MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 3210
Measure Title: HIV viral load suppression
Measure Steward: Health Resources and Services Administration - HIV/AIDS Bureau
Brief Description of Measure: 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.
Developer Rationale: Human immunodeficiency virus (HIV) is a communicable infection that leads to a progressive disease with a long asymptomatic period. Approximately 50,000 persons in the United States are newly infected with HIV each year. Without treatment, most persons develop acquired immunodeficiency syndrome (AIDS) within 10 years of HIV infection. Antiretroviral therapy delays this progression and increases the length of survival.

Antiretroviral therapy reduces HIV-associated morbidity and mortality by maximally inhibiting HIV replication (as defined by achieving and maintaining plasma HIV RNA (viral load) below levels detectable by commercially available assays). Viral suppression is a main goal of HIV treatment and an indicator of treatment success and reduction of potential HIV transmission. It is directly related to:

- Reduction in disease progression, incidence of opportunistic infections, the risk of both defining and non-AIDS- defining complications and the incidence and severity of chronic conditions.
- Reduction in the risk of transmitting HIV to a sexual or drug-using partner who does not have HIV.
- Improvement of immune function, quality of life, increase in time until development of AIDS increase in life expectancy.

Being virally suppressed is good for an HIV-positive person's overall health and preventing HIV infection from advancing to AIDS, the last stage of HIV infection.

- Durable viral suppression improves immune function and quality of life, prolongs life, lowers the risk of both AIDS-defining and non-AIDS-defining complications, and prolongs life. The proposed measure will direct providers' attention and quality improvement efforts towards this important outcome.

In 2011, the HIV community saw the emergence of the HIV care continuum. This simple model outlines the sequential steps of medical care that people living with HIV go through from initial diagnosis to achieving the goal of viral suppression. The steps include diagnosis, linkage to care, retention in care, receipt of HIV antiretroviral therapy and viral suppression. This model has been incorporated into the National HIV/AIDS Strategy as it has focused all HIV prevention, care, and treatment efforts in the United States. As outlined in the model, all though there are five different steps, each step is dependent upon each other. For instance, you cannot become virally suppressed if you are not receiving HIV antiretroviral therapy or retained in medical care.

The most recent nationwide data from CDC dated 2014 estimates that although 86% of people living with HIV have been diagnosed, only 40% are engaged in care, 37% have been prescribed HIV antiretroviral therapy, and 30% have achieved viral suppression.

Right now, we are at a very special time and place. Many states and large metropolitan areas across the United States have developed plans to end the HIV epidemic in the communities. These jurisdictions have used the HIV care continuum and its steps as the framework by which they have developed their plans.

Numerator Statement: Patients with a HIV viral load less than 200 copies/mL at last HIV viral load test during the measurement year. The outcome being measured is HIV viral suppression.

Denominator Statement: Patients, regardless of age, diagnosed with HIV during the first 3 months of the measurement year or prior to the measurement year who had at least one medical visit in the measurement year. The target population for this measure is all people living with HIV.
Denominator Exclusions: There are no patient exclusions.

Measure Type: Outcome
Data Source: Electronic Health Record (Only), Other
Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Most Recent Endorsement Date:

New Measure -- Preliminary Analysis

Criteria 1: Importance to Measure and Report

1a. Evidence

1a. Evidence. The evidence requirements for a health outcomes measure include providing rationale that supports the relationship of the health outcome to processes or structures of care. The guidance for evaluating the clinical evidence asks if the relationship between the measured health outcome and at least one clinical action is identified and supported by the stated rationale.

This measure is the new eMeasure version of the chart-abstracted measure #2082. The information provided for Evidence and Opportunity for Improvement is identical to that submitted for #2082. Measure #2082 will be discussed first – the ratings for evidence and opportunity for improvement will automatically be assigned to this eMeasure without further discussion.

Evidence Summary:

- The developer provided a diagram outlining the sequential steps of medical care that people living with HIV go through from initial diagnosis to ultimately achieving viral suppression.
- According to the developer, viral suppression is a main goal of HIV treatment and an indicator of treatment success and reduction of potential HIV transmission.
- Being virally suppressed is good for a HIV-positive person's overall health and preventing HIV infection from advancing to AIDS, the last stage of HIV infection. Viral suppression is directly related to:
    - Reduction in disease progression, incidence of opportunistic infections, the risk of both defining and non-AIDS- defining complications and the incidence and severity of chronic conditions.
    - Reduction in the risk of transmitting HIV to a sexual or drug-using partner who does not have HIV.
    - Improvement of immune function, quality of life, increase in time until development of AIDS and increase in life expectancy.
- The developer provided multiple guidelines for the administration of antiretroviral therapy and viral load monitoring intervals for adults, adolescents and pregnant women.
- The developer provided sufficient evidence demonstrating that antiretroviral therapy and viral suppression reduce morbidity and mortality associated with HIV.

Questions for the Committee:

- *Does the Committee agree that a viral load of less than 200 copies/mL leads to improved patient outcomes for patients with a diagnosis of HIV?*
- *Does the Committee agree that there is at least one thing that the provider can do to achieve a change in the viral load of patients diagnosed with HIV? If so, does the Committee agree there is no need for repeat discussion and vote on Evidence?*

Guidance from the Evidence Algorithm: Health outcome measure (Box 1)→The relationship between the outcome and at least one process is identified and supported by the stated rationale (Box 2)→Pass

Preliminary rating for evidence: ☒ **Pass** ☐ **No Pass**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- Per the developer, currently there is no performance data available from the eCQM. However, the developer provided 2014 nationwide data from the CDC that estimates that although 86% of people living with HIV have been diagnosed, only 30% have achieved viral suppression.
- The developer provided the following facility-level performance rates from the Ryan White HIV/AIDS Program Services Report (RSR) from 2010 – 2014 from the existing chart-abstracted measure, #2082:

	2014	2013	2012	2011	2010
Rate	80.8	76.1	69.9	65.5	61.8
Pts w/ ≥1 medical visit (den)	316,087	327,618	335,408	327,744	324,455
Pts w/viral suppression (num)	255,342	249,436	234,505	214,650	200,584
Mean	80.3	76.1	69.9	64.7	60.6
Median	84.2	80.7	75.6	71.4	67.8
Standard Deviation	15.5	17.0	20.3	22.1	23.8
10th percentile	65.0	57.1	40.2	31.9	19.5
90th percentile	93.1	90.2	88.0	84.9	82.8
Min, Max	0.0, 100.0	0.0, 100.0	0.0, 100.0	0.0, 100.0	0.0, 100.0
Pts w/viral load test performed	293,237 (92.8)	297,066 (90.7)	289,563 (86.3)	273,241 (83.4)	264,630 (81.6)
# of facilities	813	823	816	811	846

Disparities:

- The developer provided the following 2010 – 2014 viral suppression rates from the chart-abstracted measure, #2082:

Age	2014	2013	2012	2011	2010
<13	35.5	37.9	36.3	30.5	36.3
13-14	83.8	81.6	76.2	69.2	60.9
15-19	71.5	65.4	57.3	53.8	51.0
20-24	68.2	60.2	50.7	46.9	41.8
25-29	72.5	66.3	58.3	53.5	48.9
30-34	75.9	70.5	63.5	59.5	55.2
35-39	78.0	73.7	67.3	63.0	60.0
40-44	79.9	75.9	70.3	66.0	62.5
45-49	82.1	78.4	72.9	68.5	64.9

50-54	85.7	81.9	77.0	72.0	67.8
60-64	87.1	83.4	78.5	73.8	69.7
≥65	88.4	84.7	80.7	74.7	70.7
Race/Ethnicity					
American Indian/Alaska Native	82.4	74.2	72.3	68.2	64.9
Asian	83.4	78.5	72.4	67.6	64.8
Black/African American	77.4	72.3	66.0	61.2	56.9
Hispanic/Latino	82.6	78.2	72.7	67.6	62.8
Native Hawaiian/Pacific Islander	75.4	67.9	65.2	67.9	57.9
White	83.8	80.2	74.5	70.3	68.3
Multiple Races	83.6	78.0	71.7	66.0	66.8
Gender					
Male	80.7	76.1	70.2	65.7	62.3
Female	79.8	75.0	68.8	63.7	58.9
Transgender	73.6	72.1	66.3	60.1	55.4

- The developer also provided performance rates based on transmission risk, health care coverage, provider type and National HIV/AIDS Strategy (NHAS) populations from the existing chart-abstracted measure, #2082.

Questions for the Committee:

- *Without data from the eMeasure as specified, do you agree that there is a quality problem achieving viral suppression for patients diagnosed with HIV? Is there opportunity for improvement?*
- *Is a national performance measure warranted?*
- *Are you aware of evidence that disparities exist in this area of healthcare?*

Preliminary rating for opportunity for improvement: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence

*Same as measure 2082

*This is an outcome measure up for new eMeasure review. HIV viral load suppression is linked with decreased disease progression, incidence of OIs, and other clinically relevant outcomes. Significant evidence supports the importance of viral load suppression.

I am not aware of any new studies that alters the evidence base.

However, the viral load indicated by the viral load measure is set at <200. With improvements in antiretroviral therapy and assays to measure viral load, guidelines support viral suppression which would now be considered levels much lower than 200 (i.e. less than 20 or undetectable). Current performance data suggests there is still opportunity for improvement even with a permissive cutoff of 200. The developer should consider reassessing the cutoff range in the future as rates of compliance increase."

1b. Performance Gap

***Same as measure 2082

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

2a1. Reliability Specifications

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

Data source(s): electronic health record (EHR). This is an eMeasure.

Specifications:

- HQMF specifications for the eMeasure are included in the document set on SharePoint. See eMeasure Technical Advisor review below.
- The level of analysis is at the facility-level.
- The numerator includes the 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.
- The denominator includes the number of patients, regardless of age, with a diagnosis of HIV with at least one medical visit in the measurement year. To be included in the denominator, patients must meet all of the following conditions/events:
  - Patients of any age during the measurement year
  - Patients diagnosed with HIV during the first 3 months of the measurement year or prior to the measurement year
  - Patients who had at least one medical visit during the measurement year
- There are no patient exclusions.
- The value sets needed to calculate the numerator and denominator are included in the specifications.
- The calculation algorithm is included.

Questions for the Committee:

- *Are all the data elements clearly defined? Are all appropriate codes included?*
- *Is the calculation algorithm clear?*
- *Is it likely this measure can be consistently implemented?*

eMeasure Technical Advisor(s) review:

Submitted measure is an HQMF compliant eMeasure	The submitted eMeasure specifications follow the industry accepted format for eMeasure (HL7 Health Quality Measures Format (HQMF)). **HQMF specifications** ☒ **Yes** ☐ **No**
Documentation of HQMF or QDM limitations	**N/A – All components in the measure logic of the submitted eMeasure are represented using the HQMF and QDM**
Value Sets	**The submitted eMeasure specifications uses existing value sets when possible and uses new value sets that have been vetted through the VSAC**
Measure logic is unambiguous	**Submission includes test results from a simulated data set demonstrating the measure logic can be interpreted precisely and unambiguously**
Feasibility Testing	**The submission contains a feasibility assessment that addresses data element feasibility and follow-up with measure developer indicates that the measure logic is feasible based on assessment by EHR vendors**

2a2. Reliability Testing Testing attachment

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers.

SUMMARY OF TESTING

Reliability testing level ☐ **Measure score** ☒ **Data element** ☐ **Both**

Reliability testing performed with the data source and level of analysis indicated for this measure ☐ **Yes** ☒ **No**

Method(s) of reliability testing:

- The dataset used for testing included 34 synthetic patients created in the Bonnie testing system simulating the year 2012. The developer tested the following data elements using the Bonnie testing tool to evaluate the measure logic:
  - Patient name
  - Date of birth
  - Race
  - Ethnicity
  - Gender
  - Payer
  - Diagnosis
  - Laboratory tests and associated results
  - Encounters
- The patient bundle's demographics were designed to mimic the HIV/AIDS population, specifically drawing from the patient characteristics collected via the Ryan White HIV/AIDS Program Services Report (RSR).
- Data element validity testing was performed and will count for data element reliability – see validity testing section.
- The developer provided reliability results from the chart-abstracted measure (#2082) and stated, "Currently, there is no performance data available to test the eCQM. However, the chart-abstracted version of this measure has been in use in national quality reporting programs since as early as 2010."

Questions for the Committee:

- *Is the test sample adequate to generalize for widespread implementation?*
- *Do the results from the Bonnie tool demonstrate sufficient reliability so that differences in performance can be identified?*
- *Do you agree that the reliability results of the eMeasure will be comparable to the chart-abstracted measure (#2082)?*

Guidance from the Reliability Algorithm: Precise specifications (Box 1) → Empirical reliability testing (Box 2) → Empirical validity testing of patient-level data (Box 3) → Refer to validity testing of patient-level data elements using Bonnie tool (Box 10) → Method appropriate for legacy eMeasures (Box 11) → Moderate (highest eligible rating is MODERATE)

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

2b. Validity

2b1. Validity: Specifications

2b1. Validity Specifications. This section should determine if the measure specifications are consistent with the evidence.

Specifications consistent with evidence in 1a. ☒ **Yes** ☐ **Somewhat** ☐ **No**

Question for the Committee:

- *Are the specifications consistent with the evidence?*

2b2. Validity testing

2b2. Validity Testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

SUMMARY OF TESTING

Validity testing level ☐ **Measure score** ☒ **Data element testing against a gold standard** ☐ **Both**

Method of validity testing of the measure score:

- ☐ **Face validity only**
- ☐ **Empirical validity testing of the measure score**

Validity testing method:

- The Bonnie testing environment, with 34 synthetic patient records, were used to test the measure logic and data elements.
  - For each synthetic patient, an expected result was assigned to reflect an expected result of the measure. The synthetic patients were then run against the HQMF output loaded into Bonnie, which "calculates" a measure result for each patient and evaluates it against the expected result.
  - A patient is considered to pass Bonnie testing when the expected result matches the "calculated" result.
- The developer conducted the following testing on synthetic patients:
  - 100% logic coverage: The bundle of synthetic patients collectively includes all data elements and conditions that are specified within the measure logic.
  - Edge case testing: Data elements that test the upper or lower boundary of measure logic conditions.
  - Negative testing: Use of test cases that do not evaluate positively against the measure logic but are otherwise clinically relevant and realistic.
- The developer used references cited within the chart abstracted measure specifications to ensure the eCQM logic maintained alignment with the clinical intent of the chart abstracted measure.
- In addition to Bonnie testing, the measure specifications were reviewed independently by three eCQM experts to confirm the logic was syntactically correct, using appropriate and current versions of the eCQM standards and terminologies, and consistent with the intent of the chart-abstracted measure.

Validity testing results: See Bonnie testing results

- The testing results from the Bonnie tool reached 100% coverage and confirmed there was a test case for each pathway of logic (negative and positive test cases).
- The measure also had a 100% passing rate which confirmed that all the test cases performed as expected.

Questions for the Committee:

- *Is the test sample adequate to generalize for widespread implementation?*
- *Do the results from the Bonnie tool demonstrate sufficient validity so that conclusions about quality can be made?*
- *Do you agree that the reliability results of the eMeasure will be comparable to the chart-abstracted measure (#2082)?*

2b3-2b7. Threats to Validity

2b3. Exclusions:

N/A

2b4. Risk adjustment: **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

- The Ryan White HIV/AIDS Program is a public health, safety net program providing care to a high proportion of racial/ethnic minority, transgender, unstable housing, and low income people living with HIV. Many of people served by the Ryan White HIV/AIDS Program represent sociodemographics factors incorporated in risk adjusting models by many measures stewards. As a result, the Ryan White HIV/AIDS Program does not adjust for risk in its performance measures. Rather, it is a fundamental aspect of the Ryan White HIV/AIDS Program to identify

disparities and work to improve quality of care for subpopulations. Additionally, this measure is not used for pay-for-performance, bonuses, or penalties.

2b5. Meaningful difference *(can statistically significant and clinically/practically meaningful differences in performance measure scores can be identified)*:

- The developer provided the percentage of patients with viral suppression across providers from the chart-abstracted measure.

Question for the Committee:

- *Does the Committee agree the e-Measure will demonstrate similar results to the chart-abstracted measure?*

2b6. Comparability of data sources/methods:

- N/A

2b7. Missing Data

- Per the developer, "The HQMF standard specifies that if data are unknown or missing, they shall fail the criterion. This constraint embodies the notion that absence of evidence is evidence of absence, i.e. data not present in a structured field from which the measure draws will not be considered for measure calculation. In certain cases, missing data may have no impact on the measure outcome for a given patient. For example, a data element used in a series of OR statements will not impact the measure outcome if another data element in the OR statement is present and meets all other defined constraints."
- All Bonnie synthetic patients with missing data performed according to the HQMF standard specification and as expected.

Guidance from the Validity Algorithm: Specifications consistent with evidence (Box 1) → Threats to validity assessed (Box 2) → Empirical validity testing (Box 3) → Empirical validity testing of data elements and measure logic using Bonnie tool (Box 10) → Method appropriate for legacy eMeasures (Box 11) → Moderate (highest eligible rates is MODERATE)

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2d)

2a1. Reliability Specifications

*Data elements are clearly defined. The eMeasure technical review is positive. It is stated that the level of analysis is at the facility level.

*This measure applies at the facility level. Numerator and denominator statements are clearly defined. There are no exclusions. the calculation algorithm is a simple rate/proportion without any risk adjustment.

I have no concerns that this measure would be inconsistently applied based on the provided definitions. "

*see 2082

2a2. Reliability Testing

*Data is provided on the reliability of data obtained from 34 simluated patients. Agree with the preliminary rating for reliability.

*The Bonnie testing system was used to simulate the data elements including dob, race, ethnicity, gender, payer, diagnosis, lab tests, encounters. This data set included 34 synthetic patitients. Although a chart abstracted measure #2082 has existed for several years, corresponding data for an eCQM has not existed. Reliability testing has not been performed with real data.

*Not clear that the test sample is adequate to generalize for widespread implementation
Results from the Bonnie tool are from only a few synthetic patients so may not be of sufficient reliability

2b1. Validity Specifications

*The specifications of the measure are consistent with the evidence.

*the specifications are consistent with the evidence

2b2. Validity Testing
*The proposed eMeasure performed well on the sample of simulated patients.

*Validity testing was performed in the bonnie testing environment using 34 synthetic patients. It passed 100% logic coverage, edge case and negative testing.
The eCQM logic was reviewed by 3 independent ECQM experts.
Although 34 synthetic patients seems like a small number of patients, there was a test case for each logic pathway. "

*Not clear that the test sample is adequate to generalize for widespread implementation
The results from the Bonnie tool demonstrate sufficient validity so that conclusions about quality can be made
I agree that the reliability results of the eMeasure will be comparable to the chart-abstracted measure (#2082)"

2b3-7. Threats to Validity
*A similar question to measure 2082: It is stated that the level of analysis is at the facility level, it is somewhat unclear whether this measure is intended to be a measure of quality for individual providers or for facilities.

*There are no exclusions nor risk adjustment as the developer provides care to a diverse patient population.
As with the chart abstracted measure, this measure shows significant divergence in rates across providers. This data though is presented on the provider level, as calculated for the chart abstracted measure,. This eMeasure is testing at the facility level.
The developer has criteria for dealing with missing data. If data are unknown or missing, they shall fail the criterion. Some missing data may not impact compliance with the measure when in a series of OR statements. Missing data was in Bonnie synthetic patient testing and performed as expected. "

*2b.5 the e-Measure should demonstrate similar results to the chart-abstracted measure

Criterion 3. Feasibility

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- The developer provided information on the feasibility testing in the eMeasure Feasibility Scorecard. The developer did not identify specific EHRs used to test feasibility in ambulatory care. Instead, the developer stated that the feasibility assessment "conducted by consensus of a panel of MITRE clinical informatics, measure development, and eCQM experts."
- The developer provided a summary of the latest publicly available data on Meaningful Use EHR capabilities and provider performance on objectives and measures directly relevant for the eCQM's data elements:
    - CPOE – Meds
    - CPOE – Labs
    - Demographics
    - Lab test results
    - Problem list
- On a scale from 1 to 3 (3 = highest score), all of the data elements received a score of '3' except, "Encounter, Performed: Face-to-Face Interaction (2)" and "Patient Characteristic Payer (2)".
- Score 2 definition for data standards: Terminology standards for this data element are currently available, but it is not consistently coded to standard terminology in the EHR, or the EHR does not easily allow such coding.
    - For the Encounter, Performed: Face-to-Face Interaction data element, the developer stated that the Health IT Standards Committee recommends SNOMED-CT as the standard terminology for encounters, however SNOMED-CT isn't currently widely used for this purpose. The eCQM alllows for capture of encounters in SNOMED-CT as well as CPT, a more widely used terminology.
    - For the Patient Characteristic Payer, the developer stated that the Public Health Data Standards Consortium developed a standard to encode payer data that is increasingly being adopted. It is anticipated that this data element will be coded in a nationally accepted terminology standard (future state) because the 2015 Edition Health IT certification criteria requires the use of the source payment typology standard.
- Overall, the measure is currently 98.89% feasible and 99.44% feasible in 1 to 2 years according to the scorecard.

- The measure specifications include CPT® codes (requires a license to use) and SNOMED Clinical Terms® (requires a Unified Medical Language System (UMLS) license available for free from the National Library of Medicine).

Questions for the Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *Does the eMeasure Feasibility Score Card demonstrate acceptable feasibility in multiple EHR systems and sites?*
- *Does the feasibility scorecard demonstrate that the eMeasure can be implemented and feasibility concerns can be adequately addressed.*

Preliminary rating for feasibility: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 3: Feasibility

3. Feasibility
*Agree with the preliminary rating of high given the data provided.

*Data elements for this eMeasure are generated during routine care, although the developer did not identify specific EHRs used to test feasibility. Instead • the developer stated that the feasibility assessment "conducted by consensus of a panel of MITRE clinical informatics, measure development, and eCQM experts."
While there are standards formats for all data elements for this emeasure, utilization of EHRs that can electronically capture this data is a potential challenge if EHRs are not widely used. Meaningful use data shows some variation in % providers reporting data per the feasibility scorecard with, if interpreted correctly, some very low rates for reporting - e.g. Stage 1 MU providers in 2015 program year which only required use of EHR to order labs (not capture lab results). It appears that the number of providers that get lab results increases year over year but this is a potential threat to the measure.

*The required data elements are routinely generated and used during care delivery
The eMeasure Feasibility Score Card demonstrates acceptable feasibility in multiple EHR systems and sites"

Criterion 4: Usability and Use
Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact /improvement and unintended consequences

4. Usability and Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

Current uses of the measure
Publicly reported? ☐ **Yes** ☒ **No**

Current use in an accountability program? ☐ **Yes** ☒ **No** ☐ **UNCLEAR**
OR
Planned use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details:

- This newly developed eMeasure is not currently in an accountability program; however, it was reviewed by NQF's Measure Applications Partnership (MAP) for consideration in CMS' Merit Based Incentive Payment System (MIPS). See MAP feedback below.

Improvement results:

- Based on the chart-abstracted measure, the Ryan White HIV/AIDS Program has experienced a 20-point increase in viral suppression from 61.8% in 2010 to 80.3% in 2014. Viral suppression has increased across all demographic groups and subpopulations.

Unexpected findings (positive or negative) during implementation:

- The developer did not provide any unexpected findings during implementation.

Vetting of the measure:

- Health Resources and Services Administration worked diligently to release the annual data report in the same year it was collected (collected in April and released in December of the same year). The report is publically available on the Health Resources and Services Administration website (http://hab.hrsa.gov/data/data-reports) and is released via an accompanying webinar (recorded and archived).
- Ryan White HIV/AIDS Program national partners (national organizations that represent grant recipients, subrecipients, and patients) has provided antidotal feedback regarding the timeliness, feasibility, and usability of the release of the Ryan White HIV/AIDS Program Annual Client-Level Data Report, supplemental reports, slide decks, fact sheets, and infographics. The national partners encourage the continued release of the data in all its formats.
- During the initial development of the chart-abstracted measure, formal feedback was gathered. The measures were modified during the development phase and have not been modified since. A concerted effort was made to develop a measure that would likely stand the test of time from a scientific, clinical, and patient perspective. On an annual basis, the measures are reviewed for clinical relevance, change in scientific acceptability, and consistency with guidelines. The chart-abstracted measure has not been modified as a result of the annual reviews.

Feedback:

- The MAP agreed that this outcome measure addresses an important clinical area. However, it has not been fully tested as an e-CQM. Additionally, the performance data is in the process of being updated from the 2011 data. The measure would address an important issue regarding HIV viral suppression and would provide an additional mechanism for submitting data on this topic. MAP discussed the importance of this measure as it adds an additional outcome measure to the [MIPS] program. The MAP recommended supported this eCQM for rulemaking with the condition that it completes successful testing and the NQF Behavioral Health Standing Committee reviews the performance data to ensure a gap in care continues to exist.

***Questions for the Committee**:*

- *How can the performance results from the eCQM measure be used to further the goal of high-quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*
- *How has the eCQM been vetted in real-world settings by those being measured or others?*

Preliminary rating for usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 4: Usability and Use

4. Usability and Use

*The chart abstraction version of this measure (2082) has been in use and has demonstrated improvement over the past years.

*As a new measure, it is not currently being used in any accountability programs; however the chart review based companion measure is.
HRSA releases the data report in the same year as collected. It is sharedpublically. "

Criterion 5: Related and Competing Measures

Related or competing measures

- 0407 HIV/AIDS: HIV RNA Control After Six Months of Potent Antiretroviral Therapy (NCQA)
- 0405 HIV/AIDS: Pneumocystis Jiroveci Pneumonia (PCP) Prophylaxis (NCQA)
- 0409 HIV/AIDS: Sexually Transmitted Disease Screening for Chlamydia, Gonorrhea, and Syphilis (NCQA)
- 2079 HIV Medical Visit Frequency (HRSA)
- 2080 Gap in HIV Medical Visits (HRSA)
- 2083 Prescription of HIV Antiretroviral Therapy (HRSA)

- 3211 Prescription of HIV Antiretroviral Therapy (HRSA)
- 3210 HIV Viral Suppression (HRSA)
- 3010 HIV Medical Visit Frequency (HRSA)

Harmonization

- Per developer, harmonized with all measures except #0405 and #0409. Plans to harmonize with #0405 and #0409.

Endorsement + Designation

The "Endorsement +" designation identifies measures that exceed NQF's endorsement criteria in several key areas. After a Committee recommends a measure for endorsement, it will then consider whether the measure also meets the "Endorsement +" criteria.

This measure is a candidate for the "Endorsement +" designation IF the Committee determines that it: meets evidence for measure focus without an exception; is reliable, as demonstrated by score-level testing; is valid, as demonstrated by score-level testing (not via face validity only); and has been vetted by those being measured or other users.

Eligible for Endorsement + designation: ☐ **Yes** ☒ **No**

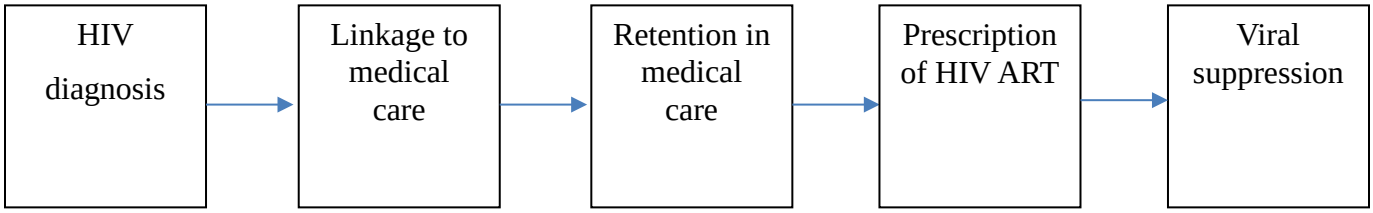
RATIONALE IF NOT ELIGIBLE: The measure is not eligible for Endorsement+ because empirical reliability and validity testing of the measure score was not conducted and the measure has not been vetted in real world settings by those being measured and other users.

Pre-meeting public and member comments

-

Measure Title: Viral Load Supression

1a.12 LOGIC MODEL



Althought the above diagram outlines the sequenctial septs of medical care that people living with HIV go through from initial diagnosis to achieving the goal of viral suppression. For some patients, this is a linear path with sustained viral suppression for many years. For other patients, there many be years between diagnosis and linkage. Yet still for others, retention in medical care is not consistent, which results in missed visits, no prescription for or adherence to HIV antiretroviral therapy (ART), and lack of viral suppression.

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process (e.g., intervention, or service).

Regularly attending medical visits (retention) is paramount to monitoring patients health status, screenings, and laboratory values. Providers need this information to make an informed decision in order to prescribe HIV antiretroviral therapy (ART). ART reduces HIV-associated morbidity and mortality by maximally inhibiting HIV replication (as defined by achieving and maintaining plasma HIV RNA (viral load) below levels detectable by commercially available assays). Durable viral suppression improves immune function and quality of life, lowers the risk of both AIDS-defining and non-AIDS-defining complications, and prolongs life. Emerging evidence also suggests that additional benefits of ART-induced viral load suppression include a reduction in HIV-associated inflammation and possibly its associated complications.

In 2011, the HIV community saw the emergence of the HIV care continuum. This simple model outlines the sequential steps of medical care that people living with HIV go through from initial diagnosis to achieving the goal of viral suppression. The steps include diagnosis, linkage to care, retention in care, receipt of HIV antiretroviral therapy and viral suppression. This model has been incorporated into the National HIV/AIDS Strategy as it has focused all HIV prevention, care, and treatment efforts in the United States. As outlined in the model, all though there are five different steps, each step is dependent upon each other. For instance, you cannot become virally suppressed if you are not receiving HIV antiretroviral therapy or retained in medical care.

The most recent nationwide data from CDC dated 2014 estimates that although 86% of people living with HIV have been diagnosed, only 40% are engaged in care, 37% have been prescribed HIV antiretroviral therapy, and 30% have achieved viral suppression.

Right now, we are at a very special time and place. Many states and large metropolitan areas across the United States have developed plans to end the HIV epidemic in the communities. These jurisdictions have used the HIV care continuum and its steps as the framework by which they have developed their plans.

In closing, the measures we have put forth are in alignment with the HIV care continuum. We see these measures as a suite – each important as individual measures, but work together as a suite to improve health outcomes for people living with HIV in the United States.

1b.1. Briefly explain the rationale for this measure (e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)

Human immunodeficiency virus (HIV) is a communicable infection that leads to a progressive disease with a long asymptomatic period. Approximately 50,000 persons in the United States are newly infected with HIV each year. Without treatment, most persons develop acquired immunodeficiency syndrome (AIDS) within 10 years of HIV infection. Antiretroviral therapy delays this progression and increases the length of survival.

Antiretroviral therapy reduces HIV-associated morbidity and mortality by maximally inhibiting HIV replication (as defined by achieving and maintaining plasma HIV RNA (viral load) below levels detectable by

commercially available assays). Viral suppression is a main goal of HIV treatment and an indicator of treatment success and reduction of potential HIV transmission. It is directly related to:

- Reduction in disease progression, incidence of opportunistic infections, the risk of both defining and non-AIDS- defining complications and the incidence and severity of chronic conditions.
- Reduction in the risk of transmitting HIV to a sexual or drug-using partner who does not have HIV.
- Improvement of immune function, quality of life, increase in time until development of AIDS increase in life expectancy. Being virally suppressed is good for an HIV-positive person's overall health and preventing HIV infection from advancing to AIDS, the last stage of HIV infection.
- Durable viral suppression improves immune function and quality of life, prolongs life, lowers the risk of both AIDS-defining and non-AIDS-defining complications, and prolongs life. The proposed measure will direct providers' attention and quality improvement efforts towards this important outcome.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Department of Health and Human Services, Accessed November 18, 2016: http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf

Recommendations for Use of Antiretroviral Drugs in Pregnant HIV-1-Infected Women for Maternal Health and Interventions to Reduce Perinatal HIV Transmission in the United States, Accessed November 18, 2016

https://aidsinfo.nih.gov/guidelines/html/1/adult-and-adolescent-arv-guidelines/0

International Advisory Panel on HIV Care Continuum Optimization (IAPAC). (2015). IAPAC Guidelines for Optimizing the HIV Care Continuum for Adults and Adolescents. Accessed November 18, 2016. http://www.iapac.org/uploads/JIAPAC-IAPAC-Guidelines-for-Optimizing-the-HIV-Care-Continuum-Supplement-Nov-Dec-2015.pdf

World Health Organization (WHO). (2016). Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. Accessed November 18, 2016: http://apps.who.int/iris/bitstream/10665/208825/1/9789241549684_eng.pdf?ua=1

Günthard HF, Saag MS, Benson CA, del Rio C, Eron JJ, Gallant JE, Hoy JF, Mugavero MJ, Sax PE, Thompson MA, Gandhi RT, Landovitz RJ, Smith DM, Jacobsen DM, Volberding PA. Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults 2016 Recommendations of the International Antiviral Society (IAS)–USA Panel. JAMA. 2016. https://www.iasusa.org/content/antiretroviral-drugs-treatment-and-prevention-hiv-infection-adults-2016-recommendations

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents

Panel's Recommendations for Initiating Antiretroviral Therapy in Treatment-Naive Patients (pE1)

Antiretroviral therapy (ART) is recommended for all HIV-infected individuals, regardless of CD4 T lymphocyte cell count, to reduce the morbidity and mortality associated with HIV infection (**AI**).

Panel's Recommendations for Acute and Recent (Early) HIV Infection (pI1)

- Antiretroviral therapy (ART) is recommended for all individuals with HIV-1 infection **(AI)** including those with early HIV-1 infection.
- Once initiated, the goal of ART is to suppress plasma HIV-1 RNA to undetectable levels **(AIII)**. Testing for plasma HIV-1 RNA levels, CD4 T lymphocyte counts, and toxicity monitoring should be performed as recommended for patients with chronic HIV-1 infection **(AII)**.

Panel's Recommendations Regarding Virologic Failure of the Treatment-Experienced Patient (pH1)

- The goal of treatment for ART-experienced patients with drug resistance who are experiencing virologic failure is to establish virologic suppression (i.e., HIV RNA below the lower limits of detection of currently used assays) **(AI)**.
- Discontinuing or briefly interrupting therapy may lead to a rapid increase in HIV RNA and a decrease in CD4 cell count and increases the risk of clinical progression. Therefore, this strategy is **not** recommended in the setting of virologic failure **(AI)**.

Laboratory Testing, Plasma HIV-1 RNA (Viral Load) and CD4 Count Monitoring (pC5)

- Viral load is the most important indicator of initial and sustained response to ART (AI) and should be measured in all HIV-infected patients at entry into care (AIII), at initiation of therapy (AIII), and on a regular basis thereafter. For those patients who choose to delay therapy, repeat viral load testing while not on ART is optional (CIII).
- Plasma viral load should be measured before initiation of ART and within 2 to 4 weeks but no later than 8 weeks after treatment initiation or modification (AIII). The purpose of the measurements is to confirm an adequate initial virologic response to ART, indicating appropriate regimen selection and patient adherence to therapy. Repeat viral load measurement should be performed at 4- to 8-week intervals until the level falls below the assay's limit of detection (BIII).
- In virologically suppressed patients in whom ART was modified because of drug toxicity or for regimen simplification. Viral load measurement should be performed within 4 to 8 weeks after changing therapy (AIII). The purpose of viral load monitoring at this point is to confirm the effectiveness of the new regimen.
- In patients on a stable, suppressive ARV regimen. Viral load should be repeated every 3 to 4 months (AIII) or as clinically indicated to confirm continuous viral suppression. Clinicians may extend the interval to 6

months for adherent patients whose viral load has been suppressed for more than 2 years and whose clinical and immunologic status is stable (AIII).

Recommendations for Use of Antiretroviral Drugs in Pregnant HIV-1-Infected Women for Maternal Health and Interventions to Reduce Perinatal HIV Transmission in the United States

https://aidsinfo.nih.gov/guidelines/html/3/perinatal-guidelines/0

Panel's Recommendations for HIV-Infected Pregnant Women Who Have Never Received Antiretroviral Drugs (pC27)

- All HIV-infected pregnant women should receive combination antiretroviral therapy (cART) to reduce the risk of perinatal transmission of HIV **(AI)**. The choice of regimen should take into account current adult treatment guidelines, what is known about the use of specific drugs in pregnancy, and the risk of teratogenicity (see Table 6 and Table 7).
- Consideration should be given to initiating cART as soon as HIV is diagnosed during pregnancy; earlier viral suppression is associated with lower risk of transmission. This decision may be influenced by CD4 T lymphocyte count, HIV RNA levels, and maternal conditions (e.g., nausea and vomiting) **(AIII)**. The benefits of early cART must be weighed against potential fetal effects of drug exposure.

Panel's Recommendations Regarding Lack of Viral Suppression During Antepartum Care (pC48)

- Because maternal antenatal viral load correlates with risk of perinatal transmission of HIV, suppression of HIV RNA to undetectable levels should be achieved as rapidly as possible **(AII)**.
- If an ultrasensitive HIV RNA assay indicates failure of viral suppression (after an adequate period of treatment):
    - Assess adherence and resistance (if HIV RNA level is high enough for resistance testing) **(AII)**.
    - Consult an HIV treatment expert and consider possible antiretroviral regimen modification **(AIII)**.
    - Scheduled cesarean delivery is recommended for HIV-infected pregnant women who have HIV RNA levels >1,000 copies/mL near the time of delivery **(AII)**.

International Advisory Panel on HIV Care Continuum Optimization

- Where possible, jurisdictions should consider longitudinal cohort measurement of HIV service utilization and treatment outcomes to identify the means to maximize viral suppression through ensuring early access to ART and retention in care. (A IV) (p4)

World Health Organization:

- ARV drugs play a key role in HIV prevention. People taking ART who achieve optimal viral suppression are extremely unlikely to pass HIV to sexual partners. ARV drugs taken by people without HIV as PrEP or PEP are highly effective in preventing HIV acquisition. (p64)
- People starting treatment and carergivers should be informed that the first ART regimen offers the best opportunity for effective viral suppression, immune recovery and consequently clinical benefit and that successful ART requires all medications to be taken as prescribed. (p72)

- Access to ART should be the first priority for all age groups, and lack of testing for monitoring treatment response should not be a barrier to initiating ART. If viral load testing capacity is limited, it should be introduced in a phased approach. Examples of phased approaches include: …
    - using viral load initially as a targeted test to confirm treatment failure;
    - prioritizing viral load testing for pregnant and breastfeeding women, especially around the time of delivery, as sustained viral suppression is critical to prevention of transmission to the child, and documented high viral load at delivery is an indication for enhanced infant prophylaxis; (p134)

Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults, 2016 Recommendations of the International Antiviral Society–USA Panel

- HIV RNA level should be monitored every 4 to 6 weeks after treatment is initiated or changed until virus is undetectable (evidence rating AIa). (Box 5)
- After viral suppression is achieved, HIV RNA should be monitored every 3 months until suppressed for 1 year and at least every 6 months thereafter for adherent patients who remain clinically stable (evidence rating AIII). (Box 5)
- When virus has been suppressed for at least 2 years and CD4 cell count is persistently above 500/μL, repeat monitoring of CD4 cell count is not recommended unless virologic failure (evidence rating AIIa) or intercurrent immunosuppressive conditions occur or immunosuppressive treatments are initiated (evidence rating AIII). (Box 5)
- If the HIV RNA level remains above the limit of quantification by 24 weeks after starting new treatment or if rebound above 50 copies/mL occurs at any time, the assay should be repeated within 4 weeks to exclude impending virologic failure (evidence rating AIIa). (Box 5)
- For patients with persistent quantifiable HIV RNA between 50 and 200 copies/mL, reassessment for causes of virologic failure, evaluation again within 4 weeks, and close monitoring are recommended (evidence rating BIII). (Box 5)

1a.4.3. Grade assigned to the quoted recommendation with definition of the grade:

Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents and Recommendations for Use of Antiretroviral Drugs in Pregnant HIV-1-Infected Women for Maternal Health and Interventions to Reduce Perinatal HIV Transmission in the United States

Basis for Recommendations

Recommendations in these guidelines are based upon scientific evidence and expert opinion. Each

recommended statement includes a letter (A, B, or C) that represents the strength of the recommendation and a Roman numeral (I, II, or III) that represents the quality of the evidence that supports the recommendation (see Table 2).

Table 2. Rating Scheme for Recommendations

Strength of Recommendation	Quality of Evidence for Recommendation
A: Strong recommendation for the statement **B:** Moderate recommendation for the statement	**I:** One or more randomized trials with clinical outcomes and/or validated laboratory endpoints

C: Optional recommendation for the statement	**II:** One or more well-designed, non-randomized trials or observational cohort studies with long-term clinical outcomes **III:** Expert opinion

International Advisory Panel on HIV Care Continuum Optimization; IAPAC Guidelines for Optimizing the HIV Care Continuum for Adults and Adolescents.

Strong (A) = Almost all patients should receive the recommended course of action.

Moderate (B) = Most patients should receive the recommended course of action. However, other choices may be appropriate for some patients.

Optional (C) There may be consideration for this recommendation based on individual patient circumstances. Not recommended routinely.

Quality of the Body of Evidence and its Interpretation:

Excellent (I) = Randomized control trial (RCT) evidence without important limitations; overwhelming evidence from observational studies

High (II) = RCT evidence with important limitations; strong evidence from observational studies

Medium (III) = RCT evidence with critical limitations; observational study without important limitations

Low (IV) = Other evidence, including extrapolations from bench research, usual practice, expert opinion, consensus guidelines; observational study evidence with important or critical limitations

World Health Organization. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection Recommendations for a public health approach - Second edition.

Process of guideline development This edition of the guidelines was revised in accordance with procedures established by the WHO Guidelines Review Committee. New clinical and operational recommendations in the guidelines are based on the GRADE (Grading of Recommendations, Assessment, Development and Evaluation) approach to reviewing evidence. Modelling, expert consultations and country case studies have all strongly informed the guidelines. The process has also identified key gaps in knowledge that will help to guide the future HIV research agenda. A strong recommendation is one for which there is confidence that the desirable effects of adherence to the recommendation clearly outweigh the undesirable effects. A conditional recommendation is one for which the Guideline Development Group concludes that the desirable effects of adherence to the recommendation probably outweigh the undesirable effects or are closely balanced, but the Groups are not confident about these trade-offs in all situations. At implementation, monitoring and rigorous evaluation is needed to address these uncertainties, which are likely to provide new evidence that may change the calculation of the balance of trade-offs and to suggest how to overcome any implementation challenges.

Quality of evidence Definition

Table 1.1. GRADE quality of evidence

Quality of evidence	Definition

High	We are very confident that the true effect lies close to that of the estimate of the effect
Middle	We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of effect, but there is a possibility that it is substantially different
Low	Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect
Very low	We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of the effect

Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults, 2016 Recommendations of the International Antiviral Society–USA Panel

Table 1. Strength of Recommendation and Quality of Evidence Rating Scale

Rating	Definition
Strength of recommendation	
A	Strong support for the recommendation
B	Moderate support for the recommendation
C	Limited support for the recommendation
Quality of evidence	
Ia	Evidence for $\geq$ 1 randomized clinical trials published in the peer-reviewed literature
Ib	Evidence for $\geq$ 1 randomized clinical trials presented in abstract form at peer-reviewed scientific meetings
IIa	Evidence from nonrandomized clinical trials or cohorts or case-control studies published in the peer-reviewed literature
IIb	Evidence from nonrandomized clinical trials or cohorts or case-control studies published in the peer-reviewed scientific meeting
III	Recommendation based on panel's analysis of the accumulated available evidnce

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system.
(*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

All grade and definitions noted in 1a.4.3.

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

Citations noted in 1a.4.1.

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

X Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1. Evidence, Performance Gap, Priority – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

VS_evidence-636177547737712934.docx

1a.1 <u>For Maintenance of Endorsement:</u> Is there new evidence about the measure since the last update/submission?

Please update any changes in the evidence attachment in red. Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. If there is no new evidence, no updating of the evidence information is needed.

Yes

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. **Briefly explain the rationale for this measure** *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

<u>IF a PRO-PM</u> (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful. (Describe how and from whom their input was obtained.)

<u>IF a COMPOSITE</u> (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and provide rationale for composite in question 1c.3 on the composite tab.

Human immunodeficiency virus (HIV) is a communicable infection that leads to a progressive disease with a long asymptomatic period. Approximately 50,000 persons in the United States are newly infected with HIV each year. Without treatment, most persons develop acquired immunodeficiency syndrome (AIDS) within 10 years of HIV infection. Antiretroviral therapy delays this progression and increases the length of survival.

Antiretroviral therapy reduces HIV-associated morbidity and mortality by maximally inhibiting HIV replication (as defined by achieving and maintaining plasma HIV RNA (viral load) below levels detectable by commercially available assays). Viral suppression is a main goal of HIV treatment and an indicator of treatment success and reduction of potential HIV transmission. It is directly related to:

- Reduction in disease progression, incidence of opportunistic infections, the risk of both defining and non-AIDS- defining complications and the incidence and severity of chronic conditions.
- Reduction in the risk of transmitting HIV to a sexual or drug-using partner who does not have HIV.
- Improvement of immune function, quality of life, increase in time until development of AIDS increase in life expectancy.

Being virally suppressed is good for an HIV-positive person's overall health and preventing HIV infection from advancing to AIDS, the last stage of HIV infection.

- Durable viral suppression improves immune function and quality of life, prolongs life, lowers the risk of both AIDS-defining and non-AIDS-defining complications, and prolongs life. The proposed measure will direct providers' attention and quality improvement efforts towards this important outcome.

In 2011, the HIV community saw the emergence of the HIV care continuum. This simple model outlines the sequential steps of medical care that people living with HIV go through from initial diagnosis to achieving the goal of viral suppression. The steps include diagnosis, linkage to care, retention in care, receipt of HIV antiretroviral therapy and viral suppression. This model has been incorporated into the National HIV/AIDS Strategy as it has focused all HIV prevention, care, and treatment efforts in the United States. As outlined in the model, all though there are five different steps, each step is dependent upon each other. For instance, you cannot become virally suppressed if you are not receiving HIV antiretroviral therapy or retained in medical care.

The most recent nationwide data from CDC dated 2014 estimates that although 86% of people living with HIV have been diagnosed, only 40% are engaged in care, 37% have been prescribed HIV antiretroviral therapy, and 30% have achieved viral suppression.

Right now, we are at a very special time and place. Many states and large metropolitan areas across the United States have developed plans to end the HIV epidemic in the communities. These jurisdictions have used the HIV care continuum and its steps as the framework by which they have developed their plans.

1b.2. **Provide performance scores on the measure as specified (current and over time) at the specified level of analysis**. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.*
See attachment "VS submission form" for formatted data.

1b.3. **If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.**
N/A

1b.4. **Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.** *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.*
Please see attachment "VS submission form" for formatted data.

1b.5. **If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4**
N/A

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. **Subject/Topic Area** *(check all the areas that apply):*
Infectious Diseases (ID) : HIV/AIDS

De.6. **Cross Cutting Areas** *(check all the areas that apply):*
«crosscutting_area»

De.7. **Target Population Category** *(Check all the populations for which the measure is specified and tested if any):*
Populations at Risk

S.1. **Measure-specific Web Page** *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*
There is no measure-specific web page for the electronic version of this measure.

S.2a. **If this is an eMeasure**, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)
This is an eMeasure **Attachment:** NQFXXX_HIVViralSuppression_Artifacts-636178423251224574.zip,NQFXXX_HIVViralSuppression_MeasureSubmissionForm-636178426319023569.docx

S.2b. **Data Dictionary, Code Table, or Value Sets** *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*
Attachment **Attachment:** HIVVLS_v4_6_Thu_Dec_15_20.35.00_CST_2016-636178423774443650.xls

S.3.1. **For maintenance of endorsement:** Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.
No

S.3.2. **For maintenance of endorsement,** please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.
Not applicable

S.4. **Numerator Statement** *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*
IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).
Patients with a HIV viral load less than 200 copies/mL at last HIV viral load test during the measurement year. The outcome being measured is HIV viral suppression.

S.5. **Numerator Details** *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).
The viral load suppression laboratory test is represented by the QDM element "Laboratory Test, Performed: HIV Viral Load" using "HIV Viral Load Grouping Value Set (2.16.840.1.113883.3.464.1003.120.12.1002)". The result of the laboratory test is modeled as an attribute of the Viral Load Suppression QDM element and represented as a numerical result associated with copies/mL as the reporting unit.

S.6. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
Patients, regardless of age, diagnosed with HIV during the first 3 months of the measurement year or prior to the measurement year who had at least one medical visit in the measurement year. The target population for this measure is all people living with HIV.

S.7. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*
IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The patient's HIV diagnosis is represented by the QDM element "Diagnosis: HIV" using "HIV Grouping Value Set (2.16.840.1.113883.3.464.1003.120.12.1003)". The patient's medical visits are represented by the following QDM elements:

- "Encounter, Performed: Face-to-Face Interaction" using "Face-to-Face Interaction Grouping Value Set (2.16.840.1.113883.3.464.1003.101.12.1048)"
- "Encounter, Performed: Office Visit" using "Office Visit Grouping Value Set (2.16.840.1.113883.3.464.1003.101.12.1001)"
- "Encounter, Performed: Outpatient Consultation" using "Outpatient Consultation Grouping Value Set (2.16.840.1.113883.3.464.1003.101.12.1008)"
- "Encounter, Performed: Preventive Care - Established Office Visit, 0 to 17" using "Preventive Care - Established Office Visit, 0 to 17 Grouping Value Set (2.16.840.1.113883.3.464.1003.101.12.1024)"
- "Encounter, Performed: Preventive Care Services - Established Office Visit, 18 and Up" using "Preventive Care Services - Established Office Visit, 18 and Up Grouping Value Set (2.16.840.1.113883.3.464.1003.101.12.1025)"
- "Encounter, Performed: Preventive Care Services-Initial Office Visit, 18 and Up" using "Preventive Care Services-Initial Office Visit, 18 and Up Grouping Value Set (2.16.840.1.113883.3.464.1003.101.12.1023)"
- "Encounter, Performed: Preventive Care- Initial Office Visit, 0 to 17" using "Preventive Care- Initial Office Visit, 0 to 17 Grouping Value Set (2.16.840.1.113883.3.464.1003.101.12.1022)"

S.8. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
There are no patient exclusions.

S.9. **Denominator Exclusion Details** *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*
There are no patient exclusions.

S.10. **Stratification Information** *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*
Not applicable

S.11. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in measure testing attachment)
No risk adjustment or risk stratification
If other:

S.12. **Type of score:**
Rate/proportion
If other:

S.13. **Interpretation of Score** *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*
Better quality = Higher score

S.14. **Calculation Algorithm/Measure Logic** (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*
1. Identify patients who meet the initial population criteria as defined by eCQM logic;
2. Identify and count subset of the initial population that meet denominator criteria as defined by eCQM logic;
3. Identify and count subset of patients in the denominator that meet numerator criteria as defined by eCQM logic.
4. Calculate the performance measure rate: by dividing the number of patients in the numerator population by the number of patients in the denominator population.

Note: the eCQM logic criteria for each population is defined in a computable format in the eCQM specifications provided as an attachment to this submission.

S.15. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF a PRO-PM, identify whether (and how) proxy responses are allowed.

Not applicable; not based on a sample.

S.16. **Survey/Patient-reported data** *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*
IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.
This measure is not based on a survey or instrument.

S.17. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.18.
Electronic Health Record (Only), Other

S.18. **Data Source or Collection Instrument** *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data is collected.)*
IF a PRO-PM, identify the specific PROM(s); and standard methods, modes, and languages of administration.
Not applicable.

S.19. **Data Source or Collection Instrument** *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
No data collection instrument provided

S.20. **Level of Analysis** *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*
Facility

S.21. **Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Clinician Office/Clinic
If other:

S.22. **COMPOSITE Performance Measure** - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*
This is not a composite measure.

2. Validity – See attached Measure Testing Submission Form
NQFXXX_ViralSuppression_BonnieTestingAttachment-636177547742392964.zip,NQFXXX_HIVViralSuppression_MeasureTestingAttatchment.docx

2.1 For maintenance of endorsement
Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. (Do not remove prior testing information – include date of new information in red.)
Yes

2.2 For maintenance of endorsement
Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. (Do not remove prior testing information – include date of new information in red.)
Yes

2.3 For maintenance of endorsement
Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes SDS factors is no longer prohibited during the SDS Trial Period (2015-2016). Please update sections 1.8, 2a2, 2b2, 2b4, and 2b6 in the Testing attachment and S.14 and S.15 in the online submission form in accordance with the requirements for the SDS Trial Period. NOTE: These sections must be updated even if SDS factors are not included in the risk-adjustment strategy. If yes, and your testing attachment does not have the additional questions for the SDS Trial please add these questions to your testing attachment:

What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or

sociodemographic factors) used in the statistical risk model or for stratification by risk (e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care)

What were the statistical results of the analyses used to select risk factors?

Describe the analyses and interpretation resulting in the decision to select SDS factors (e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects)
No - This measure is not risk-adjusted

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b2-2b7)

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number
Measure Title: HIV Viral Suppression
Date of Submission: 12/16/2016
Type of Measure:

☒ Outcome (*including PRO-PM*)	☐ Composite – ***STOP – use composite testing form***
☐Intermediate Clinical Outcome	☐ Cost/resource
☐ Process	☐ Efficiency
☐ Structure	

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.**
- **For outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b6** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b2-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 20 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address sociodemographic variables and testing in this form refer to the release notes for version 6.6 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is

precise. For **PRO-PMs and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b2. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [12]

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). [13]

2b4. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and sociodemographic factors) that influence the measured outcome and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b5. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [16] **differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b7. For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data), analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific

topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.
13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions
15. With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (***must be consistent with data sources entered in S.23***)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☐ administrative claims	☐ administrative claims
☐ clinical database/registry	☐ clinical database/registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☒ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☒ other: Synthetic Bonnie test patients	☒ other: Synthetic Bonnie test patients

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

This measure is a legacy electronic clinical quality measure (eCQM) – an NQF endorsed measure that has been respecified into eMeasures and are currently used in federal quality programs. Per NQF modified testing requirements for legacy eCQMs, the measure was tested in the Bonnie testing tool. Bonnie is designed to validate eCQM specifications (HQMF output and value sets) against the measure's expected behavior for user-developed synthetic test patients.

The synthetic patient bundle used to test this measure was designed to simulate clinically relevant, realistic patient scenarios aligned with the target population for this measure. Full details on the Bonnie synthetic patient bundle used to test this measure are included in the Bonnie testing attachment.

For more information on Bonnie, please visit https://bonnie.healthit.gov/.

1.3. What are the dates of the data used in testing? The Bonnie test environment simulates the year 2012 as the measurement period.

1.4. What levels of analysis were tested? (*testing must be provided for <u>all</u> the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (***must be consistent with levels entered in item S.26***)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☐ hospital/facility/agency
☐ health plan	☐ health plan
☐ other:	☒ other: Synthetic Bonnie test patients

1.5. How many and which <u>measured entities</u> were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

Not applicable. The Bonnie synthetic patient bundle was used to test the measure.

1.6. How many and which <u>patients</u> were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

A test bundle of 34 patients was designed and built within the Bonnie testing tool to evaluate the measure logic. Information documented for each patient within the bundle include:

- Patient name
- Date of birth
- Race
- Ethnicity
- Gender
- Payer

Additional elements contained within the patient profiles as appropriate for testing against expected outcomes include:

- Diagnosis
- Laboratory tests and associated results
- Encounters

The patient bundle's demographics were designed to mimic the HIV/AIDS population, specifically drawing from the patient characteristics collected via the Ryan White HIV/AIDS Program Services Report (RSR).

The breakdown of test bundle demographics for the 34 patients included (represented by number of patients/percentage of bundle): males 23/68%; females 11/32%; American Indian/Alaska Native 1/3%; Asian

1/3%; Black/African American 15/44%; Native Hawaiian/Pacific Islander 0/0%; White 9/26%; Hispanic/Latino 8/24%; younger than 13 1/3%; 13-17 years old 1/3%; 18-24 years old 2/6%; 25-34 years old 6/18%; 35-44 years old 6/18%; 45-54 years old 10/29%; 55-65 years old 6/18%; older than 65 2/6%.
Full details on the Bonnie synthetic patient bundle used to test this measure, including human-readable and QRDA Category 1 format documents for each synthetic patient record, are included in the Bonnie testing attachment.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

The Bonnie patient test deck was used to satisfy all testing requirements for this measure. The testing results are further supported by testing data for the chart-abstracted version of this measure collected through the Health Resources and Services Administration HIV/AIDs Bureau's Ryan White HIV/AIDS Program Services Report.

1.8 What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

Patient sociodemographic variables considered in the analysis of the chart-abstracted version of this measure were included in the eCQM specifications and modeled in the Bonnie patient bundle. These variables included age, race, ethnicity, gender and payer.

2a2. RELIABILITY TESTING
Note*: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)
☐ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)
☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

Currently, there is no performance data available to test the eCQM. However, the chart-abstracted version of this measure has been in use in national quality reporting programs since as early as 2010.

The most recent reliability analysis of the chart-abstracted measure was confirmed according to the methods outlined in a technical report prepared by J.L. Adams for the National Committee for Quality Assurance titled "The Reliability of Provider Profiling: A Tutorial" (RAND Corporation, TR-653-NCQA, 2009). In this context, reliability represents the ability of a measure to confidently distinguish the performance of one physician from another. As discussed in the report: "Conceptually, it is the ratio of signal to noise. The signal in this case is the proportion of variability in measured performance that can be explained by real differences in performance. There are 3 main drivers of reliability; sample size, differences between physicians, and measurement error."

According to this approach, reliability is estimated with a beta-binomial model. The beta-binomial model is appropriate for measuring the reliability of pass/fail measures such as those proposed here. Reliability scores

vary from 0.0 to 1.0, with a score of zero indicating that all variation is attributable to measurement error (noise, or individual accountable entity variance) whereas a reliability of 1.0 implies that all variation is caused by real difference in performance across accountable entities.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

Overall reliability scores (i.e., median of provider-level reliability [R_median], minimum [R_min], maximum [R_max]) by year, and the overall variance between sites, are summarized below.

Overall reliability scores by year, 2010-2014

Year	% suppressed	Var_between	R_median	R_min	R_max
2010	60.6	0.051	0.983	0.290	1.000
2011	64.7	0.046	0.982	0.267	1.000
2012	69.9	0.038	0.979	0.338	1.000
2013	76.1	0.020	0.967	0.211	1.000
2014	80.3	0.013	0.954	0.092	1.000

Reliability scores varied across providers by year. The proportion of providers with reliability greater than or equal to 0.9, 0.8, and 0.7 are shown below.

Distribution of provider-level reliability scores by year, 2010-2014

Year	N	≥0.9 n (%)	≥0.8 n (%)	≥0.7 n (%)
2010	846	764 (90.3)	809 (95.6)	826 (97.6)
2011	811	721 (88.9)	766 (94.5)	786 (96.9)
2012	816	713 (87.4)	775 (95.0)	794 (97.3)
2013	823	657 (79.8)	738 (89.7)	772 (93.8)
2014	813	595 (73.2)	690 (84.9)	751 (92.4)

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

There is no established cut-off for minimum reliability level. Values above 0.7 are considered sufficient to see differences between providers and the mean, and values above 0.9 are considered sufficient to see differences between pairs of providers (RAND Corporation, TR-653-NCQA, 2009).

Each year, the majority of provider-level reliability scores were greater than 0.9, and more than 90% of providers had reliability scores of 0.7 or greater. Therefore, the reliability of viral suppression can be considered to be sufficient to identify real differences in performance across providers. As previously mentioned, sample size is another driver of reliability and likely contributed to the lowest reliability scores (e.g., in 2014 site 8645 had a reliability of 0.21, and reported 3 of 4 patients with a medical visit were virally suppressed). However, median reliability was consistently over 0.95 during 2010-2014 and can help to support the conclusion that the reliability of this measure can be considered very good.

2b2. VALIDITY TESTING

2b2.1. What level of validity testing was conducted? (*may be one or both levels*)

☒ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

☐ **Empirical validity testing**

☒ **Systematic assessment of face validity of <u>performance measure score</u> as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)

2b2.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

The Bonnie testing environment was used to test the validity of the measure logic and data elements. For each Bonnie synthetic patient, an expected measure result was assigned to reflect the expected outcome of the measure given the specific patient scenario and associated data. The synthetic patients were run against the HQMF output loaded into Bonnie, which produces a measure outcome for each patient and evaluates it against the expected outcome. A patient is considered to pass Bonnie testing when the expected outcome matches the actual outcome, e.g. when a patient is expected to be in the numerator population and the computation of the synthetic patient data against the eCQM logic places the patient in the numerator.

In order to achieve a rigorous, clinically relevant test bundle, synthetic patients were designed following the below principles and test areas:

- Clinical relevance. References cited within the chart abstracted measure specification were used to design clinically relevant, realistic patient profiles for the measure's target population. This approach ensured the eCQM logic maintained alignment with the clinical intent of the chart abstracted measure.
- 100% logic coverage: The resulting bundle of synthetic patients collectively includes all data elements and conditions logic that are specified within the measure logic, including at least one patient evaluating against each measure population pathway. Fully testing the measure logic increases test rigor and mitigates risk of unexpected outcomes.
- Edge case testing. Edge cases refer to those data elements that test the upper or lower boundary of measure logic conditions, e.g. a diagnosis starting on the latest qualifying date or an HIV viral load result equal to the highest qualifying value. Edge cases are designed to test each edge that exists within each measure population.
- Negative testing. Negative testing involves use of test cases do not evaluate positively against measure logic, but are otherwise clinically relevant and realistic, e.g. scenarios where an HIV diagnosis was not documented or an HIV viral load was performed without a documented result. Negative testing further validates measure logic by accurately evaluating patients against expected outcomes and simulating the effect of missing data on measure results.

In addition to Bonnie testing, the measure specifications were reviewed independently by three eCQM experts to confirm the logic was syntactically correct, using appropriate and current versions of the eCQM standards and terminologies, and consistent with the intent of the chart-abstracted measure.

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Bonnie testing results provide logic coverage and passing rates. The synthetic bundle reached 100% coverage, confirming each logic pathway was tested. The results also showed 100% passing rate, confirming all synthetic patients performed as expected.

Full details on Bonnie testing results are contained in the Bonnie testing attachment. The attachment includes a human-readable (HTML) summary document that lists each patient within the bundle and its passing status against expected measure outcomes. The attachment also includes a summary spreadsheet for the synthetic patient bundle which lists each patient, associated demographics, expected and actual measure population outcomes, and which portions or each measure population logic the patient meets expectations for.

2b2.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

The results of measure logic testing through use of Bonnie provided confidence in the measure logic accurately representing the clinical intent and alignment with the chart abstracted measure.

2b3. EXCLUSIONS ANALYSIS (FOR MEASURS WITH EXCLUSIONS --- gap in visits and medical visit frequency)
NA ☒ no exclusions — *skip to section 2b4*

2b3.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

Not applicable.

2b3.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

Not applicable.

2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

Not applicable.

2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b4.1. What method of controlling for differences in case mix is used?
☒ **No risk adjustment or stratification**
☐ **Statistical risk model with risk factors**
☐ **Stratification by risk categories**
☐ **Other,**

2b4.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

Not applicable.

2b4.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

The Ryan White HIV/AIDS Program provides a comprehensive system of care that includes primary medical care and essential support services for people living with HIV who are uninsured or underinsured. The Program works with cities, states, and local community-based organizations to provide HIV care and treatment services to more than half a million people each year. The Program reaches approximately 52% of all people diagnosed with HIV in the United States.

As indicated in data presented earlier, the Ryan White HIV/AIDS Program is a public health, safety net program providing care to a high proportion of racial/ethnic minority, transgender, unstable housing, and low income people living with HIV. Many of people served by the Ryan White HIV/AIDS Program represent sociodemographics factors incorporate in risk adjusting models by many measures stewards. As a result, the Ryan White HIV/AIDS Program does not adjust for risk in its performance measures. Rather, it is a fundamental aspect of the Ryan White HIV/AIDS Program to identify disparities and work to improve quality of care for subpopulations. Additionally, this measure is not used for pay-for-performance, bonuses, or penalties.

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or sociodemographic factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)

Not applicable.

2b4.4a. What were the statistical results of the analyses used to select risk factors?

Not applicable.

2b4.4b. Describe the analyses and interpretation resulting in the decision to select SDS factors (e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects)

Not applicable.

2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

Not applicable.

If stratified, skip to 2b4.9

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:** Not applicable.

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*): Not applicable.

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves: Not applicable.

2b4.9. Results of Risk Stratification Analysis: Not applicable.

2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*) Not applicable.

2b4.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

Not applicable.

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

The chart-abstracted version of this measure has been in use since 2010. To examine meaningful differences in performance, we examined the distribution of the proportion of patients with viral suppression across providers, by year. Performance scores were broken into the bottom 10% and top 90% providers to better characterize the gaps that remain across providers. Moreover, performance scores were examined with respect to NHAS 2020 Indicator 6: increase the percentage of persons with diagnosed HIV infection who are virally suppressed to at least 80 percent.

2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

	% Patients with viral suppression across providers					Providers achieving ≥80% suppression		
Year	Mean	SD	Median	10th %ile	90th %ile	N	n	%
2010	60.6	23.8	67.8	19.5	82.8	846	145	17.1
2011	64.7	22.1	71.4	31.9	84.9	811	207	24.5
2012	69.9	20.3	75.6	40.2	88.0	816	277	32.7
2013	76.1	17	80.7	57.1	90.2	823	435	51.4
2014	80.3	15.5	84.2	65.0	93.1	813	530	65.2

2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

The table above demonstrates meaningful variability across providers, allowing for the identification of meaningful differences across sites. Specifically, the measure is able to detect providers with better or worse than median performance scores. In 2014, the bottom 10% of providers had viral suppression rates of 65.0% or lower; the top 90% of providers had viral suppression rates of 93.1% or higher. While this gap appears to be

narrowing over time, a meaningful difference of 28.1 percentage points remains, demonstrating the value of the measure in identifying sites based on poor performance relative to the top performers.

Provider-level performance differences observed in the table above also underscore improvements in the proportion of patients with viral suppression in achieving 80% viral suppression. In 2014, of 813 providers, 530 (65.2%) had at least 80% of patients reach viral suppression. Additionally, the overall percentage of patients with viral suppression was 80.3%; however, given the large population that the RWHAP serves, even the poorest performing sites (e.g., bottom 10%) represent a substantial number of patients.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Note: *This item is directed to measures that are risk-adjusted (with or without SDS factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without SDS factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b6.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

Not applicable.

2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

Not applicable.

2b6.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

Not applicable.

2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b7.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

The HQMF standard specifies that if data are unknown or missing, they shall fail the criterion. This constraint embodies the notion that absence of evidence is evidence of absence, i.e. data not present in a structured field

from which the measure draws will not be considered for measure calculation. In certain cases, missing data may have no impact on the measure outcome for a given patient. For example, a data element used in a series of OR statements will not impact the measure outcome if another data element in the OR statement is present and meets all other defined constraints.

2b7.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

The Bonnie synthetic patient bundle includes scenarios for missing data elements, which are a form of negative testing. All Bonnie synthetic patients with missing data performed according to the HQMF standard specification and as expected.

2b7.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

Please see response for question 2b7.1 above.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. **Data Elements Generated as Byproduct of Care Processes.**
Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)
If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. **To what extent are the specified data elements available electronically in defined fields** (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.
ALL data elements are in defined fields in electronic health records (EHRs)

3b.2. **If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources.** For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).
Not applicable.

3b.3. **If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.**
Attachment: NQFXXX_HIVViralSuppression_Feasibility_Scorecard_v1.0-636177547747228995.xlsx

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. **Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**
IF a PRO-PM, consider implications for both individuals providing PRO data (patients, service recipients, respondents) and those whose performance is being measured.
Not applicable.

3c.2. **Describe any fees, licensing, or other requirements to use any aspect of the measure as specified** *(e.g., value/code set, risk model, programming code, algorithm)*.
The measure specifications contain limited proprietary codes for convenience. Users of CPT(R) should obtain all necessary licenses from the owners of these code sets.
The use of SNOMED Clinical Terms(R) requires a Unified Medical Language System (UMLS) license. These licenses are freely available, from the National Library of Medicine.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. **Current and Planned Use**
NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
Public Reporting	
Public Health/Disease Surveillance	
Payment Program	
Quality Improvement (external benchmarking to organizations)	
Quality Improvement (Internal to the specific organization)	

4a.1. **For each CURRENT use, checked above (update for maintenance of endorsement), provide:**

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

Ryan White HIV/AIDS Program
Sponsor: Federal government
Geographic area: Nationwide
Accountable entities: Approximately 600 Ryan White HIV/AIDS Program grant recipients and their providers
Patients: Approximately 316,000 patients

Medicaid Adult Core Set
Sponsor: Federal government
Geographic area: Nationwide
Accountable entities: State Medicaid programs
Patients: Unknown

Physician Quality Report System and Value Based Modifier
Sponsor: Federal government
Geographic area: Nationwide
Accountable entities: Physicians and practitioners
Patients: Unknown

Merit-Based Incentive Payment System
Sponsor: Federal government
Geographic area: Nationwide
Accountable entities: Physicians, Physician Assistant, Nurse Practitioner, and Clinical Nurse Specialist
Patients: Unknown

National HIV/AIDS Strategy
Sponsor: Federal government
Geographic area: Nationwide
Accountable entities: Federal agencies and service providers
Patients: All people living with HIV in the United States

4a.2. **If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons?** (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)
N/A

4a.3. **If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement.** (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)
This measure is current under consideration for the Centers for Medicare and Medicaid Merit Based Incentive Payment System (MIPS).

Improvement
Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b. **Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)**
If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial

endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.
HIV viral suppression has been improving in the United States since the first release of publically available data. The Ryan White HIV/AIDS Program served more than 300,000 unduplicated patients annually between 2010-2014 across 2,000+ grant recipients and subrecipients. The Ryan White HIV/AIDS Program has experienced a 20-point increase in viral suppression from 61.8% in 2010 to 80.3% in 2014. Viral suppression has increased across all demographic groups and subpopulations.

4c. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4c.1. **Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.**
N/A

4c.2. **Please explain any unexpected benefits from implementation of this measure.**
N/A

4d1.1. **Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.**
How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.
Starting in 2015, Health Resources and Services Administration began releasing December 1st – World AIDS Day – an annual data report (Ryan White HIV/AIDS Program Annual Client-Level Data Report) that contains data similar to those presenting in the report. Building upon the success of the state profiles (http://hab.hrsa.gov/stateprofiles/), Health Resources and Services Administration worked diligently to release the annual data report in the same year it was collected (collected in April and released in December of the same year). The report is publically available on the Health Resources and Services Administration website (http://hab.hrsa.gov/data/data-reports) and is released via an accompanying webinar (recorded and archived). A supplemental report exploring data for the eligible metropolitan areas and transitional grant areas and youth/young adults has been released as well as slides sets for fact sheets by program and population, special populations (http://hab.hrsa.gov/publications/hivaids-bureau-fact-sheets), and infographics (contained in fact sheets). Additionally, grant recipient level reports are prepared and disseminated to all Ryan White HIV/AIDS Program grant recipients.

HRSA is releasing a quality module where grant recipients can voluntarily report numerator, denominator, and performance scores for a portfolio of measures. Grant recipients will be able to benchmark their performance based on a number of patient demographic and organizational factors. This measure will be included in the measure portfolio.

4d1.2. **Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.**
Starting in 2015, Health Resources and Services Administration began releasing December 1st – World AIDS Day – an annual data report (Ryan White HIV/AIDS Program Annual Client-Level Data Report) that contains data similar to those presenting in the report. Building upon the success of the state profiles (http://hab.hrsa.gov/stateprofiles/), Health Resources and Services Administration worked diligently to release the annual data report in the same year it was collected (collected in April and released in December of the same year). The report is publically available on the Health Resources and Services Administration website (http://hab.hrsa.gov/data/data-reports) and is released via an accompanying webinar (recorded and archived). A supplemental report exploring data for the eligible metropolitan areas and transitional grant areas and youth/young adults has been released as well as slides sets for fact sheets by program and population, special populations (http://hab.hrsa.gov/publications/hivaids-bureau-fact-sheets), and infographics (contained in fact sheets). Additionally, grant recipient level reports are prepared and disseminated to all Ryan White HIV/AIDS Program grant recipients.

4d2.1. **Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.**
Describe how feedback was obtained.
Antidotal feedback has been received from Ryan White HIV/AIDS Program grant recipients and subrecipients regarding the feasibility and usefulness of the data presented in the Ryan White HIV/AIDS Program Annual Client-Level Data Report. Significant feedback has been provided about the timeliness and expansions of the data release. Grant recipient report using the data for

benchmarking their program, setting goals/targets, and gaining a fuller understanding of all aspects of the Ryan White HIV/AIDS Program (i.e. other regions of the country). Grant recipients and subrecipients have also requested additional analyses. Health Resources and Services Administration responded with supplemental reports (Ryan White HIV/AIDS Program Supplemental Client-Level Data Report, Eligible Metropolitan Areas and Transitional Grant Areas; special population reports); slide decks for the overall client population and special populations; grant recipient reports; and infographics – all of which will be updated and released annually. Health Resources and Services Administration plans to release additional analyses and special reports this year based on feedback from Ryan White HIV/AIDS Program grant recipients and subrecipients.

4d2.2. **Summarize the feedback obtained from those being measured.**
See 4d2.2

4d2.3. **Summarize the feedback obtained from other users**
Ryan White HIV/AIDS Program national partners (national organizations that represent grant recipients, subrecipients, and patients) has provided antidotal feedback regarding the timeliness, feasibility, and usability of the release of the Ryan White HIV/AIDS Program Annual Client-Level Data Report, supplemental reports, slide decks, fact sheets, and infographics. The national partners encourage the continued release of the data in all its formats.

4d.3. **Describe how the feedback described in 4d.2 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.**
During the initial development of the measure, formal feedback was gathered. The measures were modified during the development phase and have not been modified since. A concerted effort was made to develop a measure that would likely stand the test of time from a scientific, clinical, and patient perspective. On an annual basis, the measure is review for clinical relevance, change in scientific acceptability, and consistency with guidelines. This measure has not been modified as a result of the annual reviews. Additionally, this measure is used by a number of measurement programs and strategies. Each of those programs require a separate annual review. No modifications have been made for those programs.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures
Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.
Yes
5.1a. **List of related or competing measures (selected from NQF-endorsed measures)**

5.1b. **If related or competing measures are not NQF endorsed please indicate measure title and steward.**
0405 HIV/AIDS: Pneumocystis Jiroveci Pneumonia (PCP) Prophylaxis
0409 HIV/AIDS: Sexually Transmitted Disease Screening for Chlamydia, Gonorrhea, and Syphilis
2079 HIV Medical Visit Frequency
2080 Gap in HIV Medical Visits
2082 HIV Viral Suppression
2083 Prescription of HIV Antiretroviral Therapy
3211 Prescription of HIV Antiretroviral Therapy
3010 HIV Medical Visit Frequency

5a. Harmonization of Related Measures
The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. **If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):**
Are the measure specifications harmonized to the extent possible?
No

5a.2. **If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.**
Harmonization exists with all measures except 405 and 409. Plan to harmonize with 405 and 409.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. **If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):**
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
None

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

No appendix **Attachment:**

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** Health Resources and Services Administration - HIV/AIDS Bureau
Co.2 **Point of Contact:** Marlene, Matosky, mmatosky@hrsa.gov, 301-443-0798-
Co.3 **Measure Developer if different from Measure Steward:** Health Resources and Services Administration - HIV/AIDS Bureau
Co.4 **Point of Contact:** Marlene, Matosky, mmatosky@hrsa.gov, 301-443-0798-

Additional Information

Ad.1 **Workgroup/Expert Panel involved in measure development**
Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
The work group members determined the measure concepts, identified the data elements, voted on the final measures, and assessed the face validity of the measures.
Bruce Agins, NYS DOH AIDS Institute, New York, NY
Judy Bradford, Fenway Community Health, Boston, MA
John Brooks, CDC, Atlanta, GA
Karen Brudney, Columbia University, New York, NY
Laura Cheever, HRSA HAB, Rockville, MD
Nikki Cockern, Wayne State University, Detroit, MI
Chinazo Cunningham, Montefiore Medical Center, New York, NY
William Cunningham, UCLA, Los Angeles, CA
Julie Dombrowski, University of Washington, Seattle, WA
Edward Gardner, Denver Health, Denver, CO
Elvin Geng, UCSF, San Francisco, CA
Thomas Giordano, Baylor College of Medicine, Houston, TX

Barb Gripshover, Cleveland ACT UP, Cleveland, OH
Deborah Konkle Parker, University of Mississippi, Jackson, MS
Tim Long, Alliance Chicago, Chicago, IL
Cheryl Lynn-Besch, Louisiana State University, New Orleans, LA
Julio Marrero, COSSMA, San Juan, PR
Brian Montague, Brown University, Providence, RI
Karam Mounzer, Philadelphia Fight, Philadelphia, PA
Michael Mugavero, University of Alabama, Birmingham, AL
Sylvia Naar King, Wayne State University, Detroit, MI
Josiah Rich, Brown University, Providence, RI
Allan Rodriguez, Miami University, Miami, FL
Amy Sitapati, UCSD, San Diego, CA
Avnish Tripathi, University of South Carolina, Charleston, SC
Gregory Winstead, Christian Community Health Center, Chicago, IL

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 **Year the measure was first released:** 2011

Ad.3 **Month and Year of most recent revision:** 05, 2016

Ad.4 **What is your frequency for review/update of this measure?** Annual

Ad.5 **When is the next scheduled review/update for this measure?** 05, 2017

Ad.6 **Copyright statement:** None

Ad.7 **Disclaimers:** None

Ad.8 **Additional Information/Comments:** None