified problem. 3 members of the face validity panel disagreed with the measure validity.
- Comparison between EMR and manual calculation is substantial

3. Usability: H-0; M-2; L-1; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

- 3a.1 - The measure having received endorsement in 2008 has yet to be used for public reporting - why? Measure steward indicates they will submit to PQRS for consideration but has it been submitted in the past? Why not? or if it was, why was it not used?
- 3a.2 - The measure steward cites the HIVQual measure yet the measure is written differently. The two are not comparable in my opinion.
- 3.2 - May be used but no guarantee
- 3b.1 and 3b.2 - Again the identification of a similar but differently specified measure cannot guarantee the usefulnessof this measure as written - interpretation versus documentation of result is too different in my opinion.

4. Feasibility: H-1; M-1; L-1; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:

- 4c.1 - The inability of the measure to capture the result stops the process, there is no certainty that the interpretation is correct and

0408: HIV/AIDS: Tuberculosis (TB) screening

that the potentially needed follow-up treatment will occur. This is not addressed and thus rated insufficient - no discussion of the provider difficulty in coding this data either which was cited earlier.

- 4d. - The measure steward indicates that the measure suffers from a lack of standardized fields at the provider level for capturing the score accurately which has led to a disparity of 20% - how can they ensure that the strategy will be implemented if there is no specification of interpretation?
- Performance and capturing this information appear challenging

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-1; N-2</mark>
Rationale:

- Currently the measure as written suffers from: Insufficient or inappropriate evidence to support impact and performance gap.   Disparity in the data source specifications and the data source used for validity testing. The validity test of face validity is not explained.   The guidelines used as evidence are also cited as contrary evidence.   Disparity data is not collected.   Significant gap is measure score in different sources (electronic clinical data vs. chart abstration).   HIVQual measure which specifies result documented is not comparable to this measure which only ensures intepetation by a non-specified interpreter (could be patient as written).   Endorsed previously yet still not used for public reporting.   As written does not ensure the health outcome of treating the patient for TB.
- I have problems with endorsement given measure validity

Additional Comments/Questions:

<mark>**Workgroup Discussion**</mark>

Importance to Measure and Report

- This measure has high impact, especially in areas where TB infection is common, i.e. along US and Mexico border.
- TB is environmentally acquired, and the measure's inability to capture persons who are incarcerated, persons who experience homelessness, or in certain institutional settings, limits the ability to target quality improvement.

Scientific Acceptability of Measure Properties

- The WG noted a significant difference between electronic health records versus manual calculation.
  - o The developer responded by explaining that the measure is based on category II codes and an issue which arose during testing was having information for the measure and structure data fields . The data is not being captured in the structure data field for an EHR or in the EHR for an eMeasure.
  - o According to the developer, the measure is still reliable and valid because the Category II code allows the implementer to search the paper medical record in order to find the appropriate information.
- One member noted that the measure as specified, requires that the results are interpreted but does not indicate by whom and how (if at all) the results are documented. The interpretation, as written, could be done by anyone and the variability among provider perceptions of acceptable interpretors is too vast.
  - o Requiring dcocumentation reults in increased accountability for the physician.
  - o The evidence cited, does not support the numerator statement.
  - o The developer noted they will make the change to the measure, clarifying that the intent of the measure which is to have the results interepreted by a medical professional not a patient.
- Predictive ability of TB test depends upon the level if immunosuppression and is not as predictive as CD4 count for PCP prophylaxis.

Feasibility

- Very labor intensive to capture this data and but having said that, there's still tremendous opportunities.

0409: HIV/AIDS: Sexually transmitted diseases – Screening for chlamydia, gonorrhea, and syphilis

Status: Maintenance, Original Endorsement: Jul 31, 2008   (Originally endorsed as two measures – one for syphilis screening and one for chlamydia and gonorrhea screening).

Description: Percentage of patients aged 13 years and older with a diagnosis of HIV/AIDS, who have received chlamydia, gonorrhea, and syphilis screenings at least once since the diagnosis of HIV infection

Numerator Statement: Patients who have received chlamydia, gonorrhea, and syphilis screenings at least once since the diagnosis of HIV infection

Denominator Statement: All patients aged 13 years and older with a diagnosis of HIV/AIDS, who had at least two visits during the measurement year, with at least 90 days between visits

Definition of "Medical Visit" - any visit with a health care professional who provides routine primary care for the patient with HIV/AIDS (may be but is not limited to a primary care clinician, ob/gyn, pediatrician, infectious diseases specialist)

Exclusions: None

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Laboratory

Measure Steward: National Committee for Quality Assurance **Other organizations:** Physician Performance Measures (Measures) and related data specifications have been developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement™ (the Consortium) and the National Committee for Quality Assurance (NCQA). The Health Resources and Services Administration (HRSA) and the Infectious Diseases Society of America also participated in the development of this measure.

0409: HIV/AIDS: Sexually transmitted diseases - Screening for chlamydia, gonorrhea, and syphilis

IMPLEMENTATION COMMENTS:
- HIV Medicine Association: We understand that NCQA is considering merging measures #0409 and #0410 to measure Chlamydia, Gonorrhea and Syphilis. We support continued NQF endorsement of a measure along these lines, as it remains clinically relevant.
  - Developer response: NCQA will be combining measures 0409 and 0410. Thank you for your support

Notes

Workgroup Preliminary Evaluations

The following evaluation ratings and comments are from the **Committee Reviewers:** Michael Farber; Kalpana Ramiah; Adam Thompson; Ed Septimus   *(comments separated by bullets)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-2
1a. Impact: H-4; M-0; L-0; I-0; **1b. Performance Gap:** H-2; M-2; L-0; I-0
Rationale:
- 1a.3 - Evidence provided suggests that the affected population experiences an disproportionate disease burden compared to general population. All three STIs are reportable diseases which indicate alignment with national public health priorities. Evidence cited indicates that untreated specified STIs can increase HIV incidence.
- 1b.1 - Data from CMS PQRS demonstrates a performance gap in quality of care
- 1b.4 - Measure steward indicates populations may experience disparities yet the measure data is not stratafied by disparities.
- Performance in this area is very disappointing

1c. Evidence (*based on decision logic*): Y-2; N-2   IF a **Health Outcome**, rationale supports: Y-1; N-1; NA-2
Quantity: H-2; M-0; L-0; I-2; **Quality:** H-0; M-2; L-0; I-2; **Consistency:** H-0; M-2; L-0; I-2
Rationale:
- 1c.5, 1c.6, 1c.7 - Quality, quantity, and consistency of evidence is not presented as specific to STIs but rather to general prevention efforts for PLWH - evidence as presented is insufficient for rating.
- 1c.14 - The contrary evidence cited is used as the body of evidence as well; why was the measure not aligned with the recommendation? ... expert opinion may not take into account assumptions by providers that patients with HIV are sexually active despite not reporting this to their physicians ... by virtue of the diagnosis many PLWH should be considered sexually active and the

0409: HIV/AIDS: Sexually transmitted diseases – Screening for chlamydia, gonorrhea, and syphilis

measure should be aligned to screen annually if in fact the measure is that important due to high incidence.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): <mark>Y-2; N-2</mark>
2a. Reliability: H-0; M-2; L-0; I-2 **2b. Validity:** H-0; M-1; L-1; I-2
Rationale:

- Face Validity is not appropriately described (no description of how the face validity was systematically assessed) - thus the validity testing is insufficient and the measure stewards submitted validity testing as the reliability testing despite (2a.1.25) not indicating that the data source is an EHR would (based on interpretation of NQF guidance) would require separate reliability testing.
- 2a.1.1 - Numerator Statement - indicates a "screening" - this language should indicate a serologic test so that providers do not assume that screening for sexual activity meets the requirement of screening. Numerator also does not specify that patients should be 13 and older which could result in a larger numerator score if under 13 report sexual activity and recieve the care.
- 2a.1.2 - The numerator time window is not aligned with the guideline used in the evidence which would require an annual screening for those reporting sexual activity
- 2a.1.4 - Definition of medical visit is good
- 2a.2.3 - Validity test results show a difference of 33% for G/C - reported as a technical glitch which is supported by the lack of difference (2%) in syphilis
- 2b2.1 - Is the sample representative? testing performed one year into measure endorsement - should a more recent testing be done?
- 2b2.3 - Face validity results showed there was dissent among the work group as to the annual versus once since diagnosis. No potential threats to validity assessed.
- Technical glitches hindered accuracy and poor performance in this area

3. Usability: <mark>H-2; M-2; L-0; I-0</mark>
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

- 3a.1 - Measure is currently used in CMS PQRS in 2010, 2011, 2012)
- 3a.2 - Use in PQRS shows performance gap, continued use demonstrates usefulness; however screening could be misinterpreted - measure should indicate serologic test (also performance rate difference between eletronic and chart though glitch could reduce usefulness - was this fixed?)
- 3b.1-3b.2 - Similar measures (specified similarly) are in use by HRSA for QI and by providers to drive improvement

4. Feasibility: <mark>H-1; M-2; L-0; I-0</mark>
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:

- 4a.1-2 - Data routinely generated during care though coded and abstracted by someone other than care provider
- 4b.1 - All data elements in electronic sources
- 4c.1 - Concern over screening versus serologic test and difference between chart and electronic data for G/C (perhaps addressed with combined measure)
- 4d.1 - While combined (bundling increases likelihood of both) the data pull from so many sources could be cumbersome.
- need to fix glitches

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-2; N-1</mark>
Rationale:

- Currently the body of evidence and reliability/validity is insufficient for grading.
- The use of screening versus language of serologic test is concerning.
- The lack of disparity data and the specifications of only once after diagnosis seems to ignore known issues of incidence rates in specific subpopulations and continued sexual activity of PLWH.
- I said yes, but validity needs to improve and glitches fixed This is an area which has great opportunity to improve

Additional Comments/Questions:

<mark>**Workgroup Discussion**</mark>

Importance to Measure and Report

- Information provided in the evidence section was not robust.
- The WG voiced concern about the screening only being done once.
  - Does not take into account assumptions by providers that patients with HIV are sexually active despite not reporting this to their physicians. If an individual is HIV positive and they did not acquire the virus either through vertical transmission or

0409: HIV/AIDS: Sexually transmitted diseases – Screening for chlamydia, gonorrhea, and syphilis

injection drug use, the sexual activity might continue.

- o By virtue of the diagnosis many people living with HIV should be considered sexually active and the measure and should be aligned to screen annually due to high incidence.
  - The developers explained that , while they did consider this, the guidelines state that you should only be doing annual training for STDs if the patient is sexually active. While many patients with HIV might be sexually active, there are some and that are not. Doing an annual screening would be considered overused for that particular population. In addition, identifying sexually active patients using claims data is difficult.
- o There is an outstanding question for the developers, as to the percentage of people that are in this category who are not sexually active.

Scientific Acceptability of Measure Properties

- This measure was tested at the measure score level; as such the highest rating can be moderate.
- The WG noted a significant difference between electronic health records versus manual calculation.
  - o The developer explained that at the particular site where the testing was being done, there was a problem in the EHR where test data was not being captured in the correct standardized field. Though, while the automated calculation was not correct, the information was available in the record.
  - o The developers have not tested it in other electronic medical records to see if the data could be more accurately captured.
- WG members expressed concern over "screening" versus serologic test and difference between chart and electronic data for Gonorrhea and Clymidia.
  - o By combining these two measures, the concern is that some providers interpret screening to be screening for sexual activity and if sexual activity is not identified, then the test is not performed. There is a high degree of variation in interpreting what constitutes a 'screening' and the way the numerator is specified is it not clear that it's not screening for sexual history indicating a test but the measure is looking for the provision of the laboratory test itself.
    - The developer agreed to clarify the screening language to reflect that the measure is intended to capture a laboratory test.

Feasibility

- Looking at a large electronic health record, electronic health system, and manual charting, it's difficult to discern who is and who is not sexually active. To make this measure feasible and operationalize, is to expect screening for everyone annually but the developer noted that would be a poor use of resources. As such, the developers determined to do it as a one-time metric.