NATIONAL QUALITY FORUM
Driving measurable health improvements together

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 #: 0041

Measure Title: Preventive Care and Screening: Influenza Immunization

Measure Steward: National Committee for Quality Assurance

Brief Description of Measure: Percentage of patients aged 6 months and older seen for a visit between October 1 and March 31 who received an influenza immunization OR who reported previous receipt of an influenza immunization

Developer Rationale: Influenza may lead to serious complications including hospitalization or death (1). Influenza vaccination is the most effective protection against influenza virus infection (1). However, data indicate that less than half of all eligible individuals receive an influenza vaccination (2). This measure promotes annual influenza vaccination for all persons aged >= 6 months.

1. Seasonal Influenza: Flu Basics. Centers for Disease Control and Prevention Web site. http://www.cdc.gov/flu/about/disease/index.htm. Updated May 4, 2016. Accessed June 23, 2016.
2. Flu vaccination coverage: United States, 2014-15 Influenza Season. Centers for Disease Control and Prevention Web site. http://www.cdc.gov/flu/fluvaxview/coverage-1415estimates.htm. Updated September 17, 2015. Accessed June 23, 2016.

Numerator Statement: Patients who received an influenza immunization OR who reported previous receipt of an influenza immunization.

Denominator Statement: All patients aged 6 months and older seen for a visit between October 1 and March 31.

Denominator Exclusions: None

Measure Type: Process

Data Source: Registry Data, Claims

Level of Analysis: Clinician: Individual

IF Endorsement Maintenance – Original Endorsement Date: 2009-08-10

Most Recent Endorsement Date: 2017-01-17

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement, endorsed measures are evaluated periodically to ensure that the measure 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 are 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 description for this measure:

- This is a maintenance process measure at the clinician level of analysis that measures the percentage of patients aged 6 months and older seen for a visit between October 1 and March 31 who received an influenza immunization OR who reported previous receipt of an influenza immunization.
- The developer provided a logic model that depicts the administration of Influenza immunization leads to the prevention of and reduction in the severity of influenza illness.

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

Summary of prior review in 2016

- The developer referenced clinical practice guidelines and used the recommendations indicated by the Advisory Committee on Immunization Practices (ACIP), which is a federal advisory committee that provides expert external advice and guidance to the CDC, as the basis of evidence and secondhand indication of evidence based on published, peer-reviewed studies.
    - The developer explained that the evidence review did not address the overall quality of the body of evidence related to this measure, the overall consistency of results across studies nor was any grade provided for the quality of the body of evidence.

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:

- The developer provided updated information on ACIP's clinical practice guideline recommendations, which provides annual recommendations for the use of influenza vaccines for the prevention and control of influenza in the United States.
    - The developer noted one of ACIP's 2021 statements on its published recommendations, i.e., routine annual influenza vaccination for all persons aged ≥6 months who do not have contraindications has been recommended by CDC and the ACIP) since 2010.
- The developer also cited studies updated studies that indicate that vaccination provides important protection from influenza illness and its potential complications. Specifically, the developer cites ACIP's reporting on six influenza seasons beginning 2010-11 through 2015-16; the findings indicated that during this period influenza vaccination prevented an estimated 1.6–6.7 million illnesses, 790,000–3.1 million outpatient medical visits, 39,000–87,000 hospitalizations, and 3,000–10,000 respiratory and circulatory deaths each season in the United States.
- The developer made additional reference to influenza activity, during the severe 2017–18 season, and notes that vaccination prevented an estimated 7.1 million illnesses, 3.7 million medical visits, 109,000 hospitalizations, and 8,000 deaths, despite an overall estimated vaccine effectiveness of 38 percent.

Exception to evidence

- N/A

Questions for the Committee:

- *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?*
- *What is the relationship of this measure to patient outcomes?*
- *How strong is the evidence for this relationship?*
- *Is the evidence directly applicable to the process of care being measured?*

Guidance from the Evidence Algorithm

Not an outcome measure (Box 1) -> Process measure based on systematic review (Box 3) -> QQC not presented (Box 4)

*Without QQC from systematic review, moderate is the highest potential rating.

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

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

Maintenance measures – increased emphasis on gap and variation

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

- The developer provided summary statistics for the year 2020 on performance data (representative of 4,032 reporting clinicians) at the clinician level. The developer posited that there is evident variation in performance that warrants continuation of the measure

- Mean: 69.81
- 10th percentile: 18
- 25th percentile: 48
- 50th percentile: 82
- 75th percentile: 99
- 90th percentile: 10

- The CDC-sourced data that the developer provides also indicates that only half of all eligible individuals receive an influenza vaccination.

Disparities

- The developer explained that while this measure is included in federal reporting programs, those programs have not yet made disparities data available for analysis.
- In the absence of disparities data, the developer provided a summary of data from the literature that addresses disparities in care on the specific focus of measurement.
    - The developer cited data from the CDC analyses that reflects differences in in flu vaccination by individual characteristics including age, gender and race/ethnicity and suggests regional differences in rates of vaccination.
        - Age: The developer reported that adults aged 18 years and older had lower rates of vaccination than children 6 months – 17 years
        - Gender: The developer reported no differences in flu vaccination coverage between male and female children 6 months through 17 years
            - The developer noted that data indicated the flu vaccination was generally higher among females than males for adults 18 years and older
        - Race/Ethnicity: The developer reported that persons classified as white had higher flu coverage than persons classified as Black, Hispanic, and people of other races.
        - Region: The developer reported on between-state variability in flu vaccination coverage among adults across reporting states.
            - The data that the developer furnished showed a range of 41.5 percent in Wyoming to 64.0 percent in Rhode Island.

Questions for the Committee:

- *Is there a gap in care that warrants a national performance measure?*
- *If limited disparities information is provided, are you aware of evidence that disparities exist in this area of healthcare?*

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

Committee Pre-evaluation Comments:

1a. Evidence

- Indirect measure- the data does not evaluate hospitalization rates for influenza in population or among those vaccinated.
- The evidence is documented and updated. The evidence is directionally the same. I do not believe that evidence needs to be reconsidered.

- Rating: Low, expect there is available QQC & graded evidence by now, systematic review evidence w/o QQC or graded evidence, would expect there is is patient outcome correlations by now-seems like we should continue see emerging evidence over the years that would inform us of measure refinement needs.
- Direct.
- Evidence is direct. Process relates to desired outcomes. Not aware of any new evidence.
- The patient self-report for this measure is a process measure and appears to apply directly to the measure.
- There was strong evidence, based on systematic reviews and practice guidelines, in support of annual influenza vaccines for individuals over 6 months of age, and the evidence base has been strengthened since this measure was last approved. It should be noted, however, that these guidelines recommend that everyone be vaccinated, not just those seen for a visit between October 1 and March 31, as stated in the Denominator statement. I understand the practical reasons for this definition, but it does shift the accountability from the plan to the members to ensure that they seek care during the flu season.
- Reasonable evidence provided.
- Evidence derived from a systematic review updated.

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

- Disparities data missing, though evidence exists they exist. It is unclear how continuing this measure "as is", without any changes, mitigates this gap. Clinician level data also not sufficient to warrant continuation. Clinicians in certain settings cannot control their panel composition, and disparities may amplify differences.
- The absence of disparities data in federal reporting requirements is unfortunate. Some disparity data is included and does demonstrate disparity. Payor status and geographic location by urban versus rural could also be helpful.
- Rating: Moderate, performance variations might be explained by risk adjustment/refinements.
- Regional, age, race disparities.
- Yes. Gap in care is substantial and warrants national performance measure. Limited recent data on disparities, but appears there are disparities regionally, by gender and white vs non-white populations, supporting continued need for the measure.
- A performance gap on the measure was not noted. No statistics on subgroups or disparities was provided.
- Overall performance is far from universal coverage, as recommended. In addition, the develop cites evidence of disparities by age, sex, race, and geographical region.
- Significant evidence of gaps.
- Data is collected at the clinician level for annual flu vaccination; of 4032 reporting clinicians, mean and percentile performance data reveal gap in vaccination for all persons 6 years and older.

Criteria 2: Scientific Acceptability of Measure Properties

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

Evaluators: Staff

2a. Reliability: Specifications and Testing

For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

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 – less emphasis if no new testing data provided.

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.

Specifications:

- Measure specifications are clear and precise.
- The developer attests that the measure specifications have not changed since the last submission.

Reliability Testing:

- Reliability testing conducted at the accountable-entity level
    - Within the January 1 – December 31, 2018, reporting period, the developer assessed a total of 7,789 practices (85 percent majority single practitioners) to evaluate variance between providers and variance within providers.
    - The developer conducted signal-to-noise reliability using the Beta-binomial model to assess how well one can confidently distinguish the performance of one reporting entity from another.
        - Mean: 0.996;
        - Min: 0.898;
        - 25th percentile: 0.995;
        - 50th percentile: 0.998;
        - 75th percentile: 0.999;
        - 90th percentile: 1.000;
        - Standard Deviation: 0.007
- The developer notes that the results demonstrate reliability.

Questions for the Committee regarding reliability:

- *Do you have any concerns that the measure cannot be consistently implemented (i.e., are measure specifications adequate)?*

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

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

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

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.

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

Validity Testing

- Empirical validity testing was conducted at the accountable entity/score level:
    - The developer reports Pearson Product Moment Correlation Coefficients to estimate the strength and direction of association between the influenza measure and another measure which assesses immunization services: Pneumococcal Vaccination Status for Older Adults.
    - The developer reports a positive correlation coefficient of 0.8111 and a p-value less than 0.0001.
- Based on the correlation coefficient of 0.8111, the developer suggests a positive and high association with the pneumococcal vaccination measure and a strong likelihood that those who perform well on the pneumococcal measure will perform well on the influenza measure.

Exclusions

- The measure does not use exclusions.

Risk-Adjustment

- The developer does not use risk-adjustment or stratification.

Meaningful Differences

- The developer calculated an inter-quartile range (IQR) between the 25th and 75th percentile for each measure and generated a testing statistic p-value of less than 0.0001.
- The developer asserts that the difference in performance between providers is statistically significant and meaningful.

Missing Data

- The developer did not report on the extent and distribution of missing data.

Comparability

- The measure only uses one set of specifications for this measure.

Questions for the Committee regarding validity:

- *Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?*

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

Committee Pre-evaluation Comments:

2a1. Reliability – Specifications

- No issues.
- The preliminary rating for reliability is high. Reliability testing was conducted at the accountable-entity level.
- Rating: high.

- No concerns.
- No problems noted/no concerns re reliability specs.
- Codes for encounters, either CPS or HCPCS are provided. Codes for Influenza Immunization Administered, Not Administered with documented and undocumented reasons are also provided.
- Reliability testing conducted at the accountable-entity level, appropriately using the beta-binomial method with good results.
- All clear.
- Signal-to-noise reliability demonstrated high level of reliability; mean 0.996, Min 0.898.

2a. Reliability – Testing

- No issues.
- No.
- No.
- No concern.
- No concerns.
- Guidance from CDC ACIP is provided about the need for the measure. No information on reliability was included.
- No concerns.
- Very high.
- No.

2b. Validity – Testing

- The use of pneumococcal as a proxy needs to be contextualized for bias it may bring- panels that have patients eligible for pneumococcal may be older, and a different demographic/risk than those that are younger, so we must ask ourselves if this is an appropriate measure for validity.
- No.
- Rating: moderate, suspect this established measure needs refinement & risk adjustment to drive improvement.
- No concern.
- No concerns.
- Guidance from CDC ACIP is provided about the need for the measure. No information on validity was included.
- No concerns.
- No concerns.
- Empirical validity testing was conducted at the accountable entity/score level: Pearson Product Moment Correlation Coefficients; developer reports a positive correlation coefficient of 0.8111 and a p-value less than 0.0001.

2b2-2b3. Threats to validity (Exclusions, Risk Adjustment)

- —
- Risk adjustment is not used.
- See comments above.
- None.
- No concerns.
- Exclusions appear to be appropriate and specific to the measure. No social risk factors were offered for risk adjustment.

- No concerns.
- No concerns.
- The developer does not use risk-adjustment or stratification.

2b4-2b7. Threats to validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)

- See above.
- No.
- Comments above.
- No concern.
- No concerns.
- No information provided.
- No concerns.
- No concerns.
- The measure does not use exclusions.; The developer did not report on the extent and distribution of missing data.

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.

- The developer indicates that the data elements are generated or collected and used by healthcare personnel during the provision of care and are coded by someone other than person obtaining original information.
- The developer states that the measure relies on data elements that are defined in a combination of electronic sources and cites no difficulties in the collection of data.

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:

3. Feasibility

- Consideration that vaccinations given in other locations such as pharmacies may not be easily accessible in EMR, dependent on state-registries.
- Data elements are collected. Since it is claims data, it may not represent uninsured individuals not receiving vaccination which is likely one of the most significant barriers.
- Rating: high.
- No concern.

- No concerns.
- The data elements appear to a part of routinely collected data during a patient encounter.
- Measure currently in wide use.
- No concerns.
- The developer states that the measure relies on data elements that are defined in a combination of electronic sources and cites no difficulties in the collection of data. High level of feasibility.

Criterion 4: Use and Usability

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 evaluates 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
Planned use in an accountability program?	☐	Yes	☐	No	☒	NA

Accountability program details

- The developer reports that the measure is in use in the Centers for Medicare and Medicaid Services (CMS) Quality Payment Program (QPP).
- The developer notes that CMS publishes measure performance results and scores, which are publicly available and identifiable by clinician and group, on its Physician Compare website annually.

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

- Developer did not report any feedback received from those being measured nor did the developer report on a mechanism for feedback by those being measured.

***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 evaluates 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

- Developer stated that trending performance data is not available for this measure yet.
- Developer added that in 2017, MIPS replaced the Physician Quality Reporting System (PQRS) which ended in 2016, and clinician-level MIPS performance results from 2017 through 2019 are not available. The developer further added that the average MIPS performance rate in 2020 was 69.8% and the most recent year of available reporting data for PQRS is 2014. Adding to this, the developer stated that the average performance rate in 2014 was 46.3% and there has been an improvement in performance between 2014 and 2020.

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

- Developer did not report any unexpected findings during implementation.

Potential harms

- Developer did not report any potential harms.

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

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

RATIONALE: Developer does not provide any performance improvement data nor any rationale for the absence of any implementation findings.

- **Updated rationale**: NQF staff changed its preliminary rating from insufficient to moderate following the developer's review of NQF staff's preliminary analysis. The developer provided feedback concerning usability; the information that the developer identified was located in the performance gap section, not under the usability section.

Committee Pre-evaluation Comments:

4a. Use

- Scant improvement data - unclear what feedback was integrated.
- This is an easily understood measure. Those being measured have been given performance results.

- Rating: moderate.
- No concerns.
- No concerns. Used as a publicly available outcome by CMS on Care Compare.
- Information about this metric are available from the CDC and are included in the Healthy People objectives as well as the Veterans Administration and Indian Health Service. Flu vaccinations for adults (most recent data is 2020), seniors (most recent data is 2018) are available on the Internet through the HEDIS users group.
- No comments.
- No concerns.
- Used for public reporting and accountability programs. Measure is in use in the Centers for Medicare and Medicaid Services (CMS) Quality Payment Program (QPP).

4a. Usability

- As with any vaccine metric and reality of vaccine hesitancy in historically marginalized communities, must be careful to consider impact of incentivizing this metric against overall impact on patients' comfort accessing/continuing care. There is no good way to measure this that I'm aware of, but guardrails can be added to mitigate risk of coercion or "cherry picking" patients who would engage.
- No identified unintended consequences.
- Rating: moderate.
- No concern.
- No concerns.
- The results of this measure could be used to encourage improvement in performance. Benefits appear to outweigh harms.
- No comments.
- No concerns.
- Developer did not report any potential harms.

Criterion 5: Related and Competing Measures

Related measures

- NQF# 0038 Childhood Immunization Status (CIS)
- NQF #0226 Influenza Immunization in the ESRD Population (Facility Level)
- NQF #0431 Influenza Vaccination Coverage among Healthcare Personnel
- NQF #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short-Stay)
- NQF #1659 Influenza Immunization
- NQF #3484 Prenatal Immunization Status
- NQF #3620 Adult Immunization Status

Harmonization

- Developer did not report on any harmonization efforts. Developer reported on the distinctions and differing target populations of the measures which it identified as related to NQF measure #0041 Preventive Care and Screening: Influenza Immunization.

Committee Pre-evaluation Comments:

5: Related and Competing Measures

- Distinct from others
- There are related measures regarding vaccination for healthcare workers. They are being harmonized.
- Group measures with distinction agreed.
- None noted.
- There are multiple related measures specific to discrete populations and/or institutional providers. No concerns about harmonization.
- None noted.
- All flu vaccine measures should have the same inclusion and exclusion criteria, consistent with NQF report of a few years ago.
- Nothing reported re harmonization.
- Many related measures. No harmonization efforts.

Public and NQF Member Comments (Submitted as of June 15, 2022)

Member Expression of Support

- Of the two NQF members who have submitted an expression of support, two expressed "support" and none expressed "do not support" for the measure.

Comments

Comment 1 by: Fern McCree, NCQA; Submitted by Bob Rehm, National Committee for Quality Assurance

In 2017, MIPS replaced the Physician Quality Reporting System (PQRS) which ended in 2016. Clinician-level MIPS performance results from 2017 through 2019 are not available. The average MIPS performance rate in 2020 was 69.8%. The most recent year of available reporting data for PQRS is 2014. The average performance rate in 2014 was 46.3%. There has been an improvement in performance between 2014 and 2020.

Comment 2 by: Submitted by Koryn Rubin, American Medical Association

The American Medical Association (AMA) appreciates the opportunity to comment on this measure. We are writing to request clarification on several items in the measure submission form. On review of the measure specifications, the developer notes that it includes a denominator exception for medical or patient reasons (see sp.13 as an example) and sp.22 outlines how these exceptions should be removed from the denominator. However, sp.16, which describes denominator exclusions, is marked "None" nor did the developer provide any analysis on the frequency of exceptions in the measure testing section (see 2b.15 through 2b.18). We believe that these inconsistencies must be addressed, and the developer must ensure that what is endorsed is aligned with the version of the measure currently in the Merit-Based Incentive Payment System (MIPS). We also request clarification on the use and usability of the measure. On our review, it does not appear that this section was updated since stewardship of the measure was transitioned from the PCPI to the National Committee for Quality Assurance. The AMA requests that these discrepancies be addressed prior to continued endorsement of this measure. We appreciate the Committee's consideration of our comments.

RELIABILITY: SPECIFICATIONS

1. **Have measure specifications changed since the last review? ☐ Yes ☒ No**
2. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented? ☒ Yes ☐ No**
3. **Briefly summarize any changes to the measure specifications and/or concerns about the measure specifications.**
   - Measure specifications are clear and precise and the developer attests that the measure specifications have not changed since the last submission.

RELIABILITY: TESTING

4. **Did the developer conduct new reliability testing? ☒ Yes ☐ No**

   4a. If no, summarize the Standing Committee's previous feedback:
   - N/A

   4b. If yes, describe any differences between the new and old testing and summarize any relevant Standing Committee's feedback from the previous review:
   - During the previous review, this measure demonstrated 0.80 reliability when evaluated at the minimum level of quality reporting events and 0.99 reliability when evaluated at the average number of quality events. The updated review indicates a minimum reliability scoring of 0.92 and a mean reliability score of 0.997.
   - In the previous review that occurred in year 2016, the developer provides a breakdown of data on the physicians reporting on the measure. In the new testing information, the developer does not provide descriptive data at the clinician level and explains that CMS does not report descriptive data at the clinician level.
5. **Reliability testing level: ☒ Accountable-Entity Level ☐ Patient/Encounter Level ☐ Neither**
6. **Reliability testing was conducted with the data source and level of analysis indicated for this measure:**

   ☒ Yes ☐ No
7. If accountable-entity level and/or patient/encounter level reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ Yes ☐ No
8. **Assess the method(s) used for reliability testing:**
   - Reliability testing conducted at the accountable-entity level. The developer assessed a total of 7,789 practices (85 percent majority single practitioners) to evaluate variance between providers and variance within providers. The developer calculated the signal-to-noise reliability ratio using the Beta-binomial model to assess the distribution of ratios and took the following steps to estimate the reliability for each provider and summarize the distribution of these estimates.
   - The developer identified the reliability score for each reporting entity. The developer averaged those reliability scores across all reporting entities. The developer produced a point estimate, a mean score, of the signal to noise reliability estimates. The developer provided the distribution of the provider level signal to noise reliability estimates.
9. **Assess the results of reliability testing**
   - The developer suggests high reliability for each practice and a strong ability for the measure to differentiate between reporting entities based on the following performance summary statistics and

attests that the reliability testing methodology is appropriate for the indicated level of analysis, and the results demonstrate strength in performance differentiation.

- Mean: 0.996;
- Min: 0.898;
- 25th percentile: 0.995;
- 50th percentile: 0.998;
- 75th percentile: 0.999;
- 90th percentile: 1.000;
- Standard Deviation: 0.007

10. 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.

    ☒ **Yes** ☐ **No** ☐ **Not applicable**

11. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

    ☐ **Yes** ☐ **No** ☒ **Not applicable** (patient/encounter level testing was not performed)

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

    ☒ **High** (NOTE: Can be HIGH only if accountable-entity level testing has been conducted)

    ☐ **Moderate** (NOTE: Moderate is the highest eligible rating if accountable-entity 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)

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

    - The developer demonstrates reliability testing at the accountable-entity level. Furthermore, the reliability testing methodology is deemed appropriate for the indicated level of analysis, and the results demonstrate strength in performance differentiation.

VALIDITY: TESTING

14. **Did the developer conduct new validity testing? ☒ Yes ☐ No**

    14a. If no, summarize the Standing Committee's previous feedback:

    - N/A

    14b. If yes, describe any differences between the new and old testing and summarize any relevant Standing Committee's feedback from the previous review:

    - In the previous review, the developer employed face validity to qualitatively assess the content validity of the measure. Furthermore, the developer assessed the extent of agreement, among a portion of their expert panel, concerning a corresponding validity statement. Referencing a scale of 1-5, where 1= Strongly Disagree; 2= Disagree 3= Neither Agree nor Disagree; 4=Agree; 5= Strongly, the developer's assessment indicated that 89% of respondents (N = 9; Mean rating = 4.1) either agree or strongly agree that this measure can accurately distinguish good and poor quality.
    - In the updated review, developer uses empirical validity testing; specifically the developer performs the Pearson correlation coefficient to determine whether the influenza measure results correlate with another immunization measure: Pneumococcal Vaccination Status for Older Adults. The developer reports positive and strong associations between the influenza measure and the pneumococcal vaccination measure ($r = 0.727$ $p < 0.001$).

15. **Validity testing level (check all that apply):**

    ☒ **Accountable-Entity Level** ☐ **Patient or Encounter-Level** ☐ **Both**

    NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

16. **If patient/encounter level validity testing was provided, was the method described and appropriate for assessing the accuracy of ALL critical data elements? NOTE:** Data element validation from the literature is acceptable.

    ☒ **Yes**

    ☐ **No**

    ☐ **Not applicable** (patient/encounter level testing was not performed)

17. **Method of establishing validity at the accountable-entity level:**

    ☐ **Face validity**

    ☒ **Empirical validity testing at the accountable-entity level**

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

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

    ☒ **Yes**

    ☐ **No**

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

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

    - Empirical validity testing was conducted at the accountable entity/score level. The developer reports Pearson Product Moment Correlation Coefficients to estimate the strength and direction of association between the influenza measure and another measure which assesses immunization services: Pneumococcal Vaccination Status for Older Adults.
    - The developer explains why the correlation of measures validates the measure and demonstrates the direction and strength of the hypothesized association. The developer also provides specific statistical tests, results, and interpretation of the analysis used; however, developer could provide additional interpretation of more of its resulting data points.

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

    - The developer reports a positive correlation coefficient of 0.8111 and a p-value less than 0.0001 as indicated by the Pearson Product Moment Correlation Coefficient test.
    - Based on the correlation coefficient of 0.8111, the developer suggests a positive and high association with the pneumococcal vaccination measure and a strong likelihood that those who perform well on the pneumococcal measure will perform well on the influenza measure.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

    - Developer notes no measure exclusions.

22. **Risk Adjustment**

    22a. **Risk-adjustment method**

    ☒ None (only answer Question 20b and 20e) ☐ Statistical model ☐ Stratification

    ☐ Other method assessing risk factors (please specify)

    22b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

22c. **Social risk adjustment:**

22c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

22c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

22c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

22d.**Risk adjustment summary:**

22d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

22d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

22d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

22d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

22d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

22e**. Assess the risk-adjustment approach**

- N/A

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

    For cost/resource use measures, does this measure identify meaningful differences about cost and resource use between the measured entities?

    - There are no concerns regarding the ability that the measure holds to identify meaningful differences in performance.

24. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**

    - The measure uses one set of specifications for this measure.

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

    - There are no concerns regarding missing data.

ADDITIONAL RECOMMENDATIONS

26. **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.**

    - There are no additional concerns.

Criteria 1: 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

---

1ma.01. Indicate whether there is new evidence about the measure since the most recent maintenance evaluation. If yes, please briefly summarize the new evidence, and ensure you have updated entries in the Evidence section as needed.

[Response Begins]

No

[Response Ends]

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:

2021 Submission:

Updated evidence information here.

2018 Submission:

Evidence from the previous submission here.

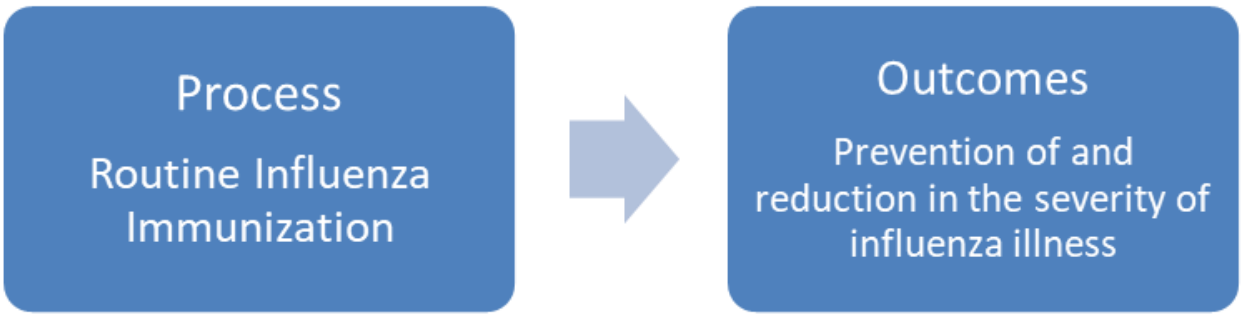
1a. Evidence

1a.01. Provide a logic model.

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.

[Response Begins]

Administering Influenza immunization leads to prevention of and reduction in the severity of influenza illness.



[Response Ends]

1a.02. Select the type of source for 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.

[Response Begins]

Clinical Practice Guideline recommendation (with evidence review)

[Response Ends]

If the evidence is not based on a systematic review, skip to the end of the section and do not complete the repeatable question group below. If you wish to include more than one systematic review, add additional tables by clicking "Add" after the final question in the group.

Evidence - Systematic Reviews Table (Repeatable)

Group 1 - Evidence - Systematic Reviews Table

1a.03. Provide the title, author, date, citation (including page number) and URL for the systematic review.

[Response Begins]

Grohskopf LA, Alyanak E, Ferdinands JM, et al. Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices, United States, 2021–22 Influenza Season. MMWR Recomm Rep 2021;70(No. RR-5):1–28. DOI: http://dx.doi.org/10.15585/mmwr.rr7005a1external icon.

[Response Ends]

1a.04. Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the systematic review.

[Response Begins]

Routine annual influenza vaccination for all persons aged ≥6 months who do not have contraindications has been recommended by CDC and the Advisory Committee on Immunization Practices (ACIP) since 2010 (CDC/Advisory Committee on Immunization Practices [ACIP], 2021).

ACIP provides annual recommendations for the use of influenza vaccines for the prevention and control of influenza in the United States. The ACIP Influenza Work Group meets by teleconference once to twice per month throughout the year. Work group membership includes several voting members of ACIP, representatives of ACIP liaison organizations, and consultants. Discussions include topics such as influenza surveillance, vaccine effectiveness and safety, vaccination coverage, program feasibility, cost-effectiveness, and vaccine supply. Presentations are requested from invited experts, and published and unpublished data are discussed. (CDC/Advisory Committee on Immunization Practices [ACIP], 2021).

The Background Document that supplements this report is updated periodically to reflect recent additions to the literature related to recommendations made in previous seasons and minor changes in guidance for the use of influenza vaccines (e.g., guidance for timing of vaccination and other programmatic issues, guidance for dosage in specific populations, guidance for selection of vaccines for specific populations that are already recommended for vaccination, and changes that reflect use that is consistent with indications and prescribing information licensed by the Food and Drug Administration [FDA]). The summary included in the Background

Document for such topics is not a systematic review; it is intended to provide an overview of current literature, with updated articles being identified primarily through a broad search for English-language articles on influenza and influenza vaccines. (CDC/Advisory Committee on Immunization Practices [ACIP], 2021).

Vaccination provides important protection from influenza illness and its potential complications. The effectiveness of influenza vaccination varies depending on several factors, such as the age and health of the recipient; the type of vaccine administered; the types, subtypes (for influenza A), and lineages (for influenza B) of circulating influenza viruses; and the degree of similarity between circulating viruses and those included in the vaccine.

CDC. How flu vaccine effectiveness and efficacy are measured: questions and answers. Atlanta, GA: US Department of Health and Human Services, CDC; 2016. https://www.cdc.gov/flu/professionals/vaccination/effectivenessqa.htm

2016 Submission

The Advisory Committee on Immunization Practices (ACIP) revised its influenza recommendations in 2010 to include a recommendation that annual vaccination be administered to all persons aged ≥6 months. This recommendation is current and has not changed as of 2016.

There is increased evidence that influenza has substantial adverse impacts in all age groups and an expectation that a simplified age-based influenza vaccine recommendation for all age groups will improve vaccine coverage levels. Published, peer-reviewed studies are the primary source of data used by the ACIP in making recommendations for the prevention and control of influenza, but unpublished data that are relevant to issues under discussion also are considered. Among studies discussed or cited, those of greatest scientific quality and those that measure influenza-specific outcomes are the most influential (CDC, 2010).

Centers for Disease Control and Prevention. Prevention and control of influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP), 2010. MMWR 2010;59(No. RR-8):1-62.

[Response Ends]

1a.05. Provide the grade assigned to the evidence associated with the recommendation, and include the definition of the grade.

[Response Begins]

2022 Submission

In general, systematic review and evaluation of evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach is performed for new recommendations or substantial changes in the current recommendations (e.g., expansion of the recommendation for influenza vaccination to new populations not previously recommended for vaccination or potential preferential recommendations for specific vaccines).

Although the body of evidence was not graded, the guidelines were developed by CDC's Advisory Committee on Immunization Practices (ACIP) who provides annual recommendations for the prevention and control of influenza. The ACIP Influenza Work Group meets by teleconference once to twice per month throughout the year. Work group membership includes several voting members of ACIP, representatives of ACIP liaison organizations, and consultants. Discussions include topics such as influenza surveillance, vaccine effectiveness and safety, vaccination coverage, program feasibility, cost-effectiveness, and vaccine supply. Presentations are requested from invited experts, and published and unpublished data are discussed. Work Group members also request periodic updates on vaccine and antiviral production, supply, safety, and efficacy from vaccinologists, epidemiologists, and manufacturers. State and local vaccination program representatives are consulted. CDC's Influenza Division (available at http://www.cdc.gov/flu) provides influenza surveillance and antiviral resistance data. The Vaccines and Related Biological Products Advisory Committee provides advice on vaccine strain

selection to the Food and Drug Administration (FDA), which selects the viral strains to be used in the annual trivalent influenza vaccines.

*The Work Group composition was as follows:

Chair: José R. Romero, MD, University of Arkansas for Medical Sciences, Little Rock, Arkansas.

Executive Secretary: Amanda Cohn, MD, National Center for Immunization and Respiratory Diseases, CDC Atlanta, Georgia.

Members: Kevin A. Ault, MD, University of Kansas Medical Center, Kansas City, Kansas; Lynn Bahta, MPH, Minnesota Department of Health, St. Paul, Minnesota; Beth P. Bell, MD, University of Washington, Seattle, Washington; Henry Bernstein, DO, Zucker School of Medicine at Hofstra/Northwell Cohen Children's Medical Center, New Hyde Park, New York; Wilbur H. Chen, MD, University of Maryland School of Medicine, Baltimore, Maryland; Matthew F. Daley, MD, Kaiser Permanente Colorado, Aurora, Colorado; Sharon E. Frey, MD, Saint Louis University Medical School, St. Louis, Missouri; Camille Nelson Kotton, MD, Harvard Medical School, Boston, Massachusetts; Grace M. Lee, MD, Lucile Packard Children's Hospital, Stanford University School of Medicine, Stanford, California; Sarah S. Long, MD, Drexel University College of Medicine, Philadelphia, Pennsylvania; Veronica V. McNally, JD, Franny Strong Foundation, West Bloomfield, Michigan; Katherine A. Poehling, MD, Wake Forest School of Medicine, Winston-Salem, North Carolina; Pablo J. Sánchez, MD, The Research Institute at Nationwide Children's Hospital, Columbus, Ohio; Helen Keipp Talbot, MD, Vanderbilt University, Nashville, Tennessee.

Ex Officio Members: Centers for Medicare and Medicaid Services, Mary Beth Hance, Baltimore, Maryland; Food and Drug Administration, Doran Fink, MD, PhD, Silver Spring, Maryland; Health Resources and Services Administration, Mary Rubin, MD, Rockville, Maryland; Indian Health Service, Thomas Weiser, MD, Portland, Oregon; Office of Infectious Disease and HIV/AIDS Policy, David Kim, MD, Washington, DC; National Institutes of Health, John Beigel, MD, Bethesda, Maryland.

2016 Submission

Although the body of evidence was not graded, the guidelines were developed by CDC's Advisory Committee on Immunization Practices (ACIP) who provides annual recommendations for the prevention and control of influenza. The ACIP Influenza Work Group (the Work Group)* meets every 2–4 weeks throughout the year to discuss newly published studies, review current guidelines, and consider revisions to the recommendations. As the Work Group reviews the annual recommendations for consideration by the full ACIP, its members discuss a variety of issues, including the burden of influenza illness; vaccine effectiveness, vaccine safety, and coverage in groups recommended for vaccination; feasibility; cost-effectiveness; and anticipated vaccine supply. Work Group members also request periodic updates on vaccine and antiviral production, supply, safety, and efficacy from vaccinologists, epidemiologists, and manufacturers. State and local vaccination program representatives are consulted. CDC's Influenza Division (available at http://www.cdc.gov/flu) provides influenza surveillance and antiviral resistance data. The Vaccines and Related Biological Products Advisory Committee provides advice on vaccine strain selection to the Food and Drug Administration (FDA), which selects the viral strains to be used in the annual trivalent influenza vaccines.

*The Work Group composition was as follows:

Chair: Kathleen Neuzil, MD, Seattle, Washington.

Members: Terry Adirim, MD, District of Columbia; William Atkinson, MD, Atlanta, Georgia; Carol Baker, MD, Houston, Texas; Beth Bell, MD, Atlanta, Georgia, Nancy Bennett, MD, Rochester, New York; Henry Bernstein, DO, Lebanon, New Hampshire; Joseph Bresee, MD, Atlanta, Georgia; Carolyn Bridges, MD, Atlanta, Georgia; Karen Broder, MD, Atlanta, Georgia; Doug Campos-Outcalt, MD, Phoenix, Arizona; Fred Cassels, MD, Rockville, Maryland; Lance Chilton, MD, Albuquerque, New Mexico; David Cho, MD, District of Columbia; Nancy Cox, PhD, Atlanta, Georgia; Therese Cvetkovich, MD, Rockville, Maryland; Sandra Dos Santos Chaves, MD, Atlanta, Georgia; Jeff Duchin, MD, Seattle, Washington; Janet Englund, MD, Seattle, Washington; Anthony Fiore, MD, Atlanta, Georgia; Sandra Fryhofer, MD, Atlanta, Georgia; Stanley Gall, MD, Louisville, Kentucky; Paul Gargiullo, PhD, Atlanta, Georgia; Steven Gordon, MD, Cleveland, Ohio; Wayne Hachey, DO, Falls Church, Virginia; John

Iskander, MD, Atlanta GA; Wendy Keitel, MD, Houston, Texas; Elyse Olshen Kharbanda, MD, New York, NY; David Lakey, MD, Austin, Texas; Susan Lett, MD, Boston, Massachusetts; Tamara Lewis, MD, Salt Lake City, Utah; Cynthia Nolletti, MD, Rockville, Maryland; Gregory Poland, MD, Rochester, Minnesota; William Schaffner, MD, Nashville, Tennessee; Robert Schechter, MD, Sacramento, California; Kenneth Schmader, MD, Durham, NC; David Shay, MD, Atlanta, Georgia; Nadine Sicard, MD, Ottawa, Canada; Danuta Skowronski, MD, Vancouver, British Columbia, Canada; Patricia Stinchfield, St. Paul, Minnesota; Ray Strikas, MD, District of Columbia; Litjen Tan, PhD, Chicago, Illinois; Mary Vernon-Smiley, MD Atlanta, Georgia; Timothy Uyeki, MD, Atlanta, Georgia; Amanda Zongrone, Atlanta, Georgia.

[Response Ends]

1a.06. Provide all other grades and definitions from the evidence grading system.

[Response Begins]

N/A

[Response Ends]

1a.07. Provide the grade assigned to the recommendation, with definition of the grade.

[Response Begins]

N/A

[Response Ends]

1a.08. Provide all other grades and definitions from the recommendation grading system.

[Response Begins]

N/A

[Response Ends]

1a.09. Detail the quantity (how many studies) and quality (the type of studies) of the evidence.

[Response Begins]

2022 Submission

The description of the evidence within the guideline, did not address the overall quantity of studies in the body of evidence. However, over 600 studies are cited in the reference section.

2016 Submission

The description of the evidence review within the guideline, did not address the overall quantity of studies in the body of evidence. However, close to 500 studies are cited in the reference section.

[Response Ends]

1a.10. Provide the estimates of benefit, and consistency across studies.

[Response Begins]

2022 Submission

Benefit and Consistency

CDC conducts studies to measure the benefits of seasonal flu vaccination each flu season to help determine how well flu vaccines are working. These vaccine effectiveness (VE) studies regularly assess the value of flu

vaccination as a public health intervention. Study results of vaccine effectiveness can vary based on the study design, the outcome(s) measured, the population studied and the season in which the flu vaccine was studied.

Vaccination provides important protection from influenza illness and its potential complications. During the six influenza seasons from 2010–11 through 2015–16, influenza vaccination prevented an estimated 1.6–6.7 million illnesses, 790,000–3.1 million outpatient medical visits, 39,000–87,000 hospitalizations, and 3,000–10,000 respiratory and circulatory deaths each season in the United States. During the severe 2017–18 season, notable for an unusually long duration of widespread high influenza activity throughout the United States and higher rates of outpatient visits and hospitalizations compared with recent seasons, vaccination prevented an estimated 7.1 million illnesses, 3.7 million medical visits, 109,000 hospitalizations, and 8,000 deaths, despite an overall estimated vaccine effectiveness of 38% (62% against influenza A[H1N1]pdm09 viruses, 22% against influenza A[H3N2] viruses, and 50% against influenza B viruses).

Grohskopf LA, Alyanak E, Ferdinands JM, et al. Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices, United States, 2021–22 Influenza Season. MMWR Recomm Rep 2021;70(No. RR-5):1–28. DOI: http://dx.doi.org/10.15585/mmwr.rr7005a1external icon.

Benefit

2016 Submission

Annual influenza vaccination is a safe and effective preventive health action with potential benefit in all age groups.

...[E]vidence from clinical trials suggests that protection against viruses that are similar antigenically to those contained in the vaccine extends for at least 6–8 months.

Influenza viruses cause disease among persons in all age groups. Rates of infection are highest among children, but the risks for complications, hospitalizations, and deaths from seasonal influenza are higher among adults aged >=65 years, children aged <5 years, and persons of any age who have medical conditions that place them at increased risk for complications from influenza.

Centers for Disease Control and Prevention. Prevention and control of influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP), 2010. MMWR 2010;59(No. RR-8):1-62.

Consistency of Results

2016 Submission

Although there is no explicit statement regarding the overall consistency of results across studies in the guideline, the recent ACIP influenza immunization recommendations represent an expansion of the previous recommendations for annual vaccination of all adults aged 19—49 years and "is supported by evidence that annual influenza vaccination is a safe and effective preventive health action with potential benefit in all age groups."

Centers for Disease Control and Prevention. Prevention and control of influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP), 2010. MMWR 2010;59(No. RR-8):1-62.

[Response Ends]

1a.11. Indicate what, if any, harms were identified in the study.

[Response Begins]

N/A

[Response Ends]

1a.12. Identify any new studies conducted since the systematic review, and indicate whether the new studies change the conclusions from the systematic review.

[Response Begins]

N/A

[Response Ends]

1a.13. If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, describe the evidence on which you are basing the performance measure.

[Response Begins]

N/A

[Response Ends]

1a.14. Briefly synthesize the evidence that supports the measure.

[Response Begins]

N/A

[Response Ends]

1a.15. Detail the process used to identify the evidence.

[Response Begins]

N/A

[Response Ends]

1a.16. Provide the citation(s) for the evidence.

[Response Begins]

N/A

[Response Ends]

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

1b.01. Briefly explain the rationale for this measure.

Explain how the measure will improve the quality of care, and list the benefits or improvements in quality envisioned by use of this measure.

[Response Begins]

Influenza may lead to serious complications including hospitalization or death (1). Influenza vaccination is the most effective protection against influenza virus infection (1). However, data indicate that less than half of all eligible individuals receive an influenza vaccination (2). This measure promotes annual influenza vaccination for all persons aged >= 6 months.

1. Seasonal Influenza: Flu Basics. Centers for Disease Control and Prevention Web site. http://www.cdc.gov/flu/about/disease/index.htm. Updated May 4, 2016. Accessed June 23, 2016.
2. Flu vaccination coverage: United States, 2014-15 Influenza Season. Centers for Disease Control and Prevention Web site. http://www.cdc.gov/flu/fluvaxview/coverage-1415estimates.htm. Updated September 17, 2015. Accessed June 23, 2016.

[Response Ends]

1b.02. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.

Include mean, std dev, min, max, interquartile range, and scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]

2022 Submission

The following data were extracted from the Merit-Based Incentive Payment System (MIPS) program reflecting claims and registry data for immunizations provided during the 2020 performance year. For 2020, 4032 clinicians reported this measure. Performance data are summarized at the clinician level and described by mean, 10th, 25th, 50th, 75th and 90th percentile.

Measure Year	**Rate Distribution**	*	*	*	*	*	*	*	*	*
*	**N**	**Mean**	**StdDev**	**Min**	**Max**	**10th**	**25th**	**50th**	**75th**	**90th**
2020	4032	69.808	31.8445	0	100	18	48	82	99	100

**Cells intentionally left empty*

In 2017, MIPS replaced the Physician Quality Reporting System (PQRS) which ended in 2016. Clinician-level MIPS performance results from 2017 through 2019 are not available.

2016 Submission

2014 PQRS Experience Report

2014 is the most recent year for which PQRS Experience Report measure data is available. The average performance rates on Preventive Care and Screening: Influenza Immunization over the last several years are as follows:

2011: 50.4%

2012: 43.9%

2013: 46.8%

2014: 46.3%

It is important to note that PQRS has been and remains a voluntary reporting program. In the early years of the PQRS program, participants received an incentive for satisfactorily reporting. However, beginning in 2015, the program will impose payment penalties for non-participants based on 2013 performance. 62% of eligible professionals participated using any reporting option in 2014. As a result, performance rates may not be nationally representative.

[Response Ends]

1b.03. If no or limited performance data on the measure as specified is reported above, 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. Include citations.

[Response Begins]

The Centers for Disease Control and Prevention (CDC) indicate a significant opportunity for improvement in the rates of influenza vaccination. Influenza vaccination is the most effective protection against influenza virus infection (Centers for Disease Control and Prevention [CDC], 2022). Influenza may lead to serious complications including hospitalization or death (CDC, 2022). Influenza vaccine is recommended for all persons aged >=6 months who do not have contraindications to vaccination. However, data indicate that only half of all eligible individuals receive an influenza vaccination (CDC, 2022). This measure promotes annual influenza vaccination for all persons aged >= 6 months.

Reference Text: Centers for Disease Control and Prevention: https://www.cdc.gov/flu/about/index.html; https://www.cdc.gov/flu/fluvaxview/coverage-2021estimates.htm

[Response Ends]

1b.04. 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.

Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]

While this measure is included in federal reporting programs, those programs have not yet made disparities data available for us to analyze and report.

[Response Ends]

1b.05. If no or limited data on disparities from the measure as specified is reported above, 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 above.

[Response Begins]

2022 Submission

The Centers for Disease Control and Prevention (CDC) analyzed data from the National Immunization Survey-Flu and Behavioral Risk Factor Surveillance System (BRFSS) to estimate rates of flu vaccination from the 2020–21 flu season (CDC, 2021). Data, summarized below, reflect differences in flu vaccination by individual characteristics including age, gender and race/ethnicity. Data also suggest regional differences in rates of vaccination.

Estimates of flu vaccination rates by age, gender and racial/ethnic groups for the 2020–21 flu season:

Vaccination rates by age: Adults aged 18 years and older had lower rates of vaccination than children 6 months – 17 years.

All People >=6 months 52.1%

Children (6 months-17 years) 58.6%

Adults (>=18 years) 50.2%

Vaccination rates by gender: There were no differences in flu vaccination coverage between male and female children 6 months through 17 years. For adults 18 years and older, flu vaccination was generally higher among females than males.

Female adults >=18 years 53.9%

Male adults >=18 years 46.3%

Vaccination rates by race/ethnicity

Among people >=6 months, white people had higher flu coverage than Black, Hispanic, and people of other races. Additionally, people of other races had higher coverage than Black and Hispanic people. Lastly, Hispanic people had higher coverage than Black people.

All races/ethnicities 52.1%

White only, non-Hispanic 56.4%

Black only, non-Hispanic 42.7%

Hispanic 44.9%

Other, non-Hispanic** 52.1%

** Includes Asian, American Indian/Alaska Native, Native Hawaiian or other Pacific Islander, multiracial, and other races.

Furthermore, there was between-state variability in flu vaccination coverage among adults across reporting states. Rates ranged from 41.5% in Wyoming to 64.0% in Rhode Island.

Flu vaccination coverage: United States, 2020-21 Influenza Season. Centers for Disease Control and Prevention Web site. https://www.cdc.gov/flu/fluvaxview/coverage-2021estimates.htm. Updated October 7, 2021. Accessed December 15, 2021.

2016 Submission

The Centers for Disease Control and Prevention (CDC) analyzed data from the National Immunization Survey-Flu and Behavioral Risk Factor Surveillance System (BRFSS) to estimate rates of flu vaccination from the 2014–15 flu season (CDC, 2015). Data, summarized below, reflect differences in flu vaccination by individual characteristics including age, gender and race/ethnicity. Data also suggest regional differences in rates of vaccination.

Estimates of flu vaccination rates by age, gender and racial/ethnic groups for the 2014–15 flu season

Vaccination rates by age: Adults aged 18 years and older had lower rates of vaccination than children 6 months – 17 years.

All People >=6 months 47.1%

Children (6 months-17 years) 59.3 %

Adults (>=18 years) 43.6 %

Vaccination rates by gender: There were no differences in flu vaccination coverage between male and female children 6 months through 17 years. For adults 18 years and older, flu vaccination was generally higher among females than males.

Female adults >=18 years 47.0%

Male adults >=18 years 40.1%

Vaccination rates by race/ethnicity

Among people >=6 months, vaccination rates for non-Hispanic whites (48.5%) and Asians (51.0%) were higher than that of non-Hispanic blacks (43.8%), Hispanics (44.3%), and people of other or multiple races (44.3%).

All races/ethnicities 47.1%

White only, non-Hispanic 48.5%

Black only, non-Hispanic 43.8%

Hispanic 44.3%

Asian 51.0%

American Indian/Alaska Native 45.2%

Other or multiple race 44.3%

Furthermore, influenza vaccination rates varied by state and ranged from 39.2% in Florida to 59.6% in South Dakota.

Flu vaccination coverage: United States, 2014-15 Influenza Season. Centers for Disease Control and Prevention Web site. http://www.cdc.gov/flu/fluvaxview/coverage-1415estimates.htm. Updated September 17, 2015. Accessed June 20, 2016.

[Response Ends]

Criteria 2: 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.

---

spma.01. Indicate whether there are changes to the specifications since the last updates/submission. If yes, update the specifications in the Measure Specifications section of the Measure Submission Form, and explain your reasoning for the changes below.

[Response Begins]

No

[Response Ends]

spma.02. Briefly describe any important changes to the measure specifications since the last measure update and provide a rationale.

For annual updates, please explain how the change in specifications affects the measure results. If a material change in specification is identified, data from re-testing of the measure with the new specifications is required for early maintenance review.

For example, specifications may have been updated based on suggestions from a previous NQF CDP review.

[Response Begins]

N/A

[Response Ends]

sp.01. Provide the measure title.

Measure titles should be concise yet convey who and what is being measured (see What Good Looks Like).

[Response Begins]

Preventive Care and Screening: Influenza Immunization

[Response Ends]

sp.02. Provide a brief description of the measure.

Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).

[Response Begins]

Percentage of patients aged 6 months and older seen for a visit between October 1 and March 31 who received an influenza immunization OR who reported previous receipt of an influenza immunization

[Response Ends]

sp.04. Check all the clinical condition/topic areas that apply to your measure, below.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Surgery: General*

[Response Begins]

Infectious Diseases (ID): Influenza

[Response Ends]

sp.05. Check all the non-condition specific measure domain areas that apply to your measure, below.

[Response Begins]

Immunization

Primary Prevention

[Response Ends]

sp.06. Select one or more target population categories.

Select only those target populations which can be stratified in the reporting of the measure's result.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Populations at Risk: Populations at Risk*

[Response Begins]

Adults (Age >= 18)

Children (Age < 18)

Elderly (Age >= 65)

[Response Ends]

sp.07. Select the levels of analysis that apply to your measure.

Check ONLY the levels of analysis for which the measure is SPECIFIED and TESTED.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]

Clinician: Individual

[Response Ends]

sp.08. Indicate the care settings that apply to your measure.

Check ONLY the settings for which the measure is SPECIFIED and TESTED.

[Response Begins]

Other

[Response Ends]

sp.09. 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. If no URL is available, indicate "none available".

[Response Begins]

The measure specifications are included with this submission. Additional measure details may be found at https://qpp.cms.gov/docs/QPP_quality_measure_specifications/Claims-Registry-Measures/2021_Measure_110_MedicarePartBClaims.pdf

[Response Ends]

sp.11. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.

Attach an excel or csv file; if this poses an issue, contact staff. Provide descriptors for any codes. Use one file with multiple worksheets, if needed.

[Response Begins]

No data dictionary/code table – all information provided in the submission form

[Response Ends]

sp.12. State the numerator.

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.

[Response Begins]

Patients who received an influenza immunization OR who reported previous receipt of an influenza immunization.

[Response Ends]

sp.13. Provide details needed to calculate the numerator.

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 sp.11.

[Response Begins]

NUMERATOR:

Patients who received an influenza immunization OR who reported previous receipt of an influenza immunization

Definition: Previous Receipt – Receipt of the current season's influenza immunization from another provider OR from same provider prior to the visit to which the measure is applied (typically, prior vaccination would include influenza vaccine given since August 1st).

Numerator Instruction:

The numerator can be met by submitting either administration of an influenza vaccination or that the patient reported previous receipt of the current season's influenza immunization. If the performance of the numerator is not met, a clinician can submit a valid denominator exception for having not administered an influenza vaccination. For clinicians submitting a denominator exception, there should be a clear rationale and documented reason for not administering an influenza immunization if the patient did not indicate previous receipt, which could include a medical reason (e.g., patient allergy), patient reason (e.g., patient declined), or system reason (e.g., vaccination not available). The system reason should be indicated only for cases of disruption or shortage of influenza vaccination supply.

Due to the changing nature of the CDC/ACIP recommendations regarding the live attenuated influenza vaccine (LAIV) for a particular flu season, this measure will not include the administration of this specific formulation of the flu vaccination. Given the variance of the timeframes for the annual update cycles, program implementation, and publication of revised recommendations from the CDC/ACIP, it has been determined that the coding for this measure will specifically exclude this formulation, so as not to inappropriately include this form of the vaccine for flu seasons when CDC/ACIP explicitly advise against it. However, it is recommended that all eligible professionals or eligible clinicians review the guidelines for each flu season to determine appropriateness of the LAIV and other formulations of the flu vaccine. Should the LAIV be recommended for administration for a particular flu season, an eligible professional or clinician may consider one of the following options: 1) satisfy the numerator by reporting previous receipt, 2) report a denominator exception, either as a patient reason (e.g., for patient preference) or a system reason (e.g., the institution only carries LAIV).

NUMERATOR NOTE: Denominator Exception(s) are determined at the time of the denominator eligible encounter during the current flu season.

Numerator Options:

Performance Met: Influenza immunization administered or previously received **(G8482)**

OR

Denominator Exception: Influenza immunization was not administered for reasons documented by clinician (e.g., patient allergy or other medical reasons, patient declined or other patient reasons, vaccine not available or other system reasons) **(G8483)**

OR

Performance Not Met: Influenza immunization was not administered, reason not given **(G8484)**

[Response Ends]

sp.14. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]

All patients aged 6 months and older seen for a visit between October 1 and March 31.

[Response Ends]

sp.15. Provide details needed to calculate the denominator.

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 sp.11.

[Response Begins]

DENOMINATOR NOTE: In order to submit on the flu season 2020-2021, the patient must have a qualifying encounter between January 1 and March 31, 2021. In order to submit on the flu season 2021-2022, the patient must have a qualifying encounter between October 1 and December 31, 2021. A qualifying encounter needs to occur within the flu season that is being submitted; any additional encounter(s) may occur at any time within the measurement period.

*Signifies that this CPT Category I code is a non-covered service under the Medicare Part B Physician Fee Schedule (PFS). These non-covered services should be counted in the denominator population for MIPS CQMs.

Denominator Criteria (Eligible Cases):

Patients aged ≥ 6 months

AND

Patient encounter during January thru March and/or October thru December (CPT or HCPCS): 90945, 90947, 90951, 90952, 90953, 90954, 90955, 90956, 90957, 90958, 90959, 90960, 90961, 90962, 90963, 90964, 90965, 90966, 90967, 90968, 90969, 90970, 99202, 99203, 99204, 99205, 99212, 99213, 99214, 99215, 99241*, 99242*, 99243*, 99244*, 99245*, 99304, 99305, 99306, 99307, 99308, 99309, 99310, 99315, 99316, 99324, 99325, 99326, 99327, 99328, 99334, 99335, 99336, 99337, 99341, 99342, 99343, 99344, 99345, 99347, 99348, 99349, 99350, 99381*, 99382*, 99383*, 99384*, 99385*, 99386*, 99387*, 99391*, 99392*, 99393*, 99394*, 99395*, 99396*, 99397*, 99401*, 99402*, 99403*, 99404*, 99411*, 99412*, 99429*, 99512*, G0438, G0439

[Response Ends]

sp.16. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]

None.

[Response Ends]

sp.17. Provide details needed to calculate the denominator exclusions.

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 sp.11.

[Response Begins]

N/A

[Response Ends]

sp.18. Provide all information required to stratify the measure results, if necessary.

Include 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 in the Data Dictionary field.

[Response Begins]

N/A

[Response Ends]

sp.19. Select the risk adjustment type.

Select type. Provide specifications for risk stratification and/or risk models in the Scientific Acceptability section.

[Response Begins]

No risk adjustment or risk stratification

[Response Ends]

sp.20. Select the most relevant type of score.

Attachment: If available, please provide a sample report.

[Response Begins]

Rate/proportion

[Response Ends]

sp.21. Select the appropriate interpretation of the measure score.

Classifies interpretation of score according to whether better quality or resource use is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score

[Response Begins]

Better quality = Higher score

[Response Ends]

sp.22. Diagram or describe the calculation of the measure score as an ordered sequence of steps.

Identify the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period of data, aggregating data; risk adjustment; etc.

[Response Begins]

1. Start with Denominator
2. Check *Patients aged greater than or equal to 6 months*:
    a. If *Patients aged greater than or equal to 6 months* equals No, do not include in *Eligible Population/Denominator*. Stop processing.
    b. If *Patients aged greater than or equal to 6 months* equals Yes, proceed to check *Patient encounter during January thru March and/or October thru December as listed in Denominator*/***.
3. Check *Patient encounter during January thru March and/or October thru December as listed in Denominator*/***:
    a. If *Patient encounter during January thru March and/or October thru December as listed in Denominator*/*** equals No, do not include in *Eligible Population/Denominator*. Stop processing.
    b. If *Patient encounter during January thru March and/or October thru December as listed in Denominator*/*** equals Yes, include in *Eligible Population/Denominator*.
4. Denominator Population:
    a. Denominator Population is all Eligible Patients in Denominator. Denominator is represented as Denominator in the Sample Calculation listed at the end of this document. Letter d equals 80 patients in the Sample Calculation.
5. Start Numerator
6. Check *Influenza immunization administered or previously received*:
    a. If *Influenza immunization administered or previously received* equals Yes, include in *Data Completeness Met and Performance Met*.
        i. *Data Completeness Met and Performance Met* letter is represented in the Data Completeness and Performance Rate in the Sample Calculation listed at the end of this document. Letter a equals 30 patients in the Sample Calculation.
    b. If *Influenza immunization administered or previously received* equals No, proceed to check *Influenza immunization was not administered for reasons documented by clinician*.
7. Check *Influenza immunization was not administered for reasons documented by clinician*:
    a. If *Influenza immunization was not administered for reasons documented by clinician* equals Yes, include in *Data Completeness Met and Denominator Exception*.
        i. *Completeness Met and Denominator Exception* letter is represented in the Data Completeness and Performance Rate in the Sample Calculation listed at the end of this document. Letter b equals 10 patients in the Sample Calculation.
    b. If *Influenza immunization was not administered for reasons documented by clinician* equals No, proceed to check *Influenza immunization was not administered, reason not given*.
8. Check *Influenza immunization was not administered, reason not given*:
    a. If *Influenza immunization was not administered, reason not given* equals Yes, include in the *Data Completeness Met and Performance Not Met*.

i. *Data Completeness Met and Performance Not Met* letter is represented in the Data Completeness in the Sample Calculation listed at the end of this document. Letter c equals 30 patients in the Sample Calculation.

b. If *Influenza immunization was not administered, reason not given* equals No, proceed to check *Data Completeness Not Met*.

9. Check *Data Completeness Not Met*:
    a. If *Data Completeness Not Met*, the Quality Data Code or equivalent was not submitted. 10 patients have been subtracted from the Data Completeness Numerator in the Sample Calculation.

[Response Ends]

sp.25. If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

[Response Begins]

N/A

[Response Ends]

sp.28. Select only the data sources for which the measure is specified.

[Response Begins]

Claims

Registry Data

[Response Ends]

sp.29. Identify the specific data source or data collection instrument.

For example, provide the name of the database, clinical registry, collection instrument, etc., and describe how data are collected.

[Response Begins]

N/A

[Response Ends]

sp.30. Provide the data collection instrument.

[Response Begins]

No data collection instrument provided

[Response Ends]

2ma.01. Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Reliability - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

Yes

[Response Ends]

2ma.02. Indicate whether additional empirical validity testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Validity - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

Yes

[Response Ends]

2ma.03. For outcome, patient-reported outcome, resource use, cost, and some process measures, risk adjustment/stratification may be conducted. Did you perform a risk adjustment or stratification analysis?

[Response Begins]

No

[Response Ends]

2ma.04. For maintenance measures in which risk adjustment/stratification has been performed, indicate whether additional risk adjustment testing has been conducted since the most recent maintenance evaluation. This may include updates to the risk adjustment analysis with additional clinical, demographic, and social risk factors.

Please update the Scientific Acceptability: Validity - Other Threats to Validity section.

Note: This section must be updated even if social risk factors are not included in the risk adjustment strategy.

[Response Begins]

No additional risk adjustment analysis included

[Response Ends]

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing

information and results should be entered in the appropriate fields in the Scientific Acceptability sections of the Measure Submission Form.

- 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.
- All required sections must be completed.
- For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.
- If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.
- An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.
- Contact NQF staff with any questions. Check for resources at the 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 the 2021 Measure Evaluation Criteria and Guidance.

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.

2a. Reliability testing 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 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;

AND

If patient preference (e.g., informed decision-making) 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).

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 non-responders) and how the specified handling of missing data minimizes bias.

2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:

2c1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and

2c2. the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

(if not conducted or results not adequate, justification must be submitted and accepted)

Definitions

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).

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.

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.

Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

Risk factors that influence outcomes should not be specified as exclusions.

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.

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Scientific Acceptability sections. For example:

2021 Submission:

Updated testing information here.

2018 Submission:

Testing from the previous submission here.

2a. Reliability

2a.01. Select only the data sources for which the measure is tested.

[Response Begins]

Claims

Registry Data

[Response Ends]

2a.02. 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).

[Response Begins]

The data source is claims and registry data from the MIPS program, provided by the Center for Medicare & Medicaid Services (CMS).

[Response Ends]

2a.03. Provide the dates of the data used in testing.

Use the following format: "MM-DD-YYYY - MM-DD-YYYY"

[Response Begins]

01-01-2020 – 12-31-2020

[Response Ends]

2a.04. Select the levels of analysis for which the measure is tested.

Testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]

Clinician: Individual

[Response Ends]

2a.05. List the measured entities 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.

[Response Begins]

2022 Submission

CMS does not report descriptive data at the clinician level.

2016 Submission

The total number of physicians reporting on this measure, via the EHR option, in 2014, is 24,299. Of those, 18,247 physicians had all the required data elements and met the minimum number of quality reporting events (10) for a total of 4,158,205 quality events. For this measure, 75.1 percent of physicians are included in the analysis, and the average number of quality reporting events after exceptions are removed is 223.6 for the remaining 4,079,421 events. The range of quality reporting events for 18,247 physicians included is from 7148 to 10. The average number of quality reporting events for the remaining 24.9 percent of physicians that aren't included is 0.10.

[Response Ends]

2a.06. Identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis), separated by level of analysis and data source; if a sample was used, describe how patients were selected for inclusion in the sample.

If there is a minimum case count used for testing, that minimum must be reflected in the specifications.

[Response Begins]

CMS does not report descriptive data at the patient level.

[Response Ends]

2a.07. 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.

[Response Begins]

The same data samples were used for all aspects of testing.

[Response Ends]

2a.08. List 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.

[Response Begins]

CMS does not report patient-level socio-demographic data.

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.07 check patient or encounter-level data; in 2a.08 enter "see validity testing section of data elements"; and enter "N/A" for 2a.09 and 2a.10.

2a.09. Select the level of reliability testing conducted.

Choose one or both levels.

[Response Begins]

Accountable Entity Level (e.g., signal-to-noise analysis)

[Response Ends]

2a.10. For each level of reliability testing 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.

[Response Begins]

We utilized the methodology described by John Adams (Adams, J.L. The Reliability of Provider Profiling: A Tutorial. Santa Monica, California: RAND Corporation. TR-653-NCQA, 2009) to calculate signal-to-noise reliability. This methodology uses the Beta-binomial model to assess how well one can confidently distinguish the performance of one reporting 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 across reporting entities in performance. The Beta-binomial model is an appropriate model when estimating the reliability of simple pass/fail rate measures, such as the influenza measure. 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 reporting entities.

For the influenza measure, the provider is the reporting entity. It is a percentage, bounded by 0 and 100, indicating the proportion of people who received an influenza vaccination shot that year.

The formula for signal-to-noise reliability is:

Signal-to-noise reliability = $\sigma^2_{\text{provider-to-provider}} / (\sigma^2_{\text{provider-to-provider}} + \sigma^2_{\text{error}})$

$$\text{Signal-to-noise reliability} = \sigma^2_{\text{provider-to-provider}} / (\sigma^2_{\text{provider-to-provider}} + \sigma^2_{\text{error}})$$

Therefore, we need to estimate two variances: 1) variance between providers ($\sigma^2_{\text{provider-to-provider}}$); 2) variance within providers (σ^2_{error}).

1. Variance between providers = $\sigma^2_{\text{provider-to-provider}} = (\alpha\ \beta) / (\alpha + \beta + 1)(\alpha + \beta)^2$

   α and β are two shape parameters of the Beta-Binomial distribution, $\alpha > 0$, $\beta > 0$

2. Variance within provider: $\sigma^2_{\text{error}} = \hat{p}\ (1- \hat{p})/n$

   $\hat{p}$ = observed rate for the provider

n = provider-specific denominator for the observed rate (most often the number of eligible patients)

Using Adams' (2009) methodology, we estimated the reliability for each reporting entity, then averaged these reliability estimates across all reporting entities to produce a point estimate of signal-to-noise reliability. We label this point estimate "mean signal-to-noise reliability". The mean signal-to-noise reliability measures how well, on average, the measure can differentiate between reporting entity performance on the measure.

Along with the point estimate of mean signal-to-noise reliability, we are also providing the distribution of the provider-level signal-to-noise reliability estimates. Each reporting unit's reliability estimate is a ratio of signal to noise, as described above [$\sigma^2_{\text{provider-to-provider}} / (\sigma^2_{\text{provider-to-provider}} + \sigma^2_{\text{error}})$]. Variability between reporting units ($\sigma^2_{\text{provider-to-provider}}$) is the same for each unit, while the specific reporting unit error (σ^2_{error}) varies. Reliability for

each reporting unit is an ordinal measure of how well one can determine where that entity lies in the distribution across reporting units, with higher estimates indicating better reliability.

This methodology allows us to estimate the reliability for each provider and summarize the distribution of these estimates.

[Response Ends]

2a.11. For each level of reliability testing checked above, what were the statistical results from reliability testing?

For example, provide the percent agreement and kappa for the critical data elements, or distribution of reliability statistics from a signal-to-noise analysis. For score-level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred (pg. 18, NQF Measure Evaluation Criteria).

[Response Begins]

2022 Submission

We estimated the reliability for each clinician for 2020 performance year reporting. The mean reliability is 0.997.

Reliability Distribution	*	*	*	*	*	*	*	*	*
N	**mean**	**min**	**10th**	**25th**	**50th**	**75th**	**90th**	**max**	**sdev**
4032	0.997	0.921	0.993	0.997	0.999	1	1	1	0.006

**Cells intentionally left empty*

[Response Ends]

2a.12. Interpret the results, in terms of how they demonstrate reliability.

(In other words, what do the results mean and what are the norms for the test conducted?)

[Response Begins]

Results indicate very good/very high reliability.

[Response Ends]

2b. Validity

2b.01. Select the level of validity testing that was conducted.

[Response Begins]

Accountable Entity Level (e.g. hospitals, clinicians)

[Response Ends]

2b.02. 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.

[Response Begins]

NCQA performed Pearson correlation to determine whether the influenza measure results correlate with another immunization measure: *Pneumococcal Vaccination Status for Older Adults*. The test estimates the strength of linear association between two continuous variables; the magnitude of correlation ranges from -1 and +1. A value of 1 indicates a strong positive linear association: an increase in values of one variable is associated with increase in value of another variable. A value of 0 indicates no linear association. A value of -1 indicates a strong negative relationship in which an increase in values of the first variable is associated with a decrease in values of the second variable. 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.

[Response Ends]

2b.03. Provide the statistical results from validity testing.

Examples may include correlations or t-test results.

[Response Begins]

The influenza measure is positively and strongly associated with the pneumococcal vaccination measure ($r = 0.727$ $p < 0.001$).

[Response Ends]

2b.04. Provide 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?)

[Response Begins]

The Influenza immunization measure performance correlates with the pneumococcal vaccination measures as indicated by the Pearson correlation test. This finding suggests that providers who perform well on the influenza measure will likely perform well on the pneumococcal measure, which is expected from a conceptual standpoint given both measures assess immunization services.

[Response Ends]

2b.05. 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 in Importance to Measure and Report: Gap in Care/Disparities.

[Response Begins]

To demonstrate meaningful differences in performance, NCQA calculated an inter-quartile range (IQR) for each measure. 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 calculated an independent sample t-test of the performance difference between two randomly selected reporting units from each group (below 25th and above 75th percentiles). The t-test method calculates a testing statistic based on the sample size, performance rate, and standardized error of each reporting unit. 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 reporting units' performance are significantly different from each other.

[Response Ends]

2b.06. Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.

Examples may include 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.

[Response Begins]

NCQA calculated the distribution of clinician-level performance for the Influenza Immunization measure. There is a 51-point gap in performance at the 25th and 75th percentiles. The difference in performance between reporting units in these percentiles is statistically significant.

Validity (t-test) Output	*	*	*	*	*	*	*
p25	**p75**	**Denominator LowQ**	**Denominator TopQ**	**Rate LowQ**	**Rate TopQ**	**z**	**p_value_interpret**
48	99	1776	724	12	100	114.123	p < 0.001

**Cells intentionally left empty*

[Response Ends]

2b.07. Provide 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.

In other words, what do the results mean in terms of statistical and meaningful differences?

[Response Begins]

The difference in performance between reporting units is statistically significant.

[Response Ends]

2b.08. Describe the method of testing conducted to identify the extent and distribution of missing data (or non-response) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders). Include how the specified handling of missing data minimizes bias.

Describe the steps—do not just name a method; what statistical analysis was used.

[Response Begins]

N/A

[Response Ends]

2b.09. Provide the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data.

For example, provide results of sensitivity analysis of the effect of various rules for missing data/non-response. If no empirical sensitivity analysis was conducted, identify the approaches for handling missing data that were considered and benefits and drawbacks of each).

[Response Begins]

N/A

[Response Ends]

2b.10. Provide 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 non-responders), and how the specified handling of missing data minimizes bias.

In other words, 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 was conducted, justify the selected approach for missing data.

[Response Begins]

N/A

[Response Ends]

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 eCQMs). 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.

2b.11. Indicate whether there is more than one set of specifications for this measure.

[Response Begins]

No, there is only one set of specifications for this measure

[Response Ends]

2b.12. 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. Indicate what statistical analysis was used.

[Response Begins]

[Response Ends]

2b.13. Provide the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications.

Examples may include correlation, and/or rank order.

[Response Begins]

[Response Ends]

2b.14. Provide your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications.

In other words, what do the results mean and what are the norms for the test conducted.

[Response Begins]

[Response Ends]

2b.15. Indicate whether the measure uses exclusions.

[Response Begins]

N/A or no exclusions

[Response Ends]

2b.16. Describe the method of testing exclusions and what was tested.

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?

[Response Begins]

N/A

[Response Ends]

2b.17. Provide 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.

[Response Begins]

N/A

[Response Ends]

2b.18. Provide your interpretation of the results, in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results.

In other words, 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.

[Response Begins]

N/A

[Response Ends]

2b.19. Check all methods used to address risk factors.

[Response Begins]

No risk adjustment or stratification

[Response Ends]

2b.20. If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.

[Response Begins]

[Response Ends]

2b.21. If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.

[Response Begins]

N/A

[Response Ends]

2b.22. Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.

[Response Begins]

[Response Ends]

2b.23. Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.

Please be sure to address the following: potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10 or other statistical tests; correlation of x or higher. Patient factors should be present at the start of care, if applicable. Also discuss any "ordering" of risk factor inclusion; note whether social risk factors are added after all clinical factors. Discuss any considerations regarding data sources (e.g., availability, specificity).

[Response Begins]

[Response Ends]

2b.24. Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.

[Response Begins]

[Response Ends]

2b.25. Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.

Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.

[Response Begins]

[Response Ends]

2b.26. 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 control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter "N/A" for questions about the statistical risk model discrimination and calibration statistics.

Validation testing should be conducted in a data set that is separate from the one used to develop the model.

[Response Begins]

[Response Ends]

2b.27. Provide risk model discrimination statistics.

For example, provide c-statistics or R-squared values.

[Response Begins]

[Response Ends]

2b.28. Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).

[Response Begins]

N/A

[Response Ends]

2b.29. Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.

The preferred file format is .png, but most image formats are acceptable.

[Response Begins]

[Response Ends]

2b.30. Provide the results of the risk stratification analysis.

[Response Begins]

[Response Ends]

2b.31. Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).

In other words, what do the results mean and what are the norms for the test conducted?

[Response Begins]

[Response Ends]

2b.32. Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.

Not required but would provide additional support of adequacy of the risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed.

[Response Begins]

[Response Ends]

Criterion 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.

3.01. Check all methods below that are used to generate the data elements needed to compute the measure score.

[Response Begins]

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-10 codes on claims)

[Response Ends]

3.02. Detail to what extent the specified data elements are available electronically in defined fields.

In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.

[Response Begins]

ALL data elements are in defined fields in a combination of electronic sources

[Response Ends]

3.03. 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 data elements not from electronic sources.

[Response Begins]

N/A

[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]

There is currently an eCQM version of this measure: CMS147.

[Response Ends]

3.06. 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.

[Response Begins]

N/A

[Response Ends]

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

3.07. Detail 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),

Attach the fee schedule here, if applicable.

[Response Begins]

COPYRIGHT:

This Physician Performance Measure (Measure) and related data specifications are owned and copyrighted by the National Committee for Quality Assurance (NCQA). NCQA is not responsible for any use of the Measure. The 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 MEASURE AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

NCQA makes no representations, warranties, or endorsement about the quality of any organization or physician that uses or reports performance measures and NCQA has no liability to anyone who relies on such measures or specifications.

The Measure can be reproduced and distributed, without modification, for noncommercial purposes (e.g., use by healthcare providers in connection with their practices) without obtaining approval from NCQA. Commercial use is defined as the sale, licensing, or distribution of the Measure for commercial gain, or incorporation of the Measure into a product or service that is sold, licensed or distributed for commercial gain. All commercial uses or requests for modification must be approved by NCQA and are subject to a license at the discretion of NCQA. The PCPI's and AMA's significant past efforts and contributions to the development and updating of the measure are acknowledged.

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

[Response Ends]

Criterion 4: Use and Usability

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.

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making.

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 demonstrating performance improvement.

4a. Use

4a.01. Check all current uses. For each current use checked, please provide:

Name of program and sponsor

URL

Purpose

Geographic area and number and percentage of accountable entities and patients included

Level of measurement and setting

[Response Begins]

Public Reporting

[Public Reporting Please Explain]

- Name of program and sponsor

  The Quality Payment Program, CMS

- URL:
  - Purpose: To assess if influenza vaccine was offered/administered to patients

  Medicare Part B Claims: https://qpp.cms.gov/docs/QPP_quality_measure_specifications/Claims-Registry-Measures/2022_Measure_110_MedicarePartBClaims.pdf

  MIPS CQM: https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2022_Measure_110_MIPSCQM.pdf

  eCQM: https://ecqi.healthit.gov/ecqm/ep/2022/cms147v11

- Geographic area and number and percentage of accountable entities and patients included: Unknown
- Level of measurement and setting:
  - Level of measurement: Clinician: Group/Practice, Clinician: Individual
  - Setting: Home Care, Other, Outpatient Services, Post-Acute Care

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Measure Currently in Use

[Response Ends]

4a.03. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing), explain why the measure is not in use.

For example, do policies or actions of the developer/steward or accountable entities restrict access to performance results or block implementation?

[Response Begins]

The PCPI supports the expanded use of this measure in government or other programs, including those intended for accountability or public reporting. The PCPI does not have any policies that would restrict access to the performance measure specifications or results or that would impede implementation of the measure for any application. We would welcome its implementation in emerging applications such as accountable care organizations (ACO), Medicare Advantage insurance plans or health plans selling on the new insurance marketplace.

[Response Ends]

4a.04. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes: used in any accountability application within 3 years, and publicly reported within 6 years of initial endorsement.

A 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.

[Response Begins]

As described above, we understand that CMS is also planning to move towards publicly reporting physician data via Physician Compare. Also, although the measure is currently in use, we support expanded use of this measure in government or other programs, including those intended for accountability or public reporting.

[Response Ends]

4a.05. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

Detail 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.

[Response Begins]

CMS publishes measure performance results, and scores on its Physician Compare website. Performance year 2020 MIPS scores are publicly available and identifiable by clinician and group. Consumers will be able to see their clinicians rated against national peers on a scale of 0 to 100.

[Response Ends]

4a.06. Describe the process for providing measure results, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

[Response Begins]

CMS publishes results annually on its Physician Compare website during the year after the performance year.

[Response Ends]

4a.07. Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how feedback was obtained.

[Response Begins]

N/A

[Response Ends]

4a.08. Summarize the feedback obtained from those being measured.

[Response Begins]

N/A

[Response Ends]

4a.09. Summarize the feedback obtained from other users.

[Response Begins]

N/A

[Response Ends]

4a.10. Describe how the feedback described has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

[Response Begins]

N/A

[Response Ends]

4b. Usability

4b.01. You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, 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, provide an explanation. 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.

[Response Begins]

Trending data for MIPS reporting is not available for this measure.

[Response Ends]

4b.02. Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.

[Response Begins]

We are not aware of any unintended consequences at this time, but we take unintended consequences very seriously and therefore continuously monitor to identify actions that can be taken to mitigate them.

[Response Ends]

4b.03. Explain any unexpected benefits realized from implementation of this measure.

[Response Begins]

N/A

[Response Ends]

Criterion 5: Related and 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.

If you are updating a maintenance measure submission for the first time in MIMS, please note that the previous related and competing data appearing in question 5.03 may need to be entered in to 5.01 and 5.02, if the measures are NQF endorsed. Please review and update questions 5.01, 5.02, and 5.03 accordingly.

5.01. Search and select all NQF-endorsed related measures (conceptually, either same measure focus or target population).

(Can search and select measures.)

[Response Begins]

3620: Adult Immunization Status

0431: INFLUENZA VACCINATION COVERAGE AMONG HEALTHCARE PERSONNEL

0226: Influenza Immunization in the ESRD Population (Facility Level)

0038: Childhood Immunization Status (CIS)

1659: Influenza Immunization

3484: Prenatal Immunization Status

0680: Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short-Stay)

[Response Ends]

5.02. Search and select all NQF-endorsed competing measures (conceptually, the measures have both the same measure focus or target population).

(Can search and select measures.)

[Response Begins]

[Response Ends]

5.03. If there are related or competing measures to this measure, but they are not NQF-endorsed, please indicate the measure title and steward.

[Response Begins]

0039 : Flu Vaccinations for Adults Ages 18 and Older; National Committee for Quality Assurance

0522 : Influenza Immunization Received for Current Flu Season (Home Health); Centers for Medicare & Medicaid Services

[Response Ends]

5.04. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s), indicate whether the measure specifications are harmonized to the extent possible.

[Response Begins]

No

[Response Ends]

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

[Response Begins]

Related measures have differing target populations from measure 0041 Preventive Care and Screening: Influenza Immunization. Measure #0041 is intended to evaluate adherence to the current recommendations of the Advisory Committee on Immunization Practices. The Committee recommends routine annual influenza vaccination for all persons aged >=6 months who do not have contraindications. Measure #0039 - Flu Vaccinations for Adults ages 18 and Older focuses on the self-reported receipt of influenza vaccination among adults using the CAHPS survey. Measure #0226 – Influenza Immunization in the ESRD Population is a facility level measure focused on influenza vaccination among end stage renal disease (ESRD) patients receiving hemodialysis or peritoneal dialysis. Measure #0431 - Influenza Vaccination Coverage Among Healthcare Personnel focuses on influenza vaccination among healthcare workers. Measure #0522 Influenza Immunization Received for Current Flu Season (Home Health) evaluates influenza immunization during home health episodes of care. Measure # 0680 Percent of Residents or Patients Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (short stay) applies to patients of Inpatient Rehabilitation Facilities and Long-Term Care Hospitals, and to short-stay nursing home residents. Measure #0681 - Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (long stay) assess influenza vaccination among long-stay nursing facility residents. Measure #1659 Influenza Immunization is limited to the assessment of influenza vaccination upon discharge from the inpatient setting.

[Response Ends]

5.06. Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality). Alternatively, justify endorsing an additional measure.

Provide analyses when possible.

[Response Begins]

N/A no competing measures

[Response Ends]