NATIONAL QUALITY FORUM

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 #: 2856

Measure Title: Pharmacotherapy Management of COPD Exacerbation

Measure Steward: National Committee for Quality Assurance

Brief Description of Measure: This measure assesses the percentage of COPD exacerbations for patients 40 years of age and older who had an acute inpatient discharge or ED visit on or between January 1-November 30 of the measurement year and who were dispensed appropriate medications. Two rates are reported:

1. Dispensed a systemic corticosteroid (or there was evidence of an active prescription) within 14 days of the event.

2. Dispensed a bronchodilator (or there was evidence of an active prescription) within 30 days of the event.

Developer Rationale: This measure assesses whether patients who had a hospitalization or an emergency department (ED) visit for a COPD exacerbation were provided appropriate medication (systemic corticosteroids and bronchodilators) to treat symptoms and prevent future exacerbations. The improvement in quality envisioned by the use of this measure is to increase the use of systemic corticosteroids and bronchodilators following a COPD exacerbation in order to improve patient outcomes, such as shorten recovery time, improve lung function and reduce the risk of early relapse, treatment failure, and length of hospital stay.

Numerator Statement: Numerator #1 (Systemic corticosteroids): The number of patients dispensed a prescription for a systemic corticosteroid on or 14 days after the Episode Date. Count systemic corticosteroids that are active on the relevant date.

Numerator #2 (Bronchodilators): The number of patients dispensed a prescription for a bronchodilator on or 30 days after the Episode Date. Count bronchodilators that are active on the relevant date.

*The Episode Date is the date of service for any acute inpatient discharge or ED claim/encounter during the 11-month intake period with a principal diagnosis of COPD.

Denominator Statement: All patients age 40 years or older as of January 1 of the measurement year with a COPD exacerbation as indicated by an acute inpatient discharge or ED encounter with a principal diagnosis of COPD.

Denominator Exclusions: This measure excludes patients who use hospice services, and patients with nonacute inpatient stays.

Measure Type: Process

Data Source: Claims

Level of Analysis: Health Plan

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a *structure, process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Evidence Summary or **Summary of prior review in [year]**

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- This measure is a process measure which assesses the percentage of patients 40 years of age and older with a new diagnosis of COPD or newly active COPD, who received appropriate spirometry testing to confirm the diagnosis.
- Developer provided a logic model that articulates the connection between acute COPD exacerbation, use of corticosteroids and inhaled SABAs, and improved health outcomes.
- Evidence from the 2015 submission was updated from the 2015 GOLD Guidelines for COPD to the 2020 GOLD Guidelines for COPD.
    - "Short-acting inhaled beta2-agonists, with or without short-acting anticholinergics, are recommended as the initial bronchodilators to treat an acute exacerbation (Evidence C)." No reference to studies; two other guidelines and position papers (NICE and ATS) are referenced, although developer alludes to the synthesis of RCTs from these other papers.

- "Systemic corticosteroids can improve lung function (FEV1), oxygenation and shorten recovery time and hospitalization duration. Duration of therapy should not be more than 5-7 days (Evidence A)." Developer states that "the body of evidence for the recommendation for systemic corticosteroids is based on eight studies, four of which are RCTs".

- Evidence from the 2015 submission was updated from the 2013 ICSI Guidelines. ICSI guidelines have been retired and ICSI has endorsed the VA/DoD Guideline for COPD from 2014.
  - "We recommend prescribing inhaled short-acting beta 2-agonists (SABAs) to patients with confirmed COPD for rescue therapy as needed. (Strong For)." Evidence for the recommendation for SABAs is based on three studies, including a systematic review with 13 trials.
  - "We recommend inhaled tiotropium as first-line therapy for patients with confirmed, stable COPD who have respiratory symptoms (e.g., dyspnea, cough) and severe airflow obstruction (I.e., post bronchodilator FEV1 < 50%) or a history of COPD exacerbations. (Strong For)." Evidence for the recommendation for tiotropium is based on a Cochrane Database systematic review.
  - "For acute COPD exacerbations, we recommend a course of systemic corticosteroids (oral preferred) of 30-40mg prednisone equivalent daily for 5-7 days (Strong For)." Evidence for the recommendation for systemic corticosteroids is based on four studies, including a large RCT and a large systematic review.

Questions for the Committee:

If the developer provided updated evidence for this measure:

- The evidence provided by the developer is updated, directionally the same, and stronger compared to that for the previous NQF review. Does the Committee agree there is no need for repeat discussion and vote on Evidence?

Guidance from the Evidence Algorithm

Per page 15 of the 2019 NQF Measure Evaluation Criteria – Clinical Evidence Algorithm: 1. Measure assesses a process → 3. It is based on a systematic review and grading →4. Summary of QQC included → 5a. Quantity: Moderate/High; Quality: High; Consistency: High → Rate as high

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

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

Maintenance measures – increased emphasis on gap and variation

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

- Developer provided analysis of HEDIS data for the most recent measurement years of Medicare, Medicaid and Commercial plans.
  - Commercial

    YEAR | N Plans | Mean Denominator Size per plan

    2016 | 263 | 179

    2017 | 254 | 178

    2018 | 255 | 165
  - Medicaid

YEAR | N Plans | Mean Denominator Size per plan

2016 | 200 | 863

2017 | 201 | 885

2018 | 189 | 937

- Medicare

  YEAR | N Plans | Mean Denominator Size per plan

  2016 | 385 | 681

  2017 | 390 | 750

  2018 | 388 | 707

- Performance for each product line trend toward improvement, but little year-over-year change between 2017 – 2018. Significant gaps remain for systemic corticosteroids:
  - Systemic Corticosteroids – Commercial Rate

    YEAR| MEAN | ST DEV

    2016 | 69.7% | 9.5%

    2017 | 75.5% | 7.6%

    2018 | 74.2% | 8.0%
  - Systemic Corticosteroids – Medicaid Rate

    YEAR| MEAN | ST DEV

    2016 | 66.0% | 11.6%

    2017 | 68.3% | 11.3%

    2018 | 68.4% | 11.7%
  - Systemic Corticosteroids – Medicare Rate

    YEAR| MEAN | ST DEV

    2016 | 67.7% | 10.6%

    2017 | 70.4% | 7.8%

    2018 | 71.4% | 9.1%
- Very little year-over-year change in bronchodilators in 2017 over 2018, but performance gap remains:
  - Bronchodilators – Commercial Rate

    YEAR| MEAN | ST DEV

    2016 | 75.8% | 8.7%

    2017 | 80.1% | 7.0%

    2018 | 79.6% | 7.4%
  - Bronchodilators – Medicaid Rate

    YEAR| MEAN | ST DEV

    2016 | 80.7% | 9.7%

    2017 | 81.4% | 10.1%

    2018 | 81.4% | 10.8%
  - Bronchodilators – Medicare Rate

    YEAR| MEAN | ST DEV

2016 | 77.9% | 9.1%

2017 | 79.6% | 8.2%

2018 | 79.8% | 8.5%

Disparities

- The CMS Office of Minority Health in collaboration with the RAND Corporation produces an annual report: CMS Racial, Ethnic, and Gender Disparities in Health Care in Medicare Advantage.
    - In the 2019 CMS report Asian/Pacific Islander patients with a COPD exacerbation were significantly more likely than White patients to receive a systemic corticosteroid with 14 days. Hispanic patients were significantly less likely than White patients to receive the same. There was no significant difference between White and Black patients.
    - Regarding the the receipt of a bronchodilator with 30 days of a COPD exacerbation, Asian/Pacific Islander patients and Black patients were more likely to receive than White patients, while Hispanic patients were less likely.
- Developer references several studies that describe disparities present in the quality of care for COPD as well as disease burden.
    - Patients who were 75 years and older, or African American, were less likely to receive comprehensive care, compared to younger patients, or white patients, respectively (Lindanauer, et al, 2006).
    - African Americans have been shown to be more likely than whites to report having a COPD exacerbation that required hospitalization in the previous year (Han, et al, 2011).
    - Lower education and lower income, often used as proxies for lower socioeconomic status, have been shown to be correlated with a greater risk of COPD exacerbations (Eisner, et al 2009).

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?
- (Performance improvement over time is reviewed under "Usability".)

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

Comments:

** I am not aware of any new evidence not already discussed in the submission. The evidence for this measure appears to be stronger than it was when previously endorsed by the NQF.

** The evidence relates well and applies directly. They included and updated the evidence to reflect the 2020 GOLD guideline for COPD and strong supportive Systematice Review.

** As a process measure that is in maintenance evaluation, endorsement was given based on prior high levels of evidence. The evidence has since been updated with more recent publications and guidelines that strengthen the recommendations rather than completely change them.

** This is a process measure that assesses the percent of COPD exacerbations for patients ≥ 40 years old with a hospitalization or ED visit from JAN 1 - NOV 30, for whom medications were dispensed: (1) systemic corticosteroids on or after 14 days after episode date; (2) bronchodilator on or after 30 after episode date.

Evidence indicates prescribing these medications appropriately is associated with improved outcomes including shortened recovery, reduced risk for relapse and readmission.

** Evidence for the treatment of COPD exacerbations with short acting inhaled beta agonists and systematic corticosteroids for 5-7 days continues to be supported by the 2020 GOLD clinical guideline review.

** Evidence is updated and rated High, including the new 2020 GOLD guidelines.

** I am not aware of any new information. My only question would be to ask if the group needs to discuss the logic model provided by the developer.

** I'm not aware of any unmentioned studies that aren't cited in the submission. The most robust updated evidence provided is the 2020 GOLD guidelines

1b. Performance Gap

Comments:

** The data provided demonstrates substantial variation between 10th percentile and 90th percentile performance as well as variation by plan type with Medicaid under performing Medicare and commercial plans.

** Yes warrants a national performance measure. Yes, current 2016-2018 performance measured. A less than optimal performance with systemic coricosteriods use show a significant gap.

** The report notes that there are still considerable performance gaps for both systemic corticosteroid and bronchodilator dispensing. As there is an apparent less than optimal performance on this measure, there needs to be a discussion on why gaps remain and methods to begin addressing.

** Opportunities for improvement persist

** The maintenance process measurement tracks whether patients of 40 y/o or older treated in hospital or in an emergency depertment setting received prescriptions for a systemic corticosteroid within 14 days of the Episode Date and a prescription for a bronchodilator within 30 days of the Episode Date. These two measurements were determined from HEDIS data from commercial, Medicaid, and Medicare health plans in years 2016-18. Performance scores in all health care group categories and in across the three years varied from the upper 60s to low 80s percentiles.

** There is a performance gap to demonstrate a gap in care.

** Yes. It showed gaps suggesting less than optimal care in multiple instances.

** Developer provided analysis of HEDIS data for the most recent measurement years of Medicare, Medicaid and Commercial plans

Disparities:

Comments:

** The measure developer lacks data to identify disparities based on race, ethnicity or language and instead refers to published studies that outcomes vary based on race and ethnicity, which Hispanic patient so be more likely to be under-treated.

** Yes disparities noted through the RAND report. Differences in care by ethnic groups in use of systemic corticosteriod prescribing. Also, quality of care for COPD vary with SDOH.

** Some data was provided on racial subgroups with COPD and the likelihood of receiving systemic corticosteroids and/or bronchodilators. It was not readily apparent how significant the differences would be in impacting measure reporting.

** The developer does not collect data on disparities. Published data indicate disparities persist based on race/ethnicity, age, gender.

** NCQA does not currently collect performance stratified by race, ethnicity, or language. The 2019 CMS Office of Minority Health study did show differences between racial groups in the percentages receiving prescriptions for systemic corticosteroids and bronchodilators (anticholinergics, beta-2 agonists, & combinations) as treatments for COPD exacerbations.

** Disparities are documented, but this measure does not address these.

** Yes. Clearly shows a difference in treatment between population groups.

** Developer references several studies that describe disparities present in the quality of care for COPD as well as disease burden.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

2b. Validity: Testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability; Missing Data

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

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. For maintenance measures – less emphasis if no new testing data provided.

Validity

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. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: Primary Care and Chronic Illness project team staff
Evaluation of Reliability and Validity: Link A (Project Team staff)

Reliability:

- The developer had the following results for reliability using HEDIS health plan data in 2018:
  - Commerical: the mean reliability result for bronchodilators and systemic corticosteroids was 0.66/0.65, however ranged 0.41 – 0.89/0.43 – 0.88 in the 10^{th} – 90^{th} percentiles.
  - Medicare: the mean reliability result for bronchodilators and systemic corticosteroids was 0.86/0.81, however ranged from 0.62 – 0.98/0.53 – 0.97 in the 10^{th} – 90^{th} percentiles.
  - Medicaid: the mean reliability result for bronchodilators and systemic corticosteroids was 0.94/0.94, with consistently high range 0.81 – 0.99/0.84 – 0.99 in the 10^{th} – 90^{th} percentiles.

Validity:

- The developer did updated validity testing from the 254 Commercial health plans, 390 Medicare plans and 201 Medicaid health plans that submitted data on this measure to HEDIS in 2017. The developer conducted performance measure score validity testing via health plan level Pearson correlation analysis between the measure's two components, as well as statin therapy measures.
    - Pearson correlation coefficients between dispensing systemic corticosteroids and bronchodilators was 0.45, 0.52, and 0.82 for Medicare, Commercial and Medicaid plans, respectively.
    - Performance on Pearson correlation coefficients between these two indicators and statin therapy measures ranged from 0.25 – 0.68.
    - This suggests mostly moderate correlations, which was the result hypothesized by the devolper and suggests some comparability in the quality constructs of the measure indicators.
- The developer previously did face validity by three expert panels in the 2016 submission of the measure.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- The staff is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The staff is satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss validity?

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

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability – Specifications

Comments:

** The measure specifications are clearly defined and their reliability appears sound. I have no concerns with definitions, codes, calculations, etc. Since this is based on standard claims, its consistent implementation is not a concern.

** Agree that the Reliability-Specifications is consistant.

** I have no concerns about the ability to implement this measure.

** Would wonder about the timeframe stipulated for prescribing of corticosteroids and bronchodilators, and the rationale underlying this. Also, such medication may be dispensed directly in ED, are such samples captured via this measure?

** Measure score reliability continues to be tested for participating health care plans.

** Reliability testing appears adequate for the measure.

** I have no concerns with the measure but do have concerns about compliance.

** The measure should be able to be implemented

2a2. Reliability – Testing

Comments:

** No concerns. The data presented by the measure developer demonstrates good reliability.

** No concerns

** I have no concerns about the reliability of the measure.

** Reliability acceptable

** No

** No.

**No

** No significant concerns

2b1. Validity – Testing

Comments:

** The measure developer did an extensive validity testing in 2004 to demonstrate data consistency between administrative and medical records. This was updated in 2017 to measure correlation with dispensing of bronchodilator and systemic corticosteroids. The developer also tested the correlation between prescribing of statins for COPD patient with Cardiovascular disease. I do not have concerns with the validity testing results.

** No concerns

** I have no concerns about the results of the validity testing.

** See above.

** Construct validity testing from 2019 continued to confirm the hypotheses that the prescribing of a systemic corticosteroid and the prescribing of a bronchodilator were significantly correlated. The correlation between appropriate statin therapy and appropriate medication treatment for COPD exacerbations (for selected patients) also continued in the 2019 validity analysis.

**No

**No

**No concerns

2b4-7. Threats to Validity

Comments:

** The measure developer provides evidence that randomly selected health plans at opposite ends of the interquartile range, which varied by 7-13%, were statistically different based on a t-test, suggesting that meaningful differences in quality exists across health plans on performance for both the use of Bronchodilators and Systemic Corticosteriods. Comparability is not required for this measure. Missing data appears to be managed appropriately by the measure developer's audit process.

** All bases on the HEDIS requirement and audited well.

** The analyses indicates that there is a quantifiable number of patients associated with the performance gap found between the 25th and 75th percentile performance rates.

** Developer included updated validity data with good validity scores.

** Meaningful differences in performance analysis was continued for each drug category and within commercial, Medicare, and Medicaid health care plans and the results continued to be statistically significant. NCQA auditing of HEDIS data was continued, not showing significant missing data or bias.

** Validity issues were covered appropriately.

** I am not aware of any potential threats to validity.

** No significant concerns.

2b2-3. Other Threats to Validity

2b2. Exclusions

2b3. Risk Adjustment

Comments:

** There are no exclusions. Risk adjustment is not required.

** Exclusions consistent with requirement and no risk adjustment.

** There was no risk adjustment or stratification for this measure.

** Hospice patients were excluded.

** Exclusions are explained and reasonable. Risk adjustment was not done.

** None

** I agree with the findings of the developer and the reviewers.

** Exclusion of hospice patients seems appropriate

Scientific Acceptability Evaluation (NQF Project Team Staff)

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 2856

Measure Title: Pharmacotherapy Management of COPD Exacerbation

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☒ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☐ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☐ **Facility** ☒ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

***NOTE**: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**

   No concerns. Per developer, there are no updates to the specifications since the last measure update (2016).

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure**
   ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical <u>VALIDITY</u> testing** of <u>patient-level data</u> conducted?

   ☐ **Yes** ☐ **No** N/A -score level reliability testing was conducted.
6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

   The developer did updated reliability testing in the maintenance using HEDIS health plan performance data for 2018. The developer conducted performance measure score reliability testing by using a beta-binomial model (i.e. signal to noise).
7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   The developer had the following results for reliability using HEDIS health plan data in 2018:

   - Commerical: the mean reliability result for bronchodilators and systemic corticosteroids was 0.66/0.65, however ranged 0.41 – 0.89/0.43 – 0.88 in the 10^{th} – 90^{th} percentiles.
   - Medicare: the mean reliability result for bronchodilators and systemic corticosteroids was 0.86/0.81, however ranged from 0.62 – 0.98/0.53 – 0.97 in the 10^{th} – 90^{th} percentiles.
   - Medicaid: the mean reliability result for bronchodilators and systemic corticosteroids was 0.94/0.94, with consistently high range 0.81 – 0.99/0.84 – 0.99 in the 10^{th} – 90^{th} percentiles.

   For signal to noise, a minimum reliability score of 0.7 is used to indicate sufficient signal strength to discriminate performance between accountable entities. The testing suggests that indicators within this measure have good reliability between 0.7 and 1.0 for Medicare and Medicaid, but not quite for Commercial lines where low-moderate reliability can be claimed.
8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)
9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

☒ **Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

    ☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

    ☐ **Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

    Reliability methodology and results are appropriate and yielded good reliability scores for Medicare and Medicaid business lines, but low-moderate for Commercial plans.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Submission document: Testing attachment, section 2b2.

    This measure excludes hospice patients and nonacute inpatient stays, neither of which carry significant concerns.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4.

    The developer indicates there is variation in performance. To determine if this difference is statistically significant, the developer calculated an independent sample t-test of the performance difference between two randomly selected plans at the 25th and 75th percentile. The results provided by the developer indicate there is a 7 – 13% gap in performance between the 25th and 75th percentile performing plans across the different age ranges and product lines.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    Submission document: Testing attachment, section 2b5.
    N/A

15. **Please describe any concerns you have regarding missing data.**

    Submission document: Testing attachment, section 2b6.

    No concerns with missing data. In 2018, NCQA auditors did not find any missing data sources for any of the health plan data submissions.

    The developer noted HEDIS addresses missing data in a structured way through its audit process. If a data source is found to be missing data, and the issues cannot be rectified, the auditor will assign a "materially biased" designation to the measure for that reporting plan, and the rate will not be used. Once measures are added to HEDIS, NCQA conducts a first-year analysis to assess the feasibility of the measure when widely implemented in the field. This analysis includes an assessment of how many plans report valid rates vs. rates that are materially biased (or have other issues, such as small denominators). These considerations are weighed in the deliberation process before measures are approved for public reporting.

16. **Risk Adjustment** N/A-no risk adjustment done for this process measure

16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

16c. **Social risk adjustment:** N/A

16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

16d.**Risk adjustment summary:** N/A

16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

N/A

VALIDITY: TESTING

17. **Validity testing level:** ☒ **Measure score** ☐ **Data element** ☐ **Both**

18. **Method of establishing validity of the measure score:**

☒ **Face validity**

☒ **Empirical validity testing of the measure score**

☐ **N/A (score-level testing not conducted)**

19. **Assess the method(s) for establishing validity**

Submission document: Testing attachment, section 2b2.2

- The developer did updated validity testing from the 254 Commercial health plans, 390 Medicare plans and 201 Medicaid health plans that submitted data on this measure to HEDIS in 2017. The developer conducted performance measure score validity testing via health plan level Pearson correlation analysis between the measure's two components, as well as statin therapy measures.
    - Pearson correlation coefficients between dispensing systemic corticosteroids and bronchodilators was 0.45, 0.52, and 0.82 for Medicare, Commercial and Medicaid plans, respectively.
    - Performance on Pearson correlation coefficients between these two indicators and statin therapy measures ranged from 0.25 – 0.68.
    - This suggests mostly moderate correlations, which was the result hypothesized by the devolper and suggests some comparability in the quality constructs of the measure indicators.
- The developer previously did face validity by three expert panels in the 2016 submission of the measure.

20. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

- The results of the Pearson correlation test (on HEDIS data of commercial, Medicare and Medicaid health plans) suggests that performance is correlated. Coefficients were an absolute value of greater then 0.3 for most comparisons, which denotes per developer moderate associations.

- 2016 submission face validity results indicate the technical expert panel showed good agreement that the measure as specified accurately.

21. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

    Submission document: Testing attachment, section 2b1.

    ☒ **Yes**

    ☐ **No**

    ☐ **Not applicable** (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?**

    NOTE that data element validation from the literature is acceptable.

    Submission document: *Testing attachment, section 2b1.*

    ☐ **Yes**

    ☐ **No**

    ☒ **Not applicable** (data element testing was not performed)

23. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

    ☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

    ☐ **Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

    Validity methodology and results are appropriate and yielded good validity scores.

ADDITIONAL RECOMMENDATIONS

25. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

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.

Data Specifications and Elements

- All data elements are in defined fields in electronic claims.
- This is not an eMeasure

Data Collection Strategy

- NCQA conducts an independent audit of all HEDIS collection and reporting processes, as well as an audit of the data which are manipulated by those processes, in order to verify that HEDIS specifications are met.
- The developer also has a Policy Clarification Support System for any inquiries on the measure.
- The measure for broad public use is encouraged by developer. Written consent would be required for any "commercial use".

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?
- Are the required data elements available in electronic form, e.g., EHR or other electronic sources?
- Is the data collection strategy ready to be put into operational use?

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

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility

Comments:

** All data is collected through electronic claims submission by providers and pharmacies to health plans.

** Data elements are all in define fields routinely generated and in electronic form and has already been used commercially as it is aligned with HEDIS>

** All elements are routinely used and I have no concerns about operational use of the data collection strategy.

** Threat to feasibility may relate to whether prescribing of corticosteroids and/or bronchodilators may not be completely captured.

** All data elements are in defined fields in electronic claims.

** Very feasibile to generate and collect the data.

** I see nothing that would cause me to think there is anything about this measure that is not completely feasible.

** I don't have concerns about feasibility

Criterion 4: Usability and Use

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

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. 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.

4a.1. 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.

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

- NCQA HEALTH PLAN RATINGS/REPORT CARDS: This measure is used in the calculation of health plan ratings, which are reported on the NCQA website annually. These ratings are based on performance on HEDIS measures among other factors.In 2019, a total of 255 Medicare health plans, 515 commercial health plans and 188 Medicaid health plans across 50 states were included in the rankings.
- NCQA HEALTH PLAN ACCREDITATION: This measure is used in scoring for accreditation of Medicare Advantage Health Plans. As of Fall 2017, a total of 184 Medicare Advantage health plans were scored for accreditation using this measure among others covering 9.2 million Medicare beneficiaries; 451 commercial health plans covering 113 million lives; and 125 Medicaid health plans covering 35 million lives. Health plans are scored based on performance compared to national benchmarks.
- NCQA QUALITY COMPASS: This measure is used in Quality Compass which is an indispensable tool used for selecting health plans, conducting competitor analysis, examining quality improvement and benchmarking plan performance. Provided in this tool is the ability to generate custom reports by selecting plans, measures, and benchmarks (averages and percentiles) for up to three trended years. Results in table and graph formats offer simple comparison of plans' performance against competitors or benchmarks.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- NCQA measures are evaluated regularly using a consensus-based process to consider input from multiple stakeholders, including but not limited to entities being measured. The developer used several methods to obtain input, including vetting of the measure with several multi-stakeholder advisory panels, 30-day public comment posting, and review of questions submitted to the Policy Clarification Support System. This information enables NCQA to assess a measure's adherence to the HEDIS Desirable Attributes of Relevance, Scientific Soundness and Feasibility.
- Feedback informed how the developer revised the measure to clarify how to identify which inpatient and ED visits should be included in the denominator.

Additional Feedback:

- The developer/steward did not provide any further feedback.

***Questions for the Committee**:*

- How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?
- How has the measure been vetted in real-world settings by those being measured or others?

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

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability 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.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- The performance rates for both numerators were higher in 2017 than in 2016 across all product lines. From 2017 to 2018, some of the same rates showed modest increases, while others showed modest decreases, and others did not change, making it difficult to identify a trend for the last 3 years.

4b2. Benefits vs. harms. 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).

Unexpected findings (positive or negative) during implementation [unexpected findings]

Potential harms

- None identified

Additional Feedback:

- None

Questions for the Committee:

- Do the performance results suggest continued opportunity for improvement?
- Do the benefits of the measure outweigh any potential unintended consequences?

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

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1. Use - Accountability and Transparency

Comments:

** The measure is publicly reported NCQA. Performance data are available to outside organizations typically aggregated at the plan level. The measures developer provides opportunity for health plans to provide feedback during the development process.

** Used publically in Health Plan Rating/Report cards, Accreditation and with Quality Compass. Regular consensus-based feedback process with multiple stakeholders.

** The measure is being utilized in several reporting programs via NCQA and Medicare; performance results are made public within scheduled timeframes. Measures are consistently evaluated via consensus-based processes which allow for feedback on components and implementation.

** Measure likely to continue to be publicly reported as we move forward with increasing quality transparency.

** Performance results are used by the participating health plans. NCQA uses the results in rating health plans and uses collective performance results to analyze delivery of health care in publications and conferences. NCQA has channels for measure feedback from stakeholders, advisory panels, and the public.

** Many opportunities for public communications.

** This measure is publically reported and is used in various public reporting and accountability programs listed by the developer.

** Feedback informed how the developer revised the measure to clarify how to identify which inpatient and ED visits should be included in the denominator

4b1. Usability – Improvement

Comments:

** The measure developer found that over the past three years "it is difficult to identify a [performance rate improvement] trend." They report that some "rates showed modest increases, while others showed modest decreases." The measure does not appear to have had a significant impact on performance among health plans or driven improvements in quality of care. It does not appear that there is harm or unintended consequences of this measure.

** Move from a modest increases in change to a higher rate of change that shows a continued opportunity for improvement. There were no potential harms identified.

** Plans can look at the populations that are consistently not being dispensed systemic corticosteroids or bronchodilators, evaluate barriers to receiving the drugs, and test interventions that will push appropriate therapies to patients.

** No issues with usability.

** There have been no steady trends from the data over the three years of analyzed HEDIS data. Unintended benefits and harms have not been detected. There are significant side effects from both categories of medications but their use in addressing COPD exacerbations consistent with clinical guidelines have benefits outweighing harms.

** Used in variety of tools by NCQA for broad communication to the public.

** Develop a ,eams for providing a higher degree of compliance with practitioners.

** Per developer, there were no identified unexpected findings (positive or negative) during testing or since implementation of this measure.

Criterion 5: Related and Competing Measures

Related or competing measures

The measure developer identified thre related or competing measures:

- 0102 : COPD: inhaled bronchodilator therapy – Percentage of patients aged 18 years or older, with a diagnosis of COPD (FEV1/FVC < 70%) who have an FEV1 < 60% predicted and have symptoms who were prescribed an inhaled bronchodilator
- 0577 : Use of Spirometry Testing in the Assessment and Diagnosis of COPD – This measure assesses the percentage of patients 40 years of age and older with a new diagnosis of COPD or newly active COPD, who received appropriate spirometry testing to confirm the diagnosis.
- 1825 : COPD - Management of Poorly Controlled COPD – The percentage of patients age 18 years or older with poorly controlled COPD, who are taking a long acting bronchodilator.

Harmonization

Developer indicates that the measures are harmonized to the extent possible. NQF staff consider these measures to be related but not competing. Measure 1825 has been withdrawn from NQF endorsement.

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

5. Related and Competing

Comments:

** The related measures do not appear to directly compete with 2856, however, the committee should discuss the potential for harmonization.

** Measures 0102 and 0577 are harmonized. Measure 1825 has been withdrawn from NQF endorsement.

** There are related measures that examine COPD management, COPD assessment and diagnosis, and inhaled bronchodilator therapy. The measures are harmonized.

** 0102 COPD: Inhaled bronchodilator therapy -- similar measure, but different age group and this process measure is focused solely on prescriptions for inhaled long acting bronchodilator.

** Three other NQF accepted measures for the process of COPD treatment are different in concept and non-competing.

** Although related, none are competing.

** There are related measures which the developer feels are harmonized to the extent possible

** Yes, related measures. I think they are harmonized to the extent possible

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 1/31/2020

- **No NQF Members have submitted support/non-support choices as of this date.**

1. Evidence and Performance Gap – 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

PCE_Evidence_Form.docx

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

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

Yes

1a. Evidence (subcriterion 1a)

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 2856

Measure Title: Pharmacotherapy Management of COPD Exacerbation

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here:

Date of Submission: 11/8/2019

Instructions

- ***Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.***
- ***Complete EITHER 1a.2, 1a.3 or 1a.4 as applicable for the type of measure and evidence.***
- ***For composite performance measures:***
  - ***A separate evidence form is required for each component measure unless several components were studied together.***
  - ***If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.***
- **All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *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.**
- **Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.**

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

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.**
- **Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.**
- **Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.**
- **Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.**
- **Efficiency: [6] evidence not required for the resource use component.**
- **For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.**
- **Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.**

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (*A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)*

☐ Intermediate clinical outcome (*e.g., lab value*):

☒ Process: The percentage of COPD exacerbations for patients 40 and older with an acute inpatient discharge or ED visit, who were dispensed appropriate medications. Rate 1: systemic corticosteroid within 14 days. Rate 2: bronchodilator within 30 days.

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

Patient has inpatient or emergency department (ED) visit for a COPD exacerbation >>> Patient receives systemic corticosteroid and bronchodilator to treat symptoms >>> Patient has improved outcomes including improved lung function and arterial hypoxemia, reduced risk of early relapse or treatment failure, and shorter recovery time and length of hospital stay

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

This measure assesses whether patients who were hospitalized or had an emergency department (ED) visit for a COPD exacerbation were dispensed appropriate medication to treat their symptoms. This measure is based on evidence that systemic corticosteroids and bronchodilators following a COPD exacerbation have been shown to improve symptoms, reduce the risk of early relapse, and shorten recovery time and length of hospital stay.

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☒ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	**2015 Submission** **Global Initiative for Chronic Obstructive Lung Disease (GOLD) Guidelines:** **Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Bethesda (MD): Global Initiative for Chronic Obstructive Lung Disease, 2015. http://www.goldcopd.org/guidelines-global-strategy-for-diagnosis-management.html** **2019 Submission** **Global Initiative for Chronic Obstructive Lung Disease (GOLD) Guidelines:** **Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Global Initiative for Chronic Obstructive Lung Disease, 2020. https://goldcopd.org/gold-reports/** **2015 Submission** **Institute for Clinical Systems Improvement (ICSI) Guidelines:** **Anderson, B., K. Conner, C. Dunn, et al. 2013. Institute for Clinical Systems Improvement. Diagnosis and Management of Chronic Obstructive Pulmonary Disease (COPD). https://www.icsi.org/guidelines__more/catalog_guidelines_and_more/catalog_guidelines/catalog_respiratory_guidelines/copd/** **The ICSI Guidelines have been retired. ICSI has endorsed the Veteran's Affairs/Department of Defense (VA/DoD) Clinical Practice Guideline for the Management of Chronic Obstructive Pulmonary Disease.** **2019 Submission** **Veteran's Affairs/Department of Defense Clinical Practice Guideline for the Management of Chronic Obstructive Pulmonary Defense:** **The Management of Chronic Obstructive Pulmonary Disease Working Group, The Office of Quality, Safety and Value, VA, Washington, DC, & Office of Evidence Based Practice, US Army Medical Command. (2014). Veterans Affairs/Department of Defense Practice Guideline for the Management of Chronic Obstructive Pulmonary Disease. https://www.healthquality.va.gov/guidelines/CD/copd/VADoDCOPDCPG2014.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	2015 Submission **GOLD GUIDELINES, 2015, MANAGEMENT OF EXACERBATIONS, page 42** **Pharmacologic Treatment** • *Short-acting Bronchodilators:* Although there are no controlled trials, short-acting inhaled beta$_2$-agonists with or without short-acting anticholinergics are usually the preferred bronchodilators for treatment of an exacerbation (Evidence C). • *Corticosteroids:* Data from studies in secondary health care indicate that systemic corticosteroids in COPD exacerbations shorten recovery time, improve lung function (FEV_1) and arterial hypoxemia (PaO_2) (Evidence A), and reduce the risk of

early relapse, treatment failure, and length of hospital stay. A dose of 40 mg prednisone per day for 5 days is recommended (Evidence B).

2019 Submission

GOLD Guidelines, 2020 (page 106)

"Short-acting inhaled beta$_2$-agonists, with or without short-acting anticholinergics, are recommended as the initial bronchodilators to treat an acute exacerbation (Evidence C)."

"Systemic corticosteroids can improve lung function (FEV_1), oxygenation and shorten recovery time and hospitalization duration. Duration of therapy should not be more than 5-7 days (Evidence A)."

2015 Submission

ICSI Guidelines, page 26-27:

- **Bronchodilators:** Albuterol and levalbuterol are the preferred bronchodilators in the setting of an acute exacerbation of COPD because of their rapid onset of action. If clinical improvement does not occur promptly, ipratropium may be added to produce additive bronchodilation and allow the use of lower doses of albuterol or levabuterol, thus diminishing dose-dependent toxicity. Administration of either agent by metered-dose inhaler and spacer or by nebulization is acceptable *(Turner, 1997 [Meta-analysis]; Moayyedi, 1995 [High Quality Evidence]; Patrick, 1990 [High Quality Evidence])*.
- **Systemic Steroids:** Studies have demonstrated the benefits of systemic glucocorticosteroids in the management of COPD exacerbations. Doses of oral prednisone at 30 to 40 mg a day for 7 to 14 days have been shown to reduce symptoms and reduce the likelihood of hospitalization. Treatment beyond two weeks does not provide any additional benefit, but does increase the likelihood of significant side effects such as hyperglycemia. There is no need to discontinue inhaled steroids while the patient is taking oral prednisone (*Aaron, 2003 [High Quality Evidence]; McEvoy, 2000 [Low Quality Evidence]; Davies, 1999 [High Quality Evidence]; Niewoehner, 1999 [High Quality Evidence]; Thompson, 1996 [High Quality Evidence])*.

2019 Submission

VA/DoD Guidelines (Pages 17, 20)

- "We recommend prescribing inhaled short-acting beta 2-agonists (SABAs) to patients with confirmed COPD for rescue therapy as needed. (Strong For)."
- "We recommend inhaled tiotropium as first-line therapy for patients with confirmed, stable COPD who have respiratory symptoms (e.g., dyspnea, cough) and severe airflow obstruction (I.e., post bronchodilator FEV1 < 50%) or a history of COPD exacerbations. (Strong For)."
- "For acute COPD exacerbations, we recommend a course of systemic corticosteroids (oral preferred) of 30-40mg prednisone equivalent daily for 5-7 days (Strong For)."

Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	2015 Submission **Global Initiative for Chronic Obstructive Lung Disease (GOLD) Levels of Evidence:** Evidence A. Randomized controlled trials. Rich body of data. Evidence is from endpoints of well-designed randomized controlled trials that provide a consistent pattern of findings in the population for which the recommendation is made. Category A requires substantial numbers of studies involving substantial numbers of participants. Evidence B. Randomized controlled trials. Limited data. Evidence is from endpoints of intervention studies that include only a limited number of patients, post hoc or subgroup analysis of randomized controlled trials, or meta-analysis of randomized controlled trials. In general, Category B pertains when few randomized trials exist, they are small in size, they were undertaken in a population that differs from the target population of the recommendation, or the results are somewhat inconsistent. Evidence C. Nonrandomized trials. Observational studies. Evidence is from outcomes of uncontrolled or nonrandomized trials or from observational studies. 2019 Submission **GOLD Guidelines, 2020** Evidence Category A: Randomized controlled trials (RCTs). Rich body of high quality evidence without any significant limitation or bias. Evidence is from endpoints of well-designed RCTs that provide consistent findings in the population for which the recommendation is made without any important limitations. Requires high quality evidence from ≥ 2 clinical trials involving a substantial number of subjects, or a single high quality RCT involving substantial numbers of patients without any bias. Evidence Category C: Non-randomized trials. Observational studies. Evidence is from outcomes of uncontrolled or non-randomized trials or from observational studies. 2015 Submission **ICSI Guidelines:** • High Quality Evidence = Further research is very unlikely to change our confidence in the estimate of effect. • Low Quality Evidence = Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate or any estimate of effect is very uncertain. 2019 Submission **VA/DoD Guidelines** Strong For indicates that the Work Group is highly confident that desirable outcomes outweigh undesirable outcomes.
Provide all other grades and definitions from the evidence grading system	2015 Submission **Global Initiative for Chronic Obstructive Lung Disease (GOLD) Levels of Evidence:** Evidence D. Panel consensus. Judgment. This category is used only in cases where the provision of some guidance was deemed valuable but the clinical literature

	addressing the subject was deemed insufficient to justify placement in one of the other categories. The Panel Consensus is based on clinical experience or knowledge that does not meet the above listed criteria. 2019 Submission **GOLD Guidelines, 2020** Evidence Category B. Randomized controlled trials (RCTs) with important limitations. Limited Body of Evidence. Evidence is from RCTs that include only a limited number of patients, post hoc or subgroup analyses of RCTs or meta analyses of RCTs. Also pertains when few RCTs exist, or important limitations are evident (methodologic flaws, small numbers, short duration, undertaken in a population that differs from the target population of the recommendation, or the results are somewhat inconsistent). Evidence Category D. Panel consensus judgment. Provision of guidance is deemed valuable but clinical literature addressing the subject is insufficient. Panel consensus is based on clinical experience or knowledge that does not meet the above stated criteria. 2015 Submission **ICSI Guidelines:** • Moderate Quality Evidence = Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. 2019 Submission **VA/DoD Guidelines** • Weak-For Recommendation: Indicates that the Work Group is less confident of the balance between desirable and undesirable outcomes • Strong-Against Recommendation: Indicates that the Work Group is confident that the undesirable consequences outweigh the desirable consequences • Weak-Against Recommendation: Indicates that the Work Group is less confident that the undesirable consequences outweigh the desirable consequences
Grade assigned to the **recommendation** with definition of the grade	2015 Submission **Global Initiative for Chronic Obstructive Lung Disease (GOLD) Levels of Evidence:** Evidence A. Randomized controlled trials. Rich body of data. Evidence is from endpoints of well-designed randomized controlled trials that provide a consistent pattern of findings in the population for which the recommendation is made. Category A requires substantial numbers of studies involving substantial numbers of participants. Evidence B. Randomized controlled trials. Limited data. Evidence is from endpoints of intervention studies that include only a limited number of patients, post hoc or subgroup analysis of randomized controlled trials, or meta-analysis of randomized controlled trials. In general, Category B pertains when few randomized trials exist, they are small in size, they were undertaken in a population that differs from the target population of the recommendation, or the results are somewhat inconsistent. Evidence C. Nonrandomized trials. Observational studies. Evidence is from outcomes of uncontrolled or nonrandomized trials or from observational studies.

	2019 Submission **GOLD Guidelines, 2020** The GOLD report does not grade evidence and recommendations separately. 2015 Submission **ICSI Guidelines:** • High Quality Evidence = Further research is very unlikely to change our confidence in the estimate of effect. • Low Quality Evidence = Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate or any estimate of effect is very uncertain. 2019 Submission **VA/DoD Guidelines** The VA/DoD Guidelines do not grade evidence and recommendations separately.
Provide all other grades and definitions from the recommendation grading system	2015 Submission **Global Initiative for Chronic Obstructive Lung Disease (GOLD) Levels of Evidence:** Evidence D. Panel consensus. Judgment. This category is used only in cases where the provision of some guidance was deemed valuable but the clinical literature addressing the subject was deemed insufficient to justify placement in one of the other categories. The Panel Consensus is based on clinical experience or knowledge that does not meet the above listed criteria. 2019 Submission **GOLD Guidelines, 2020** The GOLD report does not grade evidence and recommendations separately. 2015 Submission **ICSI Guidelines:** • Moderate Quality Evidence = Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. 2019 Submission **VA/DoD Guidelines** The VA/DoD Guidelines do not grade evidence and recommendations separately.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	2015 Submission **GOLD Guidelines, 2015: 1965-2014** The guideline developers did not provide a breakdown of the specific number of randomized control trials (RCTs) and given the number of studies included in the systematic reviews we were not able to delineate all RCTs for each recommendation. The GOLD guidelines referenced a total of 613 studies to update the previous set of guidelines from 2013. The recommendation related to systemic corticosteroid and short-acting

	bronchodilators to treat a COPD exacerbation was based on 25 studies, including randomized control trials (RCTs), meta-analyses, systematic reviews of randomized controlled trials (RCTs), and observational studies. 2015 Submission **ICSI Guidelines:** The ICSI guideline recommendations related to short-acting bronchodilators following an exacerbation referenced one meta-analysis of 12 RCTs. The ICSI guideline recommendations related to systemic corticosteroids following an exacerbation referenced four randomized control trials and one observational study. 2015 Submission (across studies) Overall, the quality of the evidence that systemic corticosteroids improve outcomes following a COPD exacerbation is high. The evidence is based on the results of multiple randomized control trials conducted in hundreds of patients with COPD. These studies have consistently found that patients who received systemic corticosteroids within two weeks following an exacerbation had improved lung function, fewer treatment failures, increased time to relapse, and shorter hospital stays than patients who did not receive systemic corticosteroids. The quality of evidence that short-acting bronchodilators improve outcomes following COPD exacerbation is moderate as the RCTs were conducted in fewer patients; however, the findings from the meta-analysis of 12 RCTs found that short-acting bronchodilators caused statistically significant increases in pulmonary function tests and are important for rapidly improving shortness of breath and wheezing. 2019 Submission **GOLD Guidelines, 2020** The body of evidence for the recommendation for bronchodilators is based primarily on 2 publications: the National Institute for Health and Care Excellence (NICE) guidelines for COPD diagnosis and management for persons over 16 years, and a summary of the American Thoracic Society/European Respiratory Society (ATS/ERS) position paper on standards for the diagnosis and treatment of patients with COPD. The NICE guidelines and the ATS/ERS position paper synthesize findings from RCTs and other studies. National Institute for Health and Care Excellence. Chronic obstructive pulmonary disease in over 16s: diagnosis and management. 2018. https://www.nice.org.uk/guidance/NG115 Celli BR, MacNee W, ATS ERS Task Force. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004; 23(6): 932-46. The body of evidence for the recommendation for systemic corticosteroids is based on eight studies, four of which are RCTs. The literature included in this 2020 edition of the GOLD Report has been updated from the previous 2019 edition to include “key peer-reviewed research publications” published from January 2018 to July 2019.

	2019 Submission **VA/DoD Guidelines** Evidence for the recommendation for SABAs is based on three studies, including a systematic review with 13 trials. Evidence for the recommendation for tiotropium is based on a Cochrane Database systematic review. Evidence for the recommendation for systemic corticosteroids is based on four studies, including a large RCT and a large systematic review. Evidence questions guided a systematic evidence review, which identified the body of evidence relevant to each evidence question. The overall quality of the body of evidence was assessed using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) methodology, which takes multiple factors (overall study quality, consistency of evidence, directness of evidence, and precision of evidence) into consideration to rate the overall quality of the evidence (i.e. high, moderate, low, very low).
Estimates of benefit and consistency across studies	2015 Submission (across studies) Studies have consistently found that the primary benefit of short-acting bronchodilators for COPD patients is increasing FEV_1 and thus are recommended as as-needed therapy including following a COPD exacerbation. Randomized control trials consistently found that patients who received systemic corticosteroids within two weeks after experiencing a COPD exacerbation had statistically significant better outcomes compared to patients who did not. The following are examples of statistically significant study findings showing the magnitude of the outcomes: • The percentage-predicted FEV1 after bronchodilation rose more frequently in corticosteroid-treated patients (from 28% to 42%) compared to - patients not treated with corticosteroid (from 26% to 32%). • Systemic glucocorticoids were associated with a shorter initial hospital stay (8.5 days, vs. 9.7 days for placebo). • The overall rate of relapse at 30 days was lower in the corticosteroid group than in the placebo group (27 percent vs. 43 percent). 2019 Submission (across studies) Studies have consistently shown the benefit of using bronchodilators in the event of a COPD exacerbation. Systemic corticosteroids have been consistently shown to improve outcomes in COPD exacerbations, such as shortening recovery time, and improving lung function.
What harms were identified?	2015 Submission (across studies) One risk associated with the use of bronchodilators is toxicity, however the benefit of use during an acute episode outweighs the risk. The risk of toxicity is mainly attributed in long term medication use in stable COPD. Patients who receive consistent systematic corticosteroid therapy are at higher risk for hyperglycemia or osteoporosis. However, there is consensus that short-term use of systemic corticosteroids to treat a COPD exacerbation outweigh the potential harms. 2019 (across studies)

	No new significant harms with either bronchodilators or systemic corticosteroids were identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new studies that change the recommendations stated above have been identified.

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

1a.4.2 What process was used to identify the evidence?

1a.4.3. Provide the citation(s) for the evidence.

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)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

This measure assesses whether patients who had a hospitalization or an emergency department (ED) visit for a COPD exacerbation were provided appropriate medication (systemic corticosteroids and bronchodilators) to treat symptoms and prevent future exacerbations. The improvement in quality envisioned by the use of this measure is to increase the use of systemic corticosteroids and bronchodilators following a COPD exacerbation in order to improve patient outcomes, such as shorten recovery time, improve lung function and reduce the risk of early relapse, treatment failure, and length of hospital stay.

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 (4b1) under Usability and Use.*

The following data are extracted from HEDIS data collection reflecting the most recent years of measurement for this measure. Performance data is summarized at the health plan level and summarized by mean, standard

deviation, minimum health plan performance, maximum health plan performance and performance at 10th, 25th, 50th, 75th, and 90th percentile. Data is stratified by year and product line (i.e. commercial, Medicare, Medicaid).

Systemic Corticosteroids – Commercial Rate

YEAR| MEAN | ST DEV | MIN | 10TH | 25TH | 50TH | 75TH | 90TH | MAX | Interquartile Range

2016 | 69.7% | 9.5% | 33.3% | 57.5% | 64.5% | 70.5% | 76.7% | 81.4% | 88.2% | 12.2%

2017 | 75.5% | 7.6% | 40.7% | 66.7% | 71.7% | 75.7% | 80.0% | 84.0% | 100% | 8.3%

2018 | 74.2% | 8.0% | 46.9% | 64.2% | 70.2% | 75.7% | 79.7% | 84.9% | 91.5% | 9.5%

Systemic Corticosteroids – Medicaid Rate

YEAR| MEAN | ST DEV | MIN | 10TH | 25TH | 50TH | 75TH | 90TH | MAX | Interquartile Range

2016 | 66.0% | 11.6% | 27.1% | 49.8% | 61.9% | 67.9% | 73.1% | 77.7% | 88.8% | 11.2%

2017 | 68.3% | 11.3% | 24.5% | 53.6% | 63.0% | 70.2% | 76.3% | 80.4% | 92.5% | 13.3%

2018 | 68.4% | 11.7% | 34.6% | 50.3% | 63.9% | 71.1% | 75.5% | 81.3% | 91.2% | 11.6%

Systemic Corticosteroids – Medicare Rate

YEAR| MEAN | ST DEV | MIN | 10TH | 25TH | 50TH | 75TH | 90TH | MAX | Interquartile Range

2016 | 67.7% | 10.6% | 13.3% | 54.9% | 62.4% | 69.7% | 74.7% | 78.9% | 87.8% | 12.3%

2017 | 70.4% | 7.8% | 27.8% | 62.5% | 67.5% | 71.6% | 74.7% | 78.4% | 87.5% | 7.2%

2018 | 71.4% | 9.1% | 18.5% | 60.6% | 68.0% | 73.0% | 76.9% | 80.5% | 90.9% | 8.9%

Bronchodilators – Commercial Rate

YEAR| MEAN | ST DEV | MIN | 10TH | 25TH | 50TH | 75TH | 90TH | MAX | Interquartile Range

2016 | 75.8% | 8.7% | 36.0% | 65.8% | 70.3% | 75.9% | 81.8% | 86.4% | 95.8% | 11.5%

2017 | 80.1% | 7.0% | 45.2% | 71.9% | 76.3% | 80.4% | 84.3% | 88.3% | 96.8% | 8.0%

2018 | 79.6% | 7.4% | 50.0% | 70.8% | 75.3% | 79.2% | 84.7% | 89.0% | 100% | 9.4%

Bronchodilators – Medicaid Rate

YEAR| MEAN | ST DEV | MIN | 10TH | 25TH | 50TH | 75TH | 90TH | MAX | Interquartile Range

2016 | 80.7% | 9.7% | 39.5% | 71.8% | 78.3% | 83.5% | 86.4% | 88.4% | 96.0% | 8.1%

2017 | 81.4% | 10.1% | 32.7% | 70.6% | 78.8% | 83.8% | 87.6% | 89.7% | 96.6% | 8.8%

2018 | 81.4% | 10.8% | 40.5% | 68.8% | 79.4% | 84.7% | 87.9% | 89.8% | 95.0% | 8.5%

Bronchodilators – Medicare Rate

YEAR| MEAN | ST DEV | MIN | 10TH | 25TH | 50TH | 75TH | 90TH | MAX | Interquartile Range

2016 | 77.9% | 9.1% | 37.7% | 67.8% | 72.8% | 78.6% | 84.5% | 88.3% | 97.2% | 11.7%

2017 | 79.6% | 8.2% | 33.3% | 71.2% | 75.7% | 79.9% | 84.9% | 89.5% | 97.2% | 9.2%

2018 | 79.8% | 8.5% | 42.2% | 70.4% | 76.1% | 80.4% | 85.4% | 89.5% | 98.1% | 9.3%

The data references are extracted from HEDIS data collection reflecting the most recent years of measurement for this measure. In 2018, HEDIS measures covered more than 190 million people. Below is a description of the denominator for this measure. It includes the number of health plans included in HEDIS data collection and the mean eligible population for the measure across health plans.

Commercial

YEAR | N Plans | Mean Denominator Size per plan

2016 | 263 | 179

2017 | 254 | 178

2018 | 255 | 165

Medicaid

YEAR | N Plans | Mean Denominator Size per plan

2016 | 200 | 863

2017 | 201 | 885

2018 | 189 | 937

Medicare

YEAR | N Plans | Mean Denominator Size per plan

2016 | 385 | 681

2017 | 390 | 750

2018 | 388 | 707

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 (4b1) under Usability and Use.*

The CMS Office of Minority Health in collaboration with the RAND Corporation produces an annual report: CMS Racial, Ethnic, and Gender Disparities in Health Care in Medicare Advantage. We provide below summary data for this measure from that report. The authors note that "for reporting HEDIS data stratified by race and ethnicity, racial and ethnic group membership is estimated using a methodology that combines information from CMS administrative data, surname, and residential location."

In the 2019 CMS report, Asian/Pacific Islander patients with a COPD exacerbation were significantly more likely than White patients to receive a systemic corticosteroid with 14 days, while Hispanic patients were significantly less likely than White patients to receive the same. There was no significant difference between White and Black patients. Regarding the the receipt of a bronchodilator with 30 days of a COPD exacerbation, Asian/Pacific Islander patients and Black patients were more likely to receive than White patients, while Hispanic patients were less likely.

2019 CMS Racial, Ethnic, and Gender Disparities in Health Care in Medicare Advantage report. https://www.cms.gov/About-CMS/Agency-Information/OMH/Downloads/2019-National-Level-Results-by-Race-Ethnicity-and-Gender.pdf

HEDIS data are stratified by type of insurance (e.g. commercial, Medicaid, Medicare). NCQA does not currently collect performance data stratified by race, ethnicity, or language. Escarce et al. have described in detail the difficulty of collecting valid data on race, ethnicity, and language at the health plan level (Escarce, 2011). While not specified in the measure, this measure can also be stratified by demographic variables, such as race/ethnicity or socioeconomic status, in order to assess the presence of health care disparities. The HEDIS Health Plan Measure Set contains two measures that can assist with stratification to assess health care disparities. The Race/Ethnicity Diversity of Membership and the Language Diversity of Membership measures were designed to promote standardized methods for collecting these data and follow Office of Management

and Budget and Institute of Medicine guidelines for collecting and categorizing race/ethnicity and language data. In addition, NCQA's Multicultural Health Care Distinction Program outlines standards for collecting, storing and using race/ethnicity and language data to assess health care disparities.

Escarce, J.J., Carreon, R., Veselovskiy, G., Lawson, E.G. Collection of Race and Ethnicity Data by Health Plans has Grown Substantially, but Opportunities Remain to Expand Efforts. Health Affairs (Millwood) 2011; 30(10):1984-91. http://www.ncbi.nlm.nih.gov/pubmed/21976343

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

Disparities are present in the quality of care of COPD, as well as in disease burden. In a study that examined the quality of care for patients who were hospitalized for acute exacerbations of COPD, several patient characteristics were shown to be associated with a decreased likelihood of receiving guidelines-based care, which included the receipt of inhaled anticholinergic bronchodilators, inhaled short-acting 2-agonists, and systemic corticosteroids. Patients who were 75 years and older, or African American, were less likely to receive comprehensive care, compared to younger patients, or white patients, respectively (Lindanauer, et al, 2006). African Americans have been shown to be more likely than whites to report having a COPD exacerbation that required hospitalization in the previous year (Han, et al, 2011). Lower education and lower income, often used as proxies for lower socioeconomic status, have been shown to be correlated with a greater risk of COPD exacerbations (Eisner, et al 2009).

References

Eisner, M. D., Blanc, P. D., Omachi, T. A., Yelin, E. H., Sidney, S., Katz, P. P., … Iribarren, C. (2009). Socioeconomic status, race and COPD health outcomes. Journal of Epidemiology & Community Health, 65(1), 26-34. doi:10.1136/jech.2009.089722

Han, M. K., Curran-Everett, D., Dransfield, M. T., Criner, G. J., Zhang, L., Murphy, J. R., … Foreman, M. G. (2011). Racial Differences in Quality of Life in Patients With COPD. Chest, 140(5), 1169-1176. doi:10.1378/chest.10-2869

Lindenauer, P. K., Pekow, P., Gao, S., Crawford, A. S., Gutierrez, B., & Benjamin, E. (2006). Quality of Care for Patients Hospitalized for acute exacerbations of chronic obstructive pulmonary disease. Annals of Internal Medicine, 144(12), 894-903.

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):*

Respiratory, Respiratory : Chronic Obstructive Pulmonary Disease (COPD)

De.6. Non-Condition Specific*(check all the areas that apply):*

Primary Prevention

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

Elderly, 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.)*

N/A

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 not an eMeasure **Attachment:**

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:** 2856_PCE_Value_Sets_Fall_2019.xlsx

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

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.

N/A

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).

Numerator #1 (Systemic corticosteroids): The number of patients dispensed a prescription for a systemic corticosteroid on or 14 days after the Episode Date. Count systemic corticosteroids that are active on the relevant date.

Numerator #2 (Bronchodilators): The number of patients dispensed a prescription for a bronchodilator on or 30 days after the Episode Date. Count bronchodilators that are active on the relevant date.

*The Episode Date is the date of service for any acute inpatient discharge or ED claim/encounter during the 11-month intake period with a principal diagnosis of COPD.

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).

Numerator 1 (Systemic Corticosteroid): Identify the number of patients dispensed a prescription for a systemic corticosteroid on or 14 days after the Episode Date.

-The Episode Date is the date of service for any acute inpatient discharge or ED visit during the 11-month intake period with a principal diagnosis of COPD.

-Count systemic corticosteroids that are active on the relevant date. A prescription is considered active if the "days supply" indicated on the date the patient filled the prescription is the number of days or more between that date and the relevant date. For an acute inpatient encounter, the relevant date is the date of admission. For an ED visit, the relevant date is the date of service.

Systemic Corticosteroid Medications List:

Glucocorticoids: cortisone-acetate, dexamethasone, hydrocortisone, methylprednisolone, prednisolone, and prednisone. See attached Value Set Excel document.

Numerator 2 (Bronchodilator): Identify the number of patients dispensed a prescription for a bronchodilator on or 30 days after the Episode Date.

-The Episode Date is the date of service for any acute inpatient discharge or ED visit during the 11-month intake period with a principal diagnosis of COPD.

-Count bronchodilators that are active on the relevant date. A prescription is considered active if the "days supply" indicated on the date the patient filled the prescription is the number of days or more between that date and the relevant date. For an acute inpatient encounter, the relevant date is the date of admission. For an ED visit, the relevant date is the date of service.

Bronchodilator Medications List:

-Anticholinergic agents: albuterol-ipratropium, aclidinium-bromide, ipratropium, tiotropium, umeclidinium

-Beta 2-agonists: albuterol, arformoterol, budesonide-formoterol, fluticasone-salmeterol, fluticasone-vilanterol, formoterol, formoterol-glycopyrrolate, indacaterol, indacaterol-glycopyrrolate, levalbuterol, formoterol-mometasone, metaproterenol, olodaterol hydrochloride, olodaterol-tiotropium, salmeterol, umeclidinium-vilanterol

-Anti-asthmatic combinations: dyphylline-guaifenesin

See attached Value Set Excel document.

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

All patients age 40 years or older as of January 1 of the measurement year with a COPD exacerbation as indicated by an acute inpatient discharge or ED encounter with a principal diagnosis of COPD.

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 eligible population for this measure is based on acute inpatient discharges and ED visits, not on patients. It is possible for the denominator to include multiple events for the same individual. The eligible population for the denominator is defined by following the series of steps below:

Step 1: Identify all patients who had either of the following during the Intake Period (an 11-month period that begins on January 1 of the measurement year and ends on November 30 of the measurement year):

1) An ED visit (ED Value Set) with a principal diagnosis of COPD (COPD Value Set), emphysema (Emphysema Value Set) or chronic bronchitis (Chronic Bronchitis Value Set). Do not include ED visits that result in an inpatient stay.

2) An acute inpatient discharge with a principal diagnosis of COPD (COPD Value Set), emphysema (Emphysema Value Set) or chronic bronchitis (Chronic Bronchitis Value Set) on the discharge claim. To identify acute inpatient discharges:

a. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set)

b. Exclude nonacute inpatient stays (Nonacute Inpatient Stay Value Set)

c. Identify the discharge date for the stay

Step 2: Identify all COPD Episodes. For each patient identified in Step 1, identify all acute inpatient discharges and ED Visits. An acute inpatient discharge and ED visit on the same date are counted as one COPD episode (ED visits that result in an inpatient stay are excluded in Step 1). Multiple ED visits on the same date are counted as one COPD episode.

Step 3: Test for direct transfers. For episodes with a direct transfer to an acute or nonacute setting for any diagnosis, the Episode Date is the discharge data from the last admission.

A direct transfer is when the discharge date from the first inpatient setting precedes the admission date to a second inpatient setting by one calendar day or less.

Use the following method to identify admission to and discharges from inpatient settings.

1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set).

2. Identify the admission and discharge dates for the stay.

See corresponding Excel file for value sets referenced above.

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

This measure excludes patients who use hospice services, and patients with nonacute inpatient stays.

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.)*

Exclude patients who use hospice services or elect to use a hospice benefit any time during the measurement year, regardless of when the services began. These patients may be identified using various methods, which may include but are not limited to enrollment data, medical record, claims/encounter data (Hospice Encounter Value Set, Hospice Intervention Value Set).

Exclude patients with nonacute inpatient stays (Nonacute Inpatient Stay Value Set).

See attached Hospice Encounter Value Set, Hospice Intervention Value Set, and Nonacute Inpatient Stay Value Set.

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.)*

N/A

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.)*

Note: The denominator for this measure is based on acute inpatient discharges and ED visits, not patients.

Step 1: Determine the eligible population: identify patients who meet the age criteria, with an ED visit or inpatient visit with a principal diagnosis of COPD, emphysema or chronic bronchitis

Step 2: Identify all COPD Episodes: for each patient identified in Step 1, identify all acute inpatient discharges and ED Visits. Multiple ED visits on the same date are counted as one COPD episode.

Step 3: Test for direct transfers.

Step 4: Determine the numerator:

Numerator 1 (Systemic Corticosteroid): identify the number of patients dispensed a prescription for a systemic corticosteroid on or 14 days after the Episode Date. Count systemic corticosteroids that are active on the relevant date.

Numerator 2 (Bronchodilator): identify the number of patients dispensed a prescription for a bronchodilator on or 30 days after the Episode Date. Count bronchodilators that are active on the relevant date.

Step 5: Calculate two rates.

A. Numerator 1/Denominator

B. Numerator 2/Denominator

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

N/A

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.)*

Specify calculation of response rates to be reported with performance measure results.

N/A

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Claims

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 are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

This measure is based on administrative claims collected in the course of providing care to health plan members. NCQA collects the Healthcare Effectiveness Data and Information Set (HEDIS) data for this measure directly from health plans via NCQA's online data submission system.

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)*

Health Plan

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Outpatient Services

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.)*

N/A

2. Validity – See attached Measure Testing Submission Form

PCE_nqf_testing_attachment_7.1.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. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

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. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

Measure Testing (subcriteria 2a2, 2b1-2b6)

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:**
Measure Title: **Pharmacotherapy Management of COPD Exacerbation**
Date of Submission: 8/15/2019

Type of Measure:

☐ **Outcome (*including PRO-PM*)**	☐ **Composite – *STOP – use composite testing form***
☐ Intermediate Clinical Outcome	☐ Cost/resource
☒ Process *(including Appropriate Use)*	☐ 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, 2b1, 2b2, and 2b4 must be completed.**
- **For outcome and resource use measures, section 2b3 also must be completed.**
- **If specified for multiple data sources/sets of specificaitons (e.g., claims and EHRs), section 2b5 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 (2b1-2b6) 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 25 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 social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.**

Note: The information provided in this form is intended to aid the Standing 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 instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

2b1. 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 instrument-based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; [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]

2b3. 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 social risk 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.**

2b4. 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.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. 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. The degree of consensus and any areas of disagreement must be provided/discussed.

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.17***)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☒ claims	☒ claims
☐ registry	☐ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other:	☐ other:

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*).

N/A

1.3. What are the dates of the data used in testing?

Initial testing: During measure development, we conducted a comprehensive field test in 2004 to assess feasibility of data collection and validity of performance data and critical data elements. This field test used data from measurement year 2003, which included health plan data spanning December 1, 2002 through January 31, 2004.

Systematic evaluation of face validity: The measure was tested for face validity throughout measure development from 2004 to 2006.

Measure score reliability and construct validity testing: We assessed measure score reliability and construct validity using data from all health plans that submitted HEDIS data to NCQA for this measure in 2018, which used data for measurement year 2017.

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.20***)	**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:

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*)

Initial testing: To assess feasibility of data collection and validity of performance data and critical data elements, 5 Commercial health plans, 1 Medicaid health plan and 3 Medicare health plans provided individual member-level data to NCQA for analysis. These plans were selected because they had the resources to generate the files, had sufficient sample of members with persistent asthma for analysis, and willingness to provide the data. The plans were geographically diverse and varied in size.

Systematic evaluation of face validity: Throughout the entire measure development process from 2004-2006, the measure was tested for face validity using panels of experts with specific clinical, methodologic and operational expertise. See Additional Information: Ad.1. Workgroup/Expert Panel Involved in Measure Development for names and affiliation of expert panel:

1) NCQA's Respiratory Measurement Advisory Panel (RMAP) is comprised of 10 experts (8 physicians, 1 pharmacist and 1 researcher) in clinical pulmonary care, including health care providers and policy makers.

2) NCQA's Technical Measurement Advisory Panel is a 12-member panel representing health plans methodologists, clinicians and HEDIS auditors.

3) NCQA's Committee on Performance Measurement (CPM) oversees the HEDIS measurement set and includes representation by purchasers, consumers, health plans, health care providers and policy makers. This panel is made up of 17 members. The CPM is organized and managed by NCQA and reports to the NCQA Board of Directors and is responsible for advising NCQA staff on the development and maintenance of performance measures. CPM members reflect the diversity of constituencies that performance measurement serves; some bring other perspectives and additional expertise in quality management and the science of measurement.

4) NCQA's HEDIS Expert Coding Panel reviewed and provided feedback on the vocabularies and definitions found in the values sets used to identify each measure component as well as the more recent mapping of ICD-9 codes to ICD-10 codes.

In 2005, the draft measure was posted for public comment, a 30-day period of review that allowed interested parties to offer feedback to NCQA about the measure. Stakeholders from various types of organizations submitted 67 comments on the measure.

2019 Update: Measure score reliability and construct validity were calculated from the 254 Commercial health plans, 390 Medicare plans and 201 Medicaid health plans that submitted data on this measure to HEDIS in 2018. The plans were geographically diverse and varied in size.

1.6. How many and which patients 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*)

Patient sample for initial measure field testing: We collected data from 5 Commercial health plans, 3 Medicare health plans and 1 Medicaid health plan. Below is a description of the sample. It includes the number of health plans that provided data for the measurement year 2003 and the median denominator for the measure across health plans. Note that the denominator is based on acute inpatient discharges and ED visits for COPD exacerbations, not members.

Product Type	Number of Plans	Median Number of Hospital/ED Visits per Plan
Commercial	5	86
Medicare	3	274
Medicaid	1	690

2019 Update

Patient sample for measure score reliability and construct validity: Data are summarized at the health plan level and stratified by product line (i.e. commercial, Medicare, Medicaid). Below is a description of the number of health plans that submitted data for this measure to HEDIS in 2018 (for measurement year 2017) and the median eligible population for the measure across health plans.

Product Type	Number of Plans	Median Number of Hospital/ED Visits per Plan
Commercial	254	178
Medicare	390	750
Medicaid	201	885

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.

During measure development, we conducted a comprehensive field test in 2004 to assess feasibility of data collection and validity of performance data and critical data elements using data submitted by 5 Commercial health plans, 1 Medicaid health plan and 3 Medicare health plans. The field test used data from measurement year 2003, which included health plan data spanning December 1, 2002 through January 31, 2004.

Face validity was demonstrated through a systematic assessment of face validity during measure development. Per NQF instructions we have described the composition of the technical expert panel, which assessed face validity in the data sample questions above.

2019 Update
No differences in the data used for reliability and construct validity testing.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

2019 and 2015 Submission:
Social risk factor data were not available in reported results. This measure is specified to be reported separately by Medicare, Medicaid and commercial plan types, which serves as a proxy for income and other socioeconomic factors.

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*)

2019 and 2015 Submission:
We utilized the Beta-binomial model (Adams 2009) to assess how well one can confidently distinguish the performance of one accountable entity from another. Conceptually, the Beta-binomial model is the ratio of signal to noise. The signal is the proportion of the variability in measured performance that can be explained by real differences in performance. The Beta-binomial model is an appropriate model when estimating the

reliability of simple pass/fail rate measures as is the case with most HEDIS measures. Reliability scores range from 0.0 to 1.0. A score of zero implies that all variation is attributed to measurement error (i.e., noise), whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance (across accountable entities).

Adams, J.L. The Reliability of Provider Profiling: A Tutorial. Santa Monica, California: RAND Corporation. TR-653-NCQA, 2009

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*)

2019 Update:

Reliability statistics for this measure were calculated using HEDIS health plan performance data for 2018. The results are as follows:

Beta-Binomial Statistics:

	Commercial			Medicare			Medicaid		
	Avg.	Overall	10th-90th	Avg.	Overall	10th-90th	Avg.	Overall	10th-90th
Bronchodilator Indicator	0.66	0.77	0.41-0.89	0.86	0.96	0.62-0.98	0.94	0.98	0.81-0.99
Systemic Corticosteroid Indicator	0.65	0.76	0.43-0.88	0.81	0.94	0.53-0.97	0.94	0.98	0.84-0.99

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?*)

Interpretation of measure score reliability testing:

Reliability scores can vary from 0.0 to 1.0. A score of zero implies that all variation is attributed to measurement error (noise) whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance (signal). Generally, a minimum reliability score of 0.7 is used to indicate sufficient signal strength to discriminate performance between accountable entities. The testing suggests that all indicators within this measure have good reliability between 0.7 and 1.0.

2b1. VALIDITY TESTING

2b1.1. 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 performance measure score 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*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. 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)*

METHOD OF ASSESSING FACE VALIDITY: NCQA has identified and refined measure management into a standardized process called the HEDIS measure life cycle.

STEP 1: NCQA staff identifies areas of interest or gaps in care. Clinical expert panels (MAPs—whose members are authorities on clinical priorities for measurement) participate in this process. Once topics are identified, a literature review is conducted to find supporting documentation on their importance, scientific soundness and feasibility. This information is gathered into a work-up format. The work-up is vetted by NCQA's Measurement Advisory Panels (MAPs), the Technical Measurement Advisory Panel (TMAP) and the Committee on Performance Measurement (CPM) as well as other panels as necessary.

This measure was developed in 2004 to assess whether patients who had a hospitalization or an emergency department visit for a COPD exacerbation were provided appropriate medication (systemic corticosteroids and bronchodilators) to treat symptoms and prevent future exacerbations. NCQA and the Respiratory Measurement Advisory Panel worked together to develop the most appropriate measure for assessing medication management of COPD exacerbations.

STEP 2: Development ensures that measures are fully defined and tested before the organization collects them. MAPs participate in this process by helping identify the best measures for assessing health care performance in clinical areas identified in the topic selection phase. Development includes the following tasks: (1) Prepare a detailed conceptual and operational work-up that includes a testing proposal and (2) Collaborate with health plans to conduct field-tests that assess the feasibility and validity of potential measures. The CPM uses testing results and proposed final specifications to determine if the measure will move forward to Public Comment.

The pharmacotherapy for COPD exacerbation measure was written and field-tested in 2004. After reviewing field test results, the CPM recommended to send the measure to public comment with a majority vote in 2005.

STEP 3: Public Comment is a 30-day period of review that allows interested parties to offer feedback to NCQA and the CPM about new measures or about changes to existing measures. NCQA MAPs and technical panels consider all comments and advise NCQA staff on appropriate recommendations brought to the CPM. The CPM reviews all comments before making a final decision about Public Comment measures. New measures and changes to existing measures approved by the CPM and NCQAs Board of Directors will be included in the next HEDIS year and reported as first-year measures.

The pharmacotherapy for COPD exacerbation measure was released for Public Comment in 2005 prior to publication in HEDIS. We received and responded to 67 comments on this measure. The CPM recommended moving this measure to first year data collection by a majority vote.

STEP 4: First-year data collection requires organizations to collect, be audited on and report these measures, but results are not publicly reported in the first year and are not included in NCQA's State of Health Care Quality, Quality Compass or in accreditation scoring. The first-year distinction guarantees that a measure can be effectively collected, reported and audited before it is used for public accountability or accreditation. This is not testing—the measure was already tested as part of its development—rather, it ensures that there are no unforeseen problems when the measure is implemented in the real world. NCQA's experience is that the first year of large-scale data collection often reveals unanticipated issues. After collection, reporting and auditing on a one-year introductory basis, NCQA conducts a detailed evaluation of first-year data. The CPM uses evaluation results to decide whether the measure should become publicly reportable or whether it needs further modifications.

The pharmacotherapy for COPD exacerbation measure was introduced to HEDIS in 2005. Organizations reported the measures in the first year and the results were analyzed for public reporting in the following year. The CPM recommended moving this measure to public reporting with a majority vote in 2006.

STEP 5: Public reporting is based on the first-year measure evaluation results. If the measure is approved, it will be publically reported and may be used for scoring in accreditation.

The pharmacotherapy for COPD exacerbation measure has been publicly reported in HEDIS since 2006.

STEP 6: Evaluation is the ongoing review of a measure's performance and recommendations for its modification or retirement. Every measure is reviewed for reevaluation at least every three years. NCQA staff continually monitors the performance of publicly reported measures. Statistical analysis, audit result review and user comments through NCQA's Policy Clarification Support portal contribute to measure refinement during re-evaluation. Information derived from analyzing the performance of existing measures is used to improve development of the next generation of measures.

Each year, NCQA prioritizes measures for re-evaluation and selected measures are researched for changes in clinical guidelines or in the health care delivery systems, and the results from previous years are analyzed. Measure work-ups are updated with new information gathered from the literature review, and the appropriate MAPs review the work-ups and the previous year's data. If necessary, the measure specification may be updated or the measure may be recommended for retirement. The CPM reviews recommendations from the evaluation process and approves or rejects the recommendation. If approved, the change is included in the next year's HEDIS Volume 2.

The clinical guideline recommendations for pharmacotherapy management following COPD exacerbation have not changed since the measure was developed in 2005; therefore, we have not made any significant changes to the measure since it was last endorsed on January 31, 2012.

Expert Participation

This measure was tested for face validity with input from three expert panels. Guidelines from the Global Initiative for Chronic Obstructive Lung Disease (GOLD) were also a strong authoritative source in applying the evidence for the measure.

We list an overview of each panel here. Please refer to Ad.1 in the submission form for the names and affiliation of experts in each panel.

1. Respiratory Measurement Advisory Panel includes 10 members (eight physicians, one pharmacist and a researcher) with expertise in respiratory care and quality measurement.
2. The Technical Measurement Advisory Panel includes 12 members, including representation by health plans, methodologists, clinician and auditors.
3. NCQA's Committee on Performance Measurement (CPM) oversees the evolution of the measurement set and includes representation by purchasers, consumers, health plans, health care providers and policy makers. This panel is made up of 16 members. The CPM is organized and managed by NCQA and reports to the NCQA Board of Directors and is responsible for advising NCQA staff on the development and maintenance of performance measures. CPM members reflect the diversity of constituencies that performance measurement serves; some bring other perspectives and additional expertise in quality management and the science of measurement.

<u>ICD-10 CONVERSION:</u>

In preparation for the national implementation of ICD-10 in 2015, NCQA conducted a systematic mapping of all value sets maintained by the organization to ensure the new values used for reporting maintained the reliability, validity and intent of the original specification.

Steps in ICD-9 to ICD-10 Conversion Process

1. NCQA first identified value sets within the measure that included ICD-9 codes. We used General Equivalence Mapping (GEM)to identify ICD-10 codes that map to ICD-9 codes and reviewed GEM mapping in both directions (ICD-9 to ICD-10 and ICD-10 to ICD-9) to identify potential trending issues.
2. NCQA then searched for additional codes (not identified by GEM mapping step) that should be considered due to the expansion of concepts in ICD-10. Using ICD-10 tabular list and ICD-10 Index, searches by diagnosis or procedure name were conducted to identify appropriate codes.
3. NCQA HEDIS Expert Coding Panel review: Updated value set recommendations were presented to for expert review and feedback.

4. NCQA RMAP clinical review: Due to increased specificity in ICD-10, new codes and definitions require review to confirm the diagnosis or procedure is consistent and appropriate given the scope of the measure.
5. New value sets containing ICD-10 code recommendations were for public review and comment in 2014 and updated in 2015. Comments received were reconciled with additional feedback from HEDIS Expert Coding Panel and MAPs as needed.
6. NCQA staff finalized value sets containing ICD-10 codes for publication in 2015.

Tools Used to Identify/Map to ICD-10

All tools used for mapping/code identification from CMS ICD-10 website (http://www.cms.gov/Medicare/Coding/ICD10/2012-ICD-10-CM-and-GEMs.html).

GEM, ICD-10 Guidelines, ICD-10-CM Tabular List of Diseases and Injuries, ICD-10-PCS Tabular List.

Expert Participation

The NCQA HEDIS Expert Coding Panel reviewed and provided feedback on staff recommendations. Names and credentials of the experts who served on these panels are listed under **Additional Information, Ad. 1. Workgroup/Expert Panel Involved in Measure Development**.

METHOD OF TESTING CRITICAL DATA ELEMENT VALIDITY: For the initial field test in 2004, validity was tested by comparing the presence of administrative claims codes for patients who had a COPD exacerbation managed in the emergency department or hospital and were discharged home (required to calculate the denominator) to documentation in the medical record, which is considered to be the "gold standard". The plans also looked at administrative claims codes for patients who had a systemic corticosteroid prescription or bronchodilator prescription (required to calculate the numerator) and searched for documentation in the medical record.

2019 Update:

METHOD OF TESTING CONSTRUCT VALIDITY**:** We tested for construct validity by exploring whether this measure was correlated with other similarly constructed process measures. We hypothesized that organizations that perform well on the measure should perform well on other similar HEDIS measures. To test these correlations, we used a Pearson correlation test. This test estimates the strength of the linear association between two continuous variables; the magnitude of correlation ranges from -1 and +1. A value of 1 indicates a perfect linear dependence in which increasing values on one variable is associated with increasing values of the second variable. A value of 0 indicates no linear association. A value of -1 indicates a perfect linear relationship in which increasing values of the first variable is associated with decreasing values of the second variable.

For this measure, we specifically hypothesized:

1. Performance on the PCE bronchodilator indicator (percent of patients who were dispensed a prescription for a bronchodilator within 30 days after an acute COPD exacerbation) would be positively correlated with performance on the corticosteroid indicator (percent of patients who were dispensed a prescription for a systemic corticosteroid within 14 days after an acute COPD exacerbation). Examining this correlation would help contribute to the validity of the measure, because both indicators assess follow-up treatment for an acute COPD exacerbation.
2. Performance on both the PCE bronchodilator and corticosteroid indicators would be positively correlated with performance on the Statin Therapy for Patients with Cardiovascular Disease measure. Examining this correlation would help contribute to the validity of the measure, because both measures assess medication therapy following a chronic disease episode.

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

RESULTS OF FACE VALIDITY ASSESSMENT:

For the initial field test conducted in 2004, we calculated the total measure performance rate, as well as performance rates based on whether patients had an active prescription that counted toward the numerator or received a new prescription. We discussed the validity of the performance results with the expert panels (Respiratory Measurement Advisory Panel and the Committee on Performance Measurement). Among the different product lines, performance rates ranged from 36 to 47 percent for the systemic corticosteroid

indicator and 45 to 64 percent for the bronchodilator indicator. Between 9 to 14 percent of the discharged members had an active systemic corticosteroid prescription and 26 to 40 percent of discharged members had an active bronchodilator prescription. Performance rates on both indicators were slightly lower for men compared to women and for people ages 75+ compared to ages 40-74. The expert panels agreed that the performance on the indicators were an accurate representation of quality performance and distinguished performance among health plans.

2004 Field Test: Performance Rates on the Systemic Corticosteroid Indicator by Product Line, Age and Gender*

	Denom.	Total Num.	Total Perf. Rate	Active Steroid Rx		New Steroid Rx	
				Num.	Perf. Rate	Num.	Perf. Rate
Product Line							
Commercial	1,085	510	47.0%	127	11.7%	434	40.0%
Medicaid	690	292	42.3%	98	14.2%	234	33.9%
Medicare	693	246	35.5%	61	8.8%	211	30.4%
Age							
40-54	601	244	40.6%	66	11.0%	204	33.9%
55-64	782	394	50.4%	108	13.8%	336	43.0%
65-74	612	258	42.2%	65	10.6%	221	36.1%
75-84	374	125	33.4%	35	9.4%	102	27.3%
85+	99	27	27.3%	12	12.1%	16	16.2%
Gender							
F	1382	602	43.6%	165	11.9%	507	36.7%
M	1086	446	41.1%	121	11.1%	372	34.3%
Total	2468	1048	42.5%	286	11.6%	879	35.6%

**Includes data submitted by 5 Commercial plans, 3 Medicare plans and 1 Medicaid plan using measurement year 2003*

2004 Field Test: Performance Rates on the Bronchodilator Indicator by Product Line, Age and Gender*

	Denom.	Total Num.	Total Perf. Rate	Active Bronch. Rx		New Bronch. Rx	
				Num.	Perf. Rate	Num.	Perf. Rate
Product Line							
Commercial	1,085	571	52.6%	303	27.9%	437	40.3%
Medicaid	690	440	63.8%	275	39.9%	309	44.8%
Medicare	693	312	45.0%	178	25.7%	231	33.3%
Age							
40-54	601	307	51.1%	163	27.1%	238	39.6%
55-64	782	473	60.5%	272	34.8%	346	44.2%
65-74	612	333	54.4%	197	32.2%	241	39.4%
75-84	374	172	46.0%	103	27.5%	124	33.2%
85+	99	38	38.4%	21	21.2%	28	28.3%
Gender							
F	1382	751	54.3%	432	31.3%	538	38.9%
M	1086	572	52.7%	324	29.8%	439	40.4%
Total	2468	1323	53.6%	756	30.6%	977	39.6%

**Includes data submitted by 5 Commercial plans, 3 Medicare plans and 1 Medicaid plan using measurement year 2003*

RESULTS OF CRITICAL DATA ELEMENT VALIDTY: Across four plans, validation of a COPD exacerbation in the medical record was 49%, with a range of 36% to 71%. The health plans were instructed to review hospital records in addition to primary care records to confirm an exacerbation, but feedback from the majority of the plans indicated that chart abstractors were not able to review hospital records and relied on primary care records to note confirmations. Since hospital records are known to be more reliable in documenting care provided in the ED/hospital, this may explain the moderate rate of denominator validation.

2004 Field Test: COPD Exacerbation Medical Record Validation by Plan and Product Line*

	Count of COPD exacerbations confirmed by administrative data	% of patients that had documentation of a COPD exacerbation in medical record	% of patients that did not have documentation of a COPD exacerbation in medical record
Plan:			
A	137	35.8%	51.1%
B	41	46.0%	48.8%
C	51	70.8%	27.5%
D	140	53.1%	43.6%
Total	369	49.1%	44.7%
Product Line:			
Commercial	240	48.8%	43.8%
Medicare	129	49.6%	46.5%

**Includes data submitted by 5 Commercial plans and 3 Medicare plans using measurement year 2003*

In four plans, there was 64.2% data consistency for steroid use between administrative and medical record data. This was calculated by adding the percent of steroid use data found in administrative data and medical record data plus the percent of steroids found in neither data source. Of note, a higher percent of steroid use data noted in medical record was not captured via administrative data. This may be due to written prescriptions for steroids that were never filled by the patient. There was 66.6% consistency for bronchodilator use between administrative and medical record data.

2004 Field Test: Systemic Corticosteroid Indicator Validation by Plan*

Plan Code	# of patients with a COPD exacerbation confirmed in both admin & medical record data	% of patients with corticosteroid confirmed in both medical record & admin data	% of patients with corticosteroid confirmed in neither admin or medical record data	% of patients with corticosteroid confirmed in admin data only	% of patients with corticosteroid confirmed in medical record data only
A	39	56.4%	17.9%	5.1%	20.5%
B	17	23.5%	29.4%	0.0%	47.1%
C	34	29.4%	14.7%	5.9%	50.0%
D	69	40.6%	30.4%	10.1%	18.8%
Total	159	40.3%	23.9%	6.9%	28.9%

**Includes data submitted by 5 Commercial plans and 3 Medicare plans using measurement year 2003*

2004 Field Test: Bronchodilator Indicator Validation by Plan*

Plan Code	# of patients with a COPD exacerbation confirmed in both admin & medical record data	% of patients with bronchodilator confirmed in both medical record & admin data	% of patients with bronchodilator confirmed in neither admin or medical record data	% of patients with bronchodilator confirmed in admin data only	% of patients with bronchodilator confirmed in medical record data only

A	39	64.1%	0.0%	2.6%	33.3%
B	17	35.3%	29.4%	5.9%	29.4%
C	34	50.0%	2.9%	2.9%	44.1%
D	69	52.2%	23.2%	11.6%	13.0%
Total	159	52.8%	13.8%	6.9%	26.4%

**Includes data submitted by 5 Commercial plans and 3 Medicare plans using measurement year 2003*

2019 Update:

STATISTICAL RESULTS OF CONSTRUCT VALIDITY

Health Plan Level Pearson Correlation Coefficients Among Pharmacotherapy Management of COPD Exacerbation Within Measure – **Commercial** Plans, 2017*

	Dispensed a systemic corticosteroid within 14 days
Dispensed a bronchodilator within 30 days	0.52

**Includes data submitted by 254 Commercial plans using measurement year 2017*

All scores were significant at p<0.05

Health Plan Level Pearson Correlation Coefficients Among Pharmacotherapy Management of COPD Exacerbation Within Measure – **Medicaid** Plans, 2017*

	Dispensed a systemic corticosteroid within 14 days
Dispensed a bronchodilator within 30 days	0.82

**Includes data submitted by 201 Medicaid plans using measurement year 2017*

All scores were significant at p<0.05

Health Plan Level Pearson Correlation Coefficients Among Pharmacotherapy Management of COPD Exacerbation Within Measure – **Medicare** Plans, 2017*

	Dispensed a systemic corticosteroid within 14 days
Dispensed a bronchodilator within 30 days	0.45

**Includes data submitted by 390 Medicare plans using measurement year 2017*

All scores were significant at p<0.05

Health Plan Level Pearson Correlation Coefficients Among Pharmacotherapy Management of COPD Exacerbation and Statin Therapy for Patients With Cardiovascular Disease – **Commercial** Plans, 2017*

	Statin Therapy Measure	
COPD Measure	Received Statin Therapy - Total	Statin Adherence 80% - Total

Dispensed a bronchodilator within 30 days	0.42	0.47
Dispensed a systemic corticosteroid within 14 days	0.48	0.31

**Includes data submitted by 254 Commercial plans using measurement year 2017*

All scores were significant at p<0.05

Health Plan Level Pearson Correlation Coefficients Among Pharmacotherapy Management of COPD Exacerbation and Statin Therapy for Patients With Cardiovascular Disease – **Medicaid** Plans, 2017*

	Statin Therapy Measure	
COPD Measure	Received Statin Therapy - Total	Statin Adherence 80% - Total
Dispensed a bronchodilator within 30 days	0.68	0.36
Dispensed a systemic corticosteroid within 14 days	0.66	0.40

**Includes data submitted by 185 Medicaid plans using measurement year 2017*

All scores were significant at p<0.05

Health Plan Level Pearson Correlation Coefficients Among Pharmacotherapy Management of COPD Exacerbation and Statin Therapy for Patients With Cardiovascular Disease – **Medicare** Plans, 2017*

	Statin Therapy Measure	
COPD Measure	Received Statin Therapy - Total	Statin Adherence 80% - Total
Dispensed a bronchodilator within 30 days	0.26	0.25
Dispensed a systemic corticosteroid within 14 days	0.25	0.29

**Includes data submitted by 374 Medicare plans using measurement year 2017*

All scores were significant at p<0.05

2b1.4. 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?*)

SYSTEMATIC ASSESSMENT OF FACE VALIDITY: The pharmacotherapy management of COPD exacerbation measure was deemed to have the desirable attributes of a HEDIS measure in 2005 (relevance, scientific soundness, and feasibility). These results indicate the technical expert panels showed good agreement that the measure as specified will accurately differentiate quality across providers. The technical

expert panels have also reviewed the measure for relevance, scientific soundness, and feasibility at regular intervals since 2005. Our interpretation is that this measure has sufficient face validity.

CRITICAL DATA ELEMENT VALIDTY: The results of the critical data element validity testing demonstrate that the administrative data elements used to calculate the measure denominator (patients who had a COPD exacerbation managed in the emergency department or hospital and were discharged home) and numerator (patients who had a systemic corticosteroid or bronchodilator prescription) had moderate to strong agreement with medical record data and are valid.

2019 Update:

CONSTRUCT VALIDITY: Coefficients with absolute value of less than 0.3 are generally considered indicative of weak associations whereas absolute values of 0.3 or higher denote moderate to strong associations. The significance of a correlation coefficient is evaluated by testing the hypothesis that an observed coefficient calculated for the sample is different from zero. The resulting p-value indicates the probability of obtaining a difference at least as large as the one observed due to chance alone. We used a threshold of 0.05 to evaluate the test results. P-values less than this threshold imply that it is unlikely that a non-zero coefficient was observed due to chance alone.

The results confirmed that the hypothesis that the measure's bronchodilator indicator (percent of patients who were dispensed a prescription for a bronchodilator within 30 days after an acute COPD exacerbation) and corticosteroid indicator (percent of patients who were dispensed a prescription for a systemic corticosteroid within 14 days after an acute COPD exacerbation) are correlated with each other, suggesting they represent the same underlying quality construct of respiratory quality of care. Both indicators were also correlated with the statin therapy measure, confirming the hypothesis that health plans with good pharmacotherapy management for COPD exacerbations also have better performance on providing statin therapy for patients with cardiovascular disease, suggesting they represent the same underlying quality construct of chronic disease quality of care.

2b2. EXCLUSIONS ANALYSIS

NA ☒ no exclusions — *skip to section 2b3*

NCQA has a policy for excluding hospice patients from HEDIS measures.

2b2.1. 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*)

N/A

2b2.2. 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*)
N/A

2b2.3. 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*)
N/A

2b3. 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 2b4.

2b3.1. 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,**

2b3.1.1 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.

2b3.2. 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.

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk 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*) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☐ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

2b3.4a. What were the statistical results of the analyses used to select risk factors?

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk 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.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

2b3.5. 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.
If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

2b3.9. Results of Risk Stratification Analysis:

2b3.10. 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*)

2b3.11. 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*)

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. 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)*

To demonstrate meaningful differences in performance, NCQA calculates an inter-quartile range (IQR) for each indicator. The IQR provides a measure of the dispersion of performance. The IQR can be interpreted as the difference between the 25th and 75th percentile on a measure. To determine if this difference is statistically significant, NCQA calculates an independent sample t-test of the performance difference between two randomly selected plans at the 25th and 75th percentile. The t-test method calculates a testing statistic based on the sample size, performance rate, and standardized error of each plan. The test statistic is then compared against a normal distribution. If the p value of the test statistic is less than .05, then the two plans' performance is significantly different from each other. Using this method, we compared the performance rates of two randomly selected plans, one plan in the 25th percentile and another plan in the 75th percentile of performance. We used these two plans as examples of measured entities. However the method can be used for comparison of any two measured entities.

2b4.2. 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*)

HEDIS 2018 Variation in Performance across Health Plans: Bronchodilator Indicator

	Avg. EP	Avg.	SD	10th	25th	50th	75th	90th	IQR	p-value
Commercial	178	80.1	7.0	71.9	76.3	80.4	84.3	88.3	8.0	<0.002
Medicare	750	79.6	8.2	71.2	75.7	79.9	84.9	89.5	9.2	<0.001
Medicaid	885	81.5	10.1	70.6	78.8	83.8	87.6	89.7	8.8	<0.001

EP: Eligible Population, the average denominator size across all plans submitting 2018 HEDIS data for this measure

IQR: Interquartile range

p-value: P-value of independent samples t-test comparing plans at the 25th percentile to plans at the 75th percentile.

HEDIS 2018 Variation in Performance across Health Plans: Systemic Corticosteroids Indicator

	Avg. EP	Avg.	SD	10th	25th	50th	75th	90th	IQR	p-value
Commercial	178	75.5	7.6	66.7	71.7	75.7	80.0	84.0	8.3	<0.001
Medicare	750	70.4	7.8	62.5	67.5	71.6	74.7	78.4	7.2	<0.001
Medicaid	885	68.3	11.3	53.6	63.0	70.2	76.2	80.4	13.2	<0.001

EP: Eligible Population, the average denominator size across all plans submitting 2018 HEDIS data for this measure

IQR: Interquartile range

p-value: P-value of independent samples t-test comparing plans at the 25th percentile to plans at the 75th percentile.

2b4.3. 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 results above indicate there is an 7-13% gap in performance between the 25th and 75th percentile performing plans across the different product lines and indicators. For most product lines and indicators, the difference between the 25th and 75th percentile performance rates is statistically significant. The highest variation in performance is in Medicare and Medicaid plans, which shows a 9-percentage point gap between 25th and 75th percentile Medicare plans for the bronchodilator indicator and a 13-percentage point gap between Medicaid plans for the systemic corticosteroid indicator.

To put these meaningful differences in performance into context, we estimated that on average 69 additional members per Medicare plan would have been discharged on a bronchodilator and 116 additional members per Medicaid plan would have been discharged on a systemic corticosteroid if plans in the 25th percentile performed as well as plans in the 75th percentile. This estimate is based on the average health plan eligible population.

2b5. 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 social risk 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 social risk 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.***

2b5.1. 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*)

N/A

2b5.2. 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*)

N/A

2b5.3. 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*)
N/A

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. 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*)

HEDIS measures apply to enrolled members in a health plan, and NCQA has a rigorous audit process to ensure the eligible population and numerator events for each measure are correctly identified and reported. The audit process is designed to verify primary data sources used to populate measures and ensure specifications are correctly implemented.

The HEDIS Compliance Audit addresses the following functions:

- Information practices and control procedures
- Sampling methods and procedures
- Data integrity
- Compliance with HEDIS specifications
- Analytic file production
- Reporting and documentation

2b6.2. 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*)

HEDIS addresses missing data in a structured way through its audit process. HEDIS measures apply to enrolled members in a health plan, and NCQA-certified auditors use standard audit methodologies to assess whether data sources are missing data. If a data source is found to be missing data, and the issues cannot be rectified, the auditor will assign a "materially biased" designation to the measure for that reporting plan, and the rate will not be used. Once measures are added to HEDIS, NCQA conducts a first-year analysis to assess the feasibility of the measure when widely implemented in the field. This analysis includes an assessment of how many plans report valid rates vs. rates that are materially biased (or have other issues, such as small denominators). These considerations are weighed in the deliberation process before measures are approved for public reporting.

2b6.3. 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*)

All of the commercial, Medicaid, and Medicare health plans that reported 2018 HEDIS data for this measure reported valid rates as determined by NCQA-certified auditors through the process described above. This means that auditors did not find any missing data sources for any of the health plan data submissions and determined that none of the rates were materially biased.

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), Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

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 claims

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).

N/A

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:

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 instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

NCQA conducts an independent audit of all HEDIS collection and reporting processes, as well as an audit of the data which are manipulated by those processes, in order to verify that HEDIS specifications are met. NCQA has developed a precise, standardized methodology for verifying the integrity of HEDIS collection and calculation processes through a two-part program consisting of an overall information systems capabilities assessment followed by an evaluation of the managed care organization's ability to comply with HEDIS specifications. NCQA-certified auditors using standard audit methodologies will help enable purchasers to make more reliable comparisons between health plans.

The HEDIS Compliance Audit addresses the following functions:

1) Information practices and control procedures
2) Sampling methods and procedures
3) Data integrity
4) Compliance with HEDIS specifications
5) Analytic file production
6) Reporting and documentation

In addition to the HEDIS audit, NCQA provides a system to allow "real-time" feedback from measure users. Our Policy Clarification Support System receives thousands of inquiries each year on over 100 measures. Through this system, NCQA responds immediately to questions and identifies possible errors or inconsistencies in the implementation of the measure. This system informs both annual updates to the measures as well as routine re-evaluation of measures. These processes include updating value sets and clarifying the specifications. Measures are re-evaluated on a periodic basis and when there is a significant change in evidence.

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)*.

Broad public use and dissemination of these measures are encouraged and NCQA has agreed with NQF that noncommercial uses do not require the consent of the measure developer. Use by health care physicians in connection with their own practices is not commercial use. Commercial use of a measure requires the prior written consent of NCQA. As used herein, "commercial use" refers to any sale, license, or distribution of a measure for commercial gain, or incorporation of a measure into any product or service that is sold, licensed, or distributed for commercial gain, even if there is no actual charge for inclusion of the measure.

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 Health Plan Rating https://www.ncqa.org/hedis/reports-and-research/ratings-2019/ Health Plan Rating https://www.ncqa.org/hedis/reports-and-research/ratings-2019/ Payment Program Medicare Advantage Plan Rating https://www.medicare.gov/find-a-plan/questions/home.aspx NCQA Health Plan Accreditation http://www.ncqa.org/tabid/123/Default.aspx Regulatory and Accreditation Programs NCQA Health Plan Accreditation https://www.ncqa.org/programs/health-plans/health-plan-accreditation-hpa/ Quality Improvement (external benchmarking to organizations) NCQA Quality Compass https://www.ncqa.org/programs/data-and-information-technology/data-purchase-and-licensing/quality-compass/

4a1.1 For each CURRENT use, checked above (update for <u>maintenance of endorsement</u>), provide:

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

NCQA HEALTH PLAN RATINGS/REPORT CARDS: This measure is used in the calculation of health plan ratings, which are reported on the NCQA website annually. These ratings are based on performance on HEDIS measures among other factors.In 2019, a total of 255 Medicare health plans, 515 commercial health plans and 188 Medicaid health plans across 50 states were included in the rankings.
NCQA HEALTH PLAN ACCREDITATION: This measure is used in scoring for accreditation of Medicare Advantage Health Plans. As of Fall 2017, a total of 184 Medicare Advantage health plans were scored for accreditation using this measure among others covering 9.2 million Medicare beneficiaries; 451 commercial health plans covering 113 million lives; and 125 Medicaid health plans covering 35 million lives. Health plans are scored based on performance compared to national benchmarks.
NCQA QUALITY COMPASS: This measure is used in Quality Compass which is an indispensable tool used for selecting health plans, conducting competitor analysis, examining quality improvement and benchmarking plan performance. Provided in this tool is the ability to generate custom reports by selecting plans, measures, and benchmarks (averages and percentiles) for up to three trended years. Results in table and graph formats offer simple comparison of plans' performance against competitors or benchmarks.

4a1.2. 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

4a1.3. 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.)

N/A

4a2.1.1. 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.

Health plans that report HEDIS calculate their rates and know their performance when submitting to NCQA. NCQA publicly reports rates across all plans and also creates benchmarks in order to help plans understand how they perform relative to other plans. Public reporting and benchmarking are effective quality improvement methods.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

NCQA publishes HEDIS results annually in our Quality Compass tool. NCQA also presents data at various conferences and webinars. For example, at the annual HEDIS Update and Best Practices Conference (now the Health Care Quality Congress), NCQA presents results from all new measures' first year of implementation or analyses from measures that have changed significantly and insight into new measure development projects. NCQA also regularly provides technical assistance on measures through its Policy Clarification Support System, as described in Section 3c.1.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

NCQA measures are evaluated regularly using a consensus-based process to consider input from multiple stakeholders, including but not limited to entities being measured. We use several methods to obtain input, including vetting of the measure with several multi-stakeholder advisory panels, 30-day public comment posting, and review of questions submitted to the Policy Clarification Support System. This information enables NCQA to comprehensively assess a measure's adherence to the HEDIS Desirable Attributes of Relevance, Scientific Soundness and Feasibility.

4a2.2.2. Summarize the feedback obtained from those being measured.

Questions received through the Policy Clarification Support system have generally centered around clarification on whether certain notation in medical record documentation is sufficient to meet measure criteria. Many of the questions ask about various scenarios concerning emergency department or inpatient visits, and whether the patient should be included in the denominator in that scenario. Other questions have sought clarification about different systemic corticosteroid and/bronchodilator prescribing or dispensing scenarios, and whether they satisfy the measure numerator.

4a2.2.3. Summarize the feedback obtained from other users

This measure has been deemed a priority measure by NCQA, as illustrated by its use in programs such as Health Plan Rating, NCQA Accreditation and Quality Compass. States, employers and regional health quality organizations value this measure (and other HEDIS measures) for shining a light on quality.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

Feedback obtained through the mechanisms described in 4a2.2.1 informed how we revised the measure to clarify how to identify which inpatient and ED visits should be included in the denominator.

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.

4b1. 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.

The performance rates, for both numerators, were higher in 2017 than in 2016, across all product lines. From 2017 to 2018, some of the same rates showed modest increases, while others showed modest decreases, making it difficult to identify a trend for the last 3 years. For the systemic corticosteroid rate, commercial and Medicare plans consistently report higher rates than Medicaid. For the bronchodilator rate, Medicaid plans consistently report higher rates than commercial and Medicare plans.

4b2. 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).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

There were no identified unintended findings for this measure during testing or since implementation.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

There were no identified unexpected benefits for this measure during testing or since implementation.

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)

0102 : COPD: inhaled bronchodilator therapy

0577 : Use of Spirometry Testing in the Assessment and Diagnosis of COPD

1825 : COPD - Management of Poorly Controlled COPD

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

N/A

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?

Yes

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

For all three related measures, there is no impact on interpretability or added burden of data collection because the focus of this measure is different. For the measures that report use of pharmacotherapy for COPD, the denominator focuses on all adults, whereas this measure focuses on older adults (40 years and over). 0102 (similar numerator, different denominator) 0102's numerator is prescription of an inhaled corticosteroid. The denominator includes certain COPD patients 18 years or older. Unlike this measure, the level of analysis for 0102 is the clinician. 0577 (different numerator, similar denominator) 0577's numerator is presence of a spirometry test to confirm a new or newly active COPD diagnosis. The denominator is persons 40 years or older with a new or newly active diagnosis of COPD. 1825 (somewhat similar numerator, different denominator) 1825's numerator is patients 18 years or older who are taking a long-acting bronchodilator. The denominator includes all patients 18 years or older with poorly controlled COPD who are taking a short-acting bronchodilator.

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.)

N/A

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): National Committee for Quality Assurance

Co.2 Point of Contact: Bob, Rehm, nqf@ncqa.org, 202-955-1728-

Co.3 Measure Developer if different from Measure Steward: National Committee for Quality Assurance

Co.4 Point of Contact: Bob, Rehm, nqf@ncqa.org, 202-955--

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.

RESPIRATORY MEASUREMENT ADVISORY PANEL (RMAP) MEMBERS:

David Au, MD, MS, (CHAIR) Associate Prof. of Medicine, VA Puget Sound Health Care System

Kurt Elward, MD, MPH, Senior Medical Director, Innovation Health

Laura Feemster, MD, MS, Investigator/Staff Physician, University of Washington Medical Center

Anne Fuhlbrigge, MD, MS, Senior Associate Dean for Clinical Affairs, University of Colorado School of Medicine

Min Joo, MD, MPH, FCCP, Assistant Professor of Medicine, University of Illinois at Chicago/ Jesse Brown VA Medical Center

Christine Joseph, PhD, MPH, BSc, Associate Director of Research, Epidemiologist, Henry Ford Health System

Todd Lee, PharmD, PhD, Primary: Senior Investigator, Secondary: Associate Professor, University of Illinois at Chicago

Allan Luskin, MD, President, Healthy Airways

Richard O´Connor, MD, Director, Dept. of Quality Management, Allergist/Immunologist, Sharp Rees-Stealy Medical Group

COMMITTEE ON PERFORMANCE MEASUREMENT (CPM) MEMBERS:

Andrew Baskin, MD, National Medical Director, Aetna

Elizabeth Drye, MD, SM, Senior Director, Yale/CORE; Research Scientist, Pediatrics, Yale University

Andrea Gelzer, MD, MS, FACP, Senior VP, Medical Affairs, AmeriHealth Caritas

Kate Goodrich, MD, MHS, Chief Medical Officer and Director, CCSQ, CMS

David Grossman, MD, MPH, Senior Associate Medical Director, Washington Permanente Medical Group

Christine S. Hunter, MD (Co- Chair), Independent Board Director, WPS Health Solutions

David K. Kelley, MD, MPA, Chief Medical Officer, Office of Medical Assistance Programs, Pennsylvania Department of Human Services

Jeffery Kelman, MMSc, MD, Chief Medical Officer, United States Department of Health and Human Services

Nancy Lane, PhD, Independent Consultant

Bernadette Loftus, MD, Freelancer

Adrienne Mims, MD, MPH, AGSF, FAAFP, VP, Chief Medical Officer, Alliant Health Solutions

Amanda Parsons, MD, MBA, Deputy Chief Medical Officer, Metroplus

Wayne Rawlins, MD, MBA, Chief Medical Officer, ConnectiCare

Misty Roberts, MSN, RN, CPHQ, PMP, Associate Vice President, Clinical Quality Officer, Humana

Rudy Saenz, MD, MMM, FACOG, Physician, Medical Director of Quality, Riverside Medical Clinic

Marcus Thygeson, MD, MPH (Co-Chair), Chief Health Officer, Bind On-Demand

JoAnn Volk, MA, Research Professor, Georgetown University

HEDIS EXPERT CODING PANEL MEMBERS:

Glen Braden, MBA, CHCA, Attest Health Care Advisors, LLC

Denene Harper, RHIA, American Hospital Association

DeHandro Hayden, BS, American Medical Association

Patience Hoag, RHIT, CPHQ, CHCA, CCS, CCS-P, Health Services Advisory Group

Nelly Leon-Chisen, RHIA, American Hospital Association

Alec McLure, RHIA, CCS-P, Verisk Health

Michele Mouradian, RN, BSN, McKesson Health Solutions

Craig Thacker, RN, CIGNA HealthCare

Mary Jane F. Toomey, RN CPC, Aetna Better Health

HEDIS EXPERT PHARMACY PANEL MEMBERS:

Gerry Hobson, RPh, Cerner Multum

Chronis Manolis, RPh, UPMC Health Plan

Cathrine Misquitta, PharmD, MBA, BCPS, BCGP, FCSHP, Envolve Pharmacy Solutions

Kevin Park, MD, Care Wisconsin

TECHNICAL MEASUREMENT ADVISORY PANEL MEMBERS:

Andy Amster, MSPH, Senior Director, Kaiser Permanente

Jennifer Brudnicki, MBA, Product Services Manager, Inovalon

Lindsay Cogan, PhD, MS, Director, Division of Quality Measurement, New York State Department of Health

Kathryn Coltin, MPH, Independent Consultant

Michael Farina, R.Ph, MBA, Director, Health Care Quality, Capital District Physicians' Health Plan

Marissa Finn, MBA, CIGNA

Scott Fox, MS, MEd, Principal, Payment Reform, The MITRE Corporation

Carlos Hernandez, Director, Quality Management, CenCalHealth

Harmon Jordan, ScD, Senior Study Director, Westat

Gigi Raney, LCSW, Center for Medicaid and CHIP Services

Lynne Rothney-Kozlak, MPH, President, Rothney-Kozlak Consulting, LLC

Laurie Spoll, Director, Aetna

The NCQA Respiratory Measurement Advisory Panel advised NCQA during measure development. They evaluated the way staff specified the measure, reviewed field test results, and assessed NCQA's overall desirable attributes of Relevance, Scientific Soundness, and Feasibility. The advisory panel consisted of a balanced group of experts. In addition to this advisory panel, we vetted the measure with a host of other stakeholders, as is our process. Thus, our measures are the result of consensus from a broad and diverse group of stakeholders.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2005

Ad.3 Month and Year of most recent revision: 07, 2019

Ad.4 What is your frequency for review/update of this measure? Approximately every 3 years, sooner if the clinical guidelines have changed significantly.

Ad.5 When is the next scheduled review/update for this measure? 12, 2020

these materials and can rescind or alter these materials at any time. These materials may not be modified by anyone other than NCQA. Anyone desiring to use or reproduce the materials without modification for a non-commercial purpose may do so without obtaining any approval from NCQA. All commercial uses must be approved by NCQA and are subject to a license at the discretion of NCQA.

Calculated measure results, based on unadjusted HEDIS specifications, may not be termed "Health Plan HEDIS rates" until they are audited and designated reportable by an NCQA-Certified Auditor. Such unaudited results should be referred to as "Unaudited Health Plan HEDIS Rates." Accordingly, "Heath Plan HEDIS rate" refers to and assumes a result from an unadjusted HEDIS specification that has been audited by an NCQA-Certified HEDIS Auditor.

Limited proprietary coding is contained in the measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. NCQA disclaims all liability for use or accuracy of any coding contained in the specifications.

Content reproduced with permission from HEDIS, Volume 2: Technical Specifications for Health Plans. To purchase copies of this publication, including the full measures and specifications, contact NCQA Customer Support at 888-275-7585 or visit

www.ncqa.org/publications.

Ad.7 Disclaimers: This HEDIS® performance measure is not a clinical guideline and does not establish a standard of medical care and has not been tested for all potential applications.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

Ad.8 Additional Information/Comments: NCQA Notice of Use. Broad public use and dissemination of these measures, without modification, are encouraged and NCQA has agreed with NQF that noncommercial uses do not require the consent of the measure developer. Modifications to, and/or commercial use of, a measure requires the prior written consent of NCQA and is subject to a license at the discretion of NCQA. As used herein, "commercial use" refers to any sale, license, or distribution of a measure for commercial gain, or incorporation of a measure into any product or service that is sold, licensed, or distributed for commercial gain, even if there is no actual charge for inclusion of the measure.