NATIONAL QUALITY FORUM

MEASURE WORKSHEET

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

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

Brief Measure Information

NQF #: 3483

Corresponding Measures:

De.2. Measure Title: Adult Immunization Status

Co.1.1. Measure Steward: National Committee for Quality Assurance

De.3. Brief Description of Measure: Percentage of adults 19 years of age and older who are up-to-date on recommended routine vaccines for influenza, tetanus and diphtheria (Td) or tetanus, diphtheria and acellular pertussis (Tdap), zoster and pneumococcal.

1b.1. Developer Rationale: The composite rate provides an overview of how many routine adult vaccines were received out of the total that were recommended for a health plan member population. The individual vaccine component rates are included to provide information on which types of adult vaccinations are being provided to members as recommended.

S.4. Numerator Statement: Adults who are up-to-date on influenza, Td or Tdap, herpes zoster and pneumococcal vaccinations based on age and recommendations.

S.6. Denominator Statement: Adults ages 19 years and older.

S.8. Denominator Exclusions: Adults who received chemotherapy, had a bone marrow transplant or were in hospice during the measurement year or those with a history of immunocompromising conditions.

De.1. Measure Type: Composite

S.17. Data Source: Claims, Electronic Health Data, Electronic Health Records, Enrollment Data, Management Data, Other, Registry Data

S.20. Level of Analysis: Health Plan

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

IF this measure is included in a composite, NQF Composite#/title: Percentage of adults 19 years of age and older who are up-to-date on recommended routine vaccines for influenza, tetanus and diphtheria (Td) or tetanus, diphtheria and acellular pertussis (Tdap), zoster and pneumococcal

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? N/A

Criteria 1: Importance to Measure and Report

1a. Evidence

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

The developer provides the following evidence for this measure:

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

Evidence Summary

- Developer cited guidelines for each of the adult vaccines which are referenced in the measure description and the Advisory Committee on Immunization Practices (ACIP) recommendations. A brief description of the body of evidence for each vaccine is below
- **Influenza Vaccine Recommendation**. The ACIP Influenza Work Group reviewed available data and evidence from 1979 to 2018 on immunogenicity, efficiency, effectiveness and safety of influenza vaccines. In total, the Work Group reviewed approximately 285 studies on the immunogenicity, efficacy and effectiveness of IIV, RIV and LAIV and 120 studies on influenza vaccine safety. This review included approximately 123 studies specifically assessing immunogenicity, efficacy and effectiveness of influenza vaccine for adults; and 36 studies assessing influenza vaccine safety for adults. These studies consist of randomized control trials, case control studies and observational studies, among others.
- **Td/Tdap Vaccine Recommendation.** The ACIP Pertussis Vaccines Work Group reviewed available published and unpublished data and evidence from 2004 to 2017, covering topics such as tetanus, diphtheria and pertussis disease epidemiology in the United States, decision analyses, cost-effectiveness, programmatic considerations, vaccine immunogenicity, vaccine safety, and postlicensure vaccine effectiveness. In total, they reviewed 110 studies consisting of randomized control trials and other types of studies on Td and Tdap vaccination.
- **Herpes Zoster Vaccination.** The ACIP Herpes Zoster Vaccines Work Group evaluated studies published from 2015-2017 on the efficacy, cost-effectiveness, and safety of both RZV and ZVL. Their review included 10 studies of RZV, including seven randomized control trials (RCTs). They reviewed 40 studies of ZVL, including 16 high-quality RCTs, 13 RCTs with noted limitations, 10 cohort studies, and 1 case control study.
- **Pneumococcal Vaccination.** The ACIP Pneumococcal Work Group evaluated studies published from 2004-2014 on benefits, harms, values and preferences, and cost-effectiveness on PCV13 for routine use among adults aged 65 years and older. Their review included 6 randomized control trials (RCTs) on immunogenicity, 3 RCTs on serious and systemic adverse events and 2 other RCTs that they determined were of high and moderate quality.

Questions for the Committee:

- *Does the Committee agree the evidence basis for the measure?*

Guidance from the Evidence Algorithm

Measure a health outcome (Box 1) No → Assess performance of intermediate outcome, process, or structure(Box 3) Yes → Empirical evidence without SR or QQC (Box 4) yes → Grade for evidence (Box 5a) Yes → High

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

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

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

The developer cited data extracted from HEDIS data collection reflecting the most most recent year of measurement (2018) for this measure. Performance data is summarized at the health plan level and summarized by mean performance and performance at 10th, 25th, 50th, 75th, and 90th percentile. The interquartile range (IQR) was also extracted, which can be interpreted as the difference between the 25th and 75th percentile. Data is stratified by product line (i.e. commercial, Medicaid and Medicare).

The following data demonstrate the variation in the rate of adult immunization across health plans. For the composite rate, there was an 11 point difference between plans in the 25th percentile and plans in the 75th percentile for Medicare plans, and 7 and 8 points for commercial and Medicaid plans, respectively. These gaps in performance underscore the opportunity for improvement.

Adult Immunization Status: Composite

Commercial, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

21.2 | 13.5 | 16.0| 18.1 | 22.9 | 30.2 | 6.9

Medicaid, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

14.0 | 2.7 | 10.0 | 13.7 | 17.5 | 20.7 | 7.5

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

19.5 | 6.6 | 9.5 | 14.4 | 20.6 | 43.8 | 11.1

Adult Immunization Status: Influenza

Commercial, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

18.7 | 11.4 | 14.9 | 18.1 | 20.6 | 26.4 | 5.7

Medicaid, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

11.6 | 2.8 | 7.8 | 11.7 | 15.3 | 20.7 | 26.0 | 7.5

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

18.3 | 5.3 | 8.3 | 12.5 | 21.7 | 30.1 | 13.4

Adult Immunization Status: Td/Tdap

Commercial, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

29.4 | 18.6 | 20.9 | 25.2 | 30.7 | 46.5 | 9.8

Medicaid, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

20.9 | 4.9 | 14.1 | 21.1 | 25.0 | 34.4 | 10.9

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

26.5 | 9.5 | 14.8 | 20.7 | 28.7 | 56.4 | 13.9

Adult Immunization Status: Herpes Zoster

Commercial, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

6.1 | 2.7 | 4.1 | 5.0 | 6.5 | 9.8 | 2.4

Medicaid, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

1.6 | 0.0 | 0.4 | 0.6 | 1.4 | 5.3 | 1.0

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

12.9 | 0.4 | 0.9 | 5.3 | 14.5 | 39.5 | 13.6

Adult Immunization Status: Pneumococcal

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

20.3 | 5.4 | 8.1 | 10.8 | 22.5 | 55.5 | 14.4

The HEDIS performance data reflect the most recent year of measurement for this measure. Below is a description of the number of health plans that reported this measure and the median eligible population for the measure (stratified by commercial, Medicaid and Medicare).

Commercial, 2018

N Plans | Median Eligible Population

71 | 80,330

Medicaid, 2018

N Plans | Median Eligible Population

21 | 36,250

Medicare, 2018

N Plans | Median Eligible Population

44 | 11,648

Disparities
The developer cited that HEDIS data are stratified by type of insurance and while disparities data was not specified in the measure, it can be stratified by demographic variables such as race/ethnicity or socioeconomic status, in order to assess the presence of health care disparities, if the data are available to a plan.

Questions for the Committee:

- If no 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:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus

- No concerns
- Evidence appears to be relevant/applicable; derived from ACIP guidelines
- Strong evidence has been provided.
- The evidence is appropriate.
- The recommendations for each specific vaccine, and their timing, are well supported by evidence-based recommendations from ACIP. Rating: High
- High rating for evidence to support measure focus
- There is substantial evidence supporting the noted adult immunizations. The exclusion for hospice patients, however, does not appear to be supported by the literature.

1b. Performance Gap

- No disparities data provided, but can be analyzed by race/ethnicty. Results at the plan level were shared and there was a preformance gap between commercial and Medicid
- Current performance data was provided. Evidence provided shows that there is a gap between the desired rate of vaccination and actual vaccination. Data on subgroups of the insured population was provided, showing some disparities between types of coverage. Would expect more substantial disparities for non-covered or newly covered, given the barriers listed, but these aren't identified subgroups.
- Strong evidence provided re IQR for performance - much opportunity for improvement. There are significant and well documented inequities in adult vaccination rates (https://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2017.304257 & https://www.sciencedirect.com/science/article/pii/S0264410X16312713 etc. So many others!) There should absolutely be an exploration of the need to social risk adjust / stratify as a means to highlight these disparities.
- I agree with the moderate rating.
- The submission focuses on variation among health plans as the primary evidence of performance gaps. While this is important, it should also be noted how low the current rates are across the board for vaccines that are universally recommended. The submission does not address disparities, but does present data documenting differences between Commercial, Medicaid, and Medicare plans, which can be a proxy for disparities. I strongly suspect that there are major disparities by race, ethnicity, education, income, geography and other socio-demographic characteristics. Thus, I do believe that there are major performance gaps to address, the submission does not do a good job of presenting them. Rating: Moderate

- Significant gaps in performance underscore the opportunity for improvement for this important measure.
- Gaps across plans were noted. However, disparities related to race, ethnicity, socioeconomic status, georgraphic, etc., were not noted. Disparities withihn and across different groups may play a significant role in the effectiness of certain vaccines. In addition, data on the uninsured and immigrant populations also may significantly impact vaccine effectiveness. No data were provided about gaps other than gaps across insurance plans. Data should be required to better understand and address disparities. The measure developers also claim that the insurance plan individuals have serves as a "proxy for socioeconomic status." This might be the case for Medicaid, but certainly is not the case for Medicare or commercial plans.

1c. Composite Performance Measure

- Measure is a composite of critical routine immunizations, however this could be complicated by the fact that these are not routine or yearly immunizations. Immunizations may have occured at different time points and the patient may have changes in enrollment status over the time period.
- Yes
- This composite makes sense.
- No concerns noted
- The rationale for the composite measure is straightforward: all of the ACIP-recommended adult vaccines are included. This rationale is not, however, addressed in the submission. It should be noted that the definition of the composite is unusual: it is neither an all-or-nothing measure nor an average. Rather, it is defined for each age group as: Denominator: total number of recommended vaccines in the population monitored Numerator: total number of vaccines received in the same population For example, if there were a group of 100 individuals in an age-group in which 2 vaccines were recommended, and 20 and 40 individuals had received vaccines A and B respectively, the measure would be 60/200 = 0.30. Note that although this is a proportion, it does not have a binomial distribution since any individual's receiving vaccine A and B are probably not independent events. Rating: Moderate
- The stated quality construct is logical.
- There are drammatic differences across vaccines related to the potential for spreading the disease when unvaccinated. The spread of influenza, for instance, is quite different from shingles. The analysis of gaps and the composite measure do not address the issue of herd immunity and the spread of disease within population groups and geographic areas. In addition, the individual and societal impact of unvaccinated individuals is not equal across all of the diseases addressed by each of the vaccines in this composite set.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

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

2c. For composite measures: empirical analysis support composite approach

Reliability

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

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

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

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

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

Evaluators:

Methods Panel Review (Combined)

Methods Panel Evaluation Summary:

This measure was reviewed by the Scientific Methods Panel and discussed on the call. A summary of the measure and the Panel discussion is provided below.

Reliability

RELIABILITY: SPECIFICATIONS

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

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

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

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

Panel member 1: Following was my assessment on the measure specification in the prior submission:

Members are required to have continuous enrollment in the measurement period; however, assessment of immunization status may be much before the measurement period. For example, for influenza vaccine (see S.5), adults 19 years or older who received the flu vaccine on or between July 1st of the year prior to the measurement year and June 30th of the measurement year can qualify. This is even more of an issue for the other measures as, for example, for Tdap, as long as a member had the Tdap vaccine 9 years prior to the measurement period through the end of measurement period, the member would be considered eligible for the measure. How do we know that a member receiving a Tdap vaccine 5 years ago was part of the health plan unless it is required that the member had continuous enrollment going back to 5 years prior to the measurement period? Contrastingly, what if the health plan does not have immunization information of a

relatively new member despite the fact that this member had immunization (e.g., Tdap) 5 years ago, and had changed jobs multiple times and hence had different health plans? Thus, some clarity is needed as to how the past immunization statuses of individuals will be ascribed to a health plan.

I don't think this issues have been addressed in this submission. Only way by which a health plan can know about a member's prior vaccination status is through immunization registries that might be maintained by State or other third parties. However, this possibility is not explained in the testing document.

Panel member 2: No Concerns

Panel member 3: SMO: I think the calculation methods are relatively easy to understand or surmise but the description could be made more explicit. To avoid any ambiguity, the developers might want to modify the numerator statement to make it clear that each numerator is always a subset of a corresponding denominator. Exclusions such as adults with a history of adverse reactions could potentially be incorporated in the numerator/denominator statements. Details like "continuously enrolled" could also be in the denominator statements. If I am reading correctly, it seems like patients meeting the exclusion criteria are not actually excluded from the denominator but are instead included and counted as meeting the numerator criteria.

Panel member 4: None

Panel member 5: No response

Panel member 6: Specification lacks precision and likely could not be implemented as presented in the MIF. For example, in the numerator, the term 'up-to-date' is vague and should be replace with more precise language, as is 'based on age and recommendations'. More precision will make the statements longer but they will less open to interpretation. This seems like the bailiwick of the standing committee and they should weigh in on this issue, along with alignment with recent guidance for these

RELIABILITY: TESTING

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

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure**
   ☒ **Yes** ☒ **No**
   Panel member 4: The measure developer indicated integrated delivery system as a level of analysis; I did not see any data that demonstrated testing with integrated delivery systems.
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical <u>VALIDITY</u> testing** of <u>patient-level data</u> conducted?
   ☐ **Yes** ☐ **No**
6. **Assess the method(s) used for reliability testing**

   Panel member 1: Beta binomial model to estimate signal to noise ratio.

   Panel member 2: Beta-binomial model-ratio of signal -to -noise was appropriate

   Panel member 3: SMO: The near perfect reliability results seemed potentially too good to be true in the 2018 submission and I wondered if this might be an artifact of using data from only 3 plans per stratum when estimating parameters of the beta-binomial model. The addition of HEDIS data from ~135 plans in the 2019 cycle addressed this concern, and I think the near perfectly reliability estimates seem plausible. One possible way to provide greater assurance about this (if there is lingering concern about it) would be to report a confidence interval around the estimated reliability. It may also be helpful to report a graphical and/or tabular summary of the raw data (e.g. distribution of sample sizes and measure results) and the parameter estimates from the beta-binomial model. Strictly speaking, the beta-binomial model does not

seem to be a literally correct model for this measure because it ignores within-patient correlation across the 4 measures. In theory, I think this could lead to over-estimating reliability. On the other hand, I don't think this is the source of the near perfect reliability estimates and think the beta-binomial approach is a practical and acceptable approximation. One detail that should be provided is how developers dealt with unequal sample sizes per plan in their reliability calculations.

Panel member 4: Used beta-binominal for, which is given that the measure is a pass/fail measure.

Panel member 5: Beta-binomial model

Panel member 6: The developer used a signal-to-noise analysis with a beta-binomial model to estimate composite measure score reliability, and tested separately for Medicaid plans, Medicare plan and commercial plans. There were sufficient number of plans for testing. The approach used by the deverloper is acceptable and appropriate.

Submission document: Testing attachment, section 2a2.2

7. **Assess the results of reliability testing**

Panel member 1: Except for herpes zoster for Medicaid plans (Table 4), all of the estimated Beta-binomial coefficients are close to 1 indicating near-perfect reliability.

Panel member 2: Demonstrated strong statistical reliability

Panel member 3: SMO: The results indicate

Panel member 4: For all three plan types, the reliability statistic was 0.999 or higher for 50th percentile plus. May be worth discussing as a subgroup if that is "normal" or if we should dig deeper.

Panel member 5: Arrived at nearly perfect reliability.

Panel member 6: Taken at face value, the results indicate very high reliability for this composite measure, with minimal measurement error. The results for each payer type indicate that for the median plan, 100% of variability is due to difference in measure performance rather than measurement error (median reliability estimate of 1.0). This pattern is also seen for the composite measure components. These somewhat implausible results suggest there may be something specified incorrectly in the model, but we would need more information to determine this, at face value, the results suggest very high reliability.

Submission document: Testing attachment, section 2a2.3

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

Submission document: Testing attachment, section 2a2.2

☒ **Yes**

☐ **No**

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

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

Submission document: Testing attachment, section 2a2.2

☒ **Yes**

☐ **No**

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

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

☒**High** (NOTE: Can be HIGH <u>only if</u> score-level testing has been conducted)

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

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

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

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

Panel member 1: My rating for overall reliability is "moderate" because of the **measure with** reasons explained in #2 above. Unless it is specifically stated as to how an **integrated delivery system.**individual's immunization status from a time prior to the measurement period will be ascribed to a health plan either through continuous enrollment requirement or through other ways, the measure might lead to incorrect performance attribution to the health plan.

Panel member 2: No concerns, submitors analysis supported high reliability

Panel member 3: SMO: Rationale provided above

Panel member 4: The measure developer indicated integrated delivery system as a level of, but it appears as if they did not provide testing

Panel member 5: Based on method and result.

Panel member 6: Taken at face value, the results indicate very high reliability for this composite measure, with minimal measurement error.

Validity

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

Panel member 1: No exclusions

Panel member 2: No concerns

Panel member 3: SMO: None

Panel member 4: None

Panel member 5: None

Panel member 6: None

Submission document: Testing attachment, section 2b2.

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

Panel member 1: None

Panel member 2: No concerns

Panel member 3: SMO: None

Panel member 4: None. See substantial variation across all vaccination types and all plan types.

Panel member 5: Appears the IQRs are statistically significant (likely influenced by large sample size); however, practical difference across health plans isn't described.

Panel member 6: The composite easure appears to achieve a modest level of dispersion of health plans, with the largest IQRs observed among Medicare plans. The standing committee should evaluate whether these differences are meaningful from a clinical perspective.

Submission document: Testing attachment, section 2b4.

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

Panel member 1: N/A
Panel member 2: No response
Panel member 3: SMO: None:
Panel member 4: Not applicable.
Panel member 5: None
Panel member 6: No response

Submission document: Testing attachment, section 2b5.

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

Panel member 1: None
Panel member 2: No concerns
Panel member 3: SMO: None
Panel member 4: None. NCQA has standard processes for ensuring data are captured.
Panel member 5: None
Panel member 6: None

Submission document: Testing attachment, section 2b6.

16. **Risk Adjustment**

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

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

☒ Yes ☐ No ☒ Not applicable (**process measure**)

16c. **Social risk adjustment:**

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

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

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

16d.**Risk adjustment summary:**

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

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

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

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

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

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

Panel member 1: Only natural risk adjustment is through separate testing analysis for commercial, Medicaid and Medicare data, which is a potential proxy for socioeconomic status. No additional socioeconomic risks were adjusted primarily because such data are not typically collected in health plans.

Panel member 2: No response

Panel member 3: SMO: This is a process measure and was not risk-adjusted. In theory, differences across plans could be explained in part by case mix or by differences in the mix of eligibility status for the different vaccine measures in the composite.

Panel member 4: No response

Panel member 5: N/A

Panel member 6: N/A

VALIDITY: TESTING

17. **Validity testing level: ☒ Measure score ☒ Data element ☐ Both**
18. **Method of establishing validity of the measure score:**
    - ☒ **Face validity**
    - ☒ **Empirical validity testing of the measure score**
    - ☐ **N/A (score-level testing not conducted)**
19. **Assess the method(s) for establishing validity**

    Panel member 1: Pearson correlation coefficient was used to assess construct validity, while systematic face validity was established through different advisory panels and through public comments.

    Panel member 2: Pearson correlation for Empiric Testing was appropriate. Face Validity of Composite and Component Performance Measure Score thorough.

    Panel member 3: SMO: The developers assessed correlations between each individual measure in the composite and the overall composite and also compared score-level results for this measure to other HEDIS vaccine measures.

    Panel member 4: Face validity- process used does not seem to match NQF's criteria; process does not specifically address the questions that NQF wants asked
    Empirical testing – compared whether the indicators in the measure correlate with each other; also looked at performance on this measure vs. other HEDIS vaccination measures

    Panel member 5: Pearson correlation (component measures with whole and whole with other measures)

    Panel member 6: The developer examined the correlation among the AIS measure with other HEDIS vaccine measures and the intercorrelation of the composite components. This is a lower-bar of validity testing, but still acceptable. The developer also assessed face validity but did so using a non-systematic approach that does not conform with NQF's recommended approach of a systematic, standardized assessment.

    Submission document: Testing attachment, section 2b2.2

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

    Panel member 1: None

    Panel member 2: Health-Plan Level Pearson Correlation Coefficients of performance scores demonstrated moderate to high correlation values using the submitors thresholds.

Panel member 3: SMO: Results indicated that each individual measure was highly correlated with the composite. In addition, there was a positive correlation between performance on this composite and other HEDIS vaccine measures.

Panel member 4: Face validity –results provided only reflected NCQA's internal processes for measurement development

Empirical testing – generally showed strong correlations, **with indicators within the measure (Table 6); correlations with other HEDIS vaccination measures**. **were generally moderate (Tables 7-9)**

Panel member 5: Moderate to high correlations, with higher correlations among components with the whole.

Panel member 6: Composite measure component intercorrelations were strong and in the right direction, while the composite score and positive vaccine

Submission document: Testing attachment, section 2b2.3

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

    Submission document: Testing attachment, section 2b1.

    ☒ **Yes**

    ☒ **No**

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

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?**
    NOTE that data element validation from the literature is acceptable.

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

    ☒ **Yes**

    ☐ **No**

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

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

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

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

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

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

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

Panel member 1: No response

Panel member 2: The overall results in table 6 demonstrated strong correlation results but when the results were stratified by the individual health plans (tables 7-9) were significantly lower but in the "moderate" range as defined by the submitors. I felt that the defined "moderate" range was a wide

spread when compared to weak and strong ranges. There wasn't a good explanation of how they determined the thresholds for these levels. My question; is this acceptable by NQF standards?

Panel member 3: SMO: This measure appears to measure what it purports to measure and is related to quality.

Panel member 4: Measure developer used appropriate empirical approach; strong correlations with indicators within the measure (Table 6); correlations with other HEDIS vaccination measures were generally moderate (Tables 7-9). I'm not sure why the measure developer hypothesized that vaccination success with adults would translate to peds and adolescents (my experience is the engagement strategies vary across populations).

Panel member 5: Validating a measure with itself is not a strong approach. As expected the correlations dropped significantly when comparing to alternate measures (down to .29).

Panel member 6: Composite measure component intercorrelations were strong and in the right direction, while the composite score showed strong and positive correlations with other vaccine measures. However, this approach is lacking in methodological rigor.

FOR COMPOSITE MEASURES ONLY: Empirical analyses to support composite construction

25. **What is the level of certainty or confidence that the empirical analysis demonstrates that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct?**

    ☒ **High**

    ☐ **Moderate**

    ☒ **Low**

    ☐ **Insufficient**

26. **Briefly explain rationale for rating of EMPIRICAL ANALYSES TO SUPPORT COMPOSITE CONSTRUCTION**

Panel member 1: High internal consistency (Cronbach's alpha) of the components (e.g., individual vaccine rates) of the composite measure was demonstrated for each types of health plans (Commercial, Medicare, and Medicaid).

Panel member 2: Cronbach's alpha statistic demonstrated high internal consistency. Comparison of 2 different constructs was very thorough and supported their current construct.

Panel member 3: SMO: Each individual measure was highly correlated with the overall composite

Panel member 4: The analysis supports that health plans that do well on components of the also do well on the score..

Panel member 5: It is difficult to assess, possibly insufficient, but it appears problematic to have such variability in which vaccines are driving the composite score across plans. It is positied that this is due to differing quality improvement efforts—can this be tested? It seems important to validate what is driving the scores.

Panel member 6: Cronbach's alpha scores were uniformly high (>=0.95) across plan populations.

ADDITIONAL RECOMMENDATIONS

27. **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.**

Panel member 1: The testing documentation needs to clarify how it would attribute a member's immunization status from prior years without requiring continuous enrollment in the health plan (See my comment in #2 & #11).

Panel member 2: No response

Panel member 3: No response

Panel member 4: No response

Panel member 5: No response

Panel member 6: Significant mprecision exists in the specifications, for example, in the numerator, the term 'up-to-date' is vague and should be replace with more precise language, as is 'based on age and recommendations'. More precision will make the statements longer but they will less open to interpretation. This seems like the bailiwick of the standing committee and they weigh in on this issue.

In addition, the composite easure appears to achieve a modest level of dispersion of health plans, with the largest IQRs observed among Medicare plans. The standing committee should evaluate whether these differences are meaningful from a clinical perspective.

The standing committee will need to weigh in on the alignment between the specifications and the most recent guidance for these immunizations.

Questions for the Committee regarding reliability:

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

Questions for the Committee regarding validity:

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

Questions for the Committee regarding composite construction:

- Do you have any concerns regarding the composite construction approach (e.g., do the component measures fit the quality construct and add value to the overall composite? Are the aggregation and weighting rules consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible?)?
- The Scientific Methods Panel is satisfied with the composite construction. Does the Committee think there is a need to discuss and/or vote on the composite construction approach?

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

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

Preliminary rating for composite construction: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

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

2a1. Reliability-Specifications

- I have concerns regarding the specifications and echo the comments from Panel members #1 and 6.
- No concerned regarding consistent implementation
- I appreciate Panel Member 1's stated concerns about timing and "assingment" to a health plan.
- Reliability is appropriate.
- As noted by Scientific Methods Panel (SMP) member #6, the term 'up-to-date' is vague and should be replace with more precise language, as is 'based on age and recommendations'. Even if this change were made, though, it is hard to know how plans should or would deal with vaccines that are recommended, say, every 10 years when individuals were not covered by the plan for 10 years.
- Specs from all lines of businesses (Medicaid/Commercial/Medicare) need to be reliable, in addition to reliability on different provider group categories (IPAs, medical group/foundation/staff model, integrated health systems, ACOs).
- The reliability is sound for what is being measured. However, as expressed in other comments, there is concern about the reationale for the measure itself. Social risk factors have not been addressed, and with vaccinations, social risk factors are essential to consider. This is a new HEDIS measure, and no data were provided to demonstrate reliability. Another reliability issue not addressed is the impact on consistent collection of these data when services are provided by pharmacy clinics or in other less traditional settings. While reliability may not be a problem, it would have been helpful to see this issue addressed.

2a2. Reliability – Testing, any concerns

- I think this should be discussed and reviewed by the committee and the methods panel.
- No
- Methods used make sense - no concerns
- I agree with high rating.
- Testing was done appropriately at the measure score level, and with appropriate data. However, as noted above, the vague specification of the term 'up-to-date' is a major concern. The beta-binomial method is appropriate for the individual vaccines, but since the composite is not a binomial proportion as noted above, it is not strictly appropriate for the composite measure. However, since the results are so strong, I don't see a problem. Some of the SMP members questioned the near-perfect reliability results, but given median plan sizes ranging from 11,648 to 80,330 and the broad range of coverage rates, this is not surprising. Overall rating of reliability: Low (specifications are NOT precise, unambiguous, and complete)
- Given my previous comment, is there INSUFFICIENT rating for reliability or simply MODERATE? I don't have enough information to make a determination at this time.
- This is a new HEDIS measure (implemented 2018) and no data were provided on reliability across health plans.

2b1. Validity -Testing, any concerns

- Based on the comments of the methods panel validity needs to be discussed
- No
- No concerns
- Validity is appropriate.
- No. Face validity approach appropriate.
- No

- The measure sponsor states that this is a population health measure. However, because the approach to used for assuring vaccination related to each of the four (or three, depending on the age group), may be different. The impact on the population also is different for each vaccine. Having an composite measure does not enable a plan to address population health issues as they relate to specific types of vaccines, which may require different types of stratgegies to assure appropriate vacinnations are administered and received.

2b4-7; 2b2-3 Threats to Validity

- They evaluated the different components and are not using an all or nothing approach for this measure. This may be an 'in the weeds' discussion, but I think that we may want to dig into the validity/reliablity of this measure.
- IQRs appear to indicate fair opportunity for improvement. No threat from missing data
- No concerns
- No concerns noted.
- Data used for testing are appropriate. Missing data not a problem.
- N/A
- Some issues to consider: 1) comparing this composite measure across plans may suggest similar quality, but the types of vaccines administered maybe different. This may result in a greater impact on health from one plan to another. 2) it is unclear why hospice patient data are excluded; 3) Vaccinations received in less traditional settings may impact data collection.
- It looks like the patients with vaccine reactions are included in the numerator and are not excluded from the denominator. This could be challenging given the longitudinal nature of these measures. It would be helpful to see the Measure algorithm for the denominator and numerator definitions.
- Exclusions are consistent. No inappropriate exclusions are evident. Would appreciate the clarification noted by one panelist on continuous plan enrollment.
- Again, I think we should discuss social risk adjustment and stratification for this measure, particularly given our prevention and pop health focus
- No concerns noted.
- Exclusions not a problem.
- I didn't see any risk adjustment noted.
- As mentioned above, it is unclear why hospice patients are excluded. Regarding social risk factor variables, the unvaccinated population (whether or not covered by an included health plan), can impact the health of those who have been vaccinated. The measure sponsor does not address the impact of the unvaccinated population on the vaccinated population.

2c. Composite Performance Measure

- The overall goal of the is measure is of value, but I'm not sure if the construct of the composite achieves the goal.
- Yes
- Strong Chronbach alpha stats
- No concerns noted.
- Correlations between coverage fractions for vaccines in the adult composite are high, and higher than with vaccines for the pre-natal and adolescent populations, as one would expect. This supports the concept of the composite. Overall rating of validity: High
- It appears the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

- This composite measure may have unanticpated consequences. The most significat is variation of types of vaccinations received being masked by using the composite measure. This may mask health plan quality as well as individual and social impact.

Criterion 3. Feasibility

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

ALL data elements are in defined fields in a combination of electronic sources. Data Elements generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score), Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry).

The measure developer notes that an independent audit of all HEDIS collection and reporting processes is conducted, as well as an audit of the data which are manipulated by those processes, in order to verify that HEDIS specifications are met. In addition to the HEDIS audit, the measure developer provides a system to allow "real-time" feedback from measure users.

Questions for the Committee:

- None

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

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

- Should be routinely generated during care, however patient churn may impact this mesure.
- No feasibility concerns
- No concerns
- I agree with moderate rating.
- The individual vaccine measures are currently in use in HEDIS with no apparent problems. Rating: High
- Rating of feasibility is moderate given the necessity of HEDIS collection and reporting processes (e.g., health plan "chart chase")
- This is a new HEDIS measure, and feasibility has not yet been fully demonstrated.

Criterion 4: Usability and Use

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

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

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

The developer provided a table of current and planned use

Specific Plan for Use	Current Use (for current use provide URL)
Public Reporting Quality Improvement (external benchmarking to organizations)	Quality Improvement (Internal to the specific organization) Healthcare Effectiveness Data and Information Set (HEDIS) https://www.ncqa.org/hedis/using-hedis-measures/

Publicly reported? ☒ **Yes** ☐ **No**

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

OR

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

Accountability program details

The HEDIS set is one of health care's most widely used performance improvement tools and are used by health plans and other various levels of the health care system for quality improvement initiatives. This measure was a new HEDIS measure in 2018. NCQA's standard process is to evaluate data for all new measures prior to use for public reporting, benchmarking and/or other programs.

The developer notes that the plan for implementation includes working with a multi-stakeholder advisory panels to assess the number of plans that have shown they can report the measure; whether measure results match what we expect; whether results seem indicative of true performance; and whether performance indicates an opportunity for improvement for the industry overall.

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

The measure developer noted that during a recent public comment posting which was held during the development process, measured entities supported the new measure and found it to be relevant and clearly specified.

Additional Feedback: This measure has been deemed a priority measure by NCQA and other entities such as the Centers for Disease Control and Prevention, the federal National Vaccine Program Office and the American Immunization Registry Association.

***Questions for the Committee**:*

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

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

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

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

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

Improvement results New measure, therefore no data on improvement over time. Adoption of this measure has the potential to improve the immunization rates for adults.

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 None

Potential harms None identified

Additional Feedback:

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**

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

4a1. Use - Accountability and Transparency

- New Hedis Measure in 2018. Potentially very useful as surveillance of adult immunization rates can be challenging.
- New measure to NCQA in 2018, still undergoing panel review and pending NCQA internal assessment. Feedback obtained through comment period according to submission, and has informed modifications to specs to better align with ACIP guidelines.
- No concerns.
- I agree with passing rating.
- The individual vaccine measures are currently used effectively in HEDIS. Rating: High
- Feedback on measure is possible and considered by measure developer.
- It is not clear how those being measured will use a single composite measure.

4b1. Usability – Improvement

- New measure
- Improvement potential is stated, but not elaborated upon in submission..
- I do not perceive any unintended harms
- I agree with moderate rating.
- NA

- Rating for usability is high.
- It is unclear how this composite measure can/will be used by health plans to improve performance. For instance, for one vaccine, availability may be a factor. For another vaccine, acceptance by the patient may be a factor. A composite measure will not be useful for addressing these types of issues.

Criterion 5: Related and Competing Measures

Related or competing measures

0041 : Preventive Care and Screening: Influenza Immunization
0043 : Pneumococcal Vaccination Status for Older Adults (PNU)
0431 : INFLUENZA VACCINATION COVERAGE AMONG HEALTHCARE PERSONNEL
0680 : Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)
0681 : Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)
0682 : Percent of Residents or Patients Assessed and Appropriately Given the Pneumococcal Vaccine (Short-Stay)
0683 : Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long-Stay)
1653 : Pneumococcal Immunization
1659 : Influenza Immunization

Harmonization

The measure developer notes that the specifications are not harmonized. This is a population-based measure that assesses vaccines provided in the outpatient setting at the health plan level. Most of the other NQF-endorsed vaccination measures focus only on either pneumococcal or influenza vaccination. These measures specifically apply to inpatient populations, residents in long-term care/skilled nursing facilities or healthcare personnel or are specified at the provider-level. Moreover, our proposed measure is specified to use electronic clinical data, while other related measures (e.g., NQF 0039) are specified to use survey data in which patients must recall whether they had received a vaccine.

The developer also mentions that the measure proposed measure is more specific than several of the other adult vaccination measures because it assesses whether health plan members received the appropriate type and doses of vaccines at the right time according to clinical guidelines.

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

- It is a composite measure and some components are similar to this measure, but this measure is aiming to provide a composite estimate of adult immunization rates (persumabely stratified by age).
- Submission identifies many competing/related measures. No additional harmonization appears to be needed; measure is more specific in terms of populations, collection method, and target audience (health plans).
- No concerns noted.
- NA
- The measure developer notes that the specifications are not harmonized. This is a population-based measure that assesses vaccines provided in the outpatient setting at the health plan level--different from the individual vaccine measures in circulation.

- There are related measures and this composite measure does not seem to be in conflict with those measures.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: Month/Day/Year

- **Of the XXX NQF members who have submitted a support/non-support choice:**
  - XX support the measure
  - YY do not support the measure

Brief Measure Information

NQF #: 3483

Corresponding Measures:

De.2. Measure Title: Adult Immunization Status

Co.1.1. Measure Steward: National Committee for Quality Assurance

De.3. Brief Description of Measure: Percentage of adults 19 years of age and older who are up-to-date on recommended routine vaccines for influenza, tetanus and diphtheria (Td) or tetanus, diphtheria and acellular pertussis (Tdap), zoster and pneumococcal.

1b.1. Developer Rationale: See question 1c.3.

S.4. Numerator Statement: Adults who are up-to-date on influenza, Td or Tdap, herpes zoster and pneumococcal vaccinations based on age and recommendations.

S.6. Denominator Statement: Adults ages 19 years and older.

S.8. Denominator Exclusions: Adults who received chemotherapy, had a bone marrow transplant or were in hospice during the measurement year or those with a history of immunocompromising conditions.

De.1. Measure Type: Composite

S.17. Data Source: Claims, Electronic Health Data, Electronic Health Records, Enrollment Data, Management Data, Other, Registry Data

S.20. Level of Analysis: Health Plan

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

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? N/A

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

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

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

AIS_Evidence_Form_.docx

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

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

No

1a. Evidence (subcriterion 1a)

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 3483
Measure Title: Adult Immunization Status
IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title
Date of Submission: 11/15/2019

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

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

1a. Evidence to Support the Measure Focus

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

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

Notes

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

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

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

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

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

☐ Outcome: Click here to name the health outcome

☐Patient-reported outcome (PRO): Click here to name the PRO

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

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☐ Process: Click here to name what is being measured

☐ Appropriate use measure: Click here to name what is being measured

☐ Structure: Click here to name the structure

☒ Composite: Percentage of adults 19 years of age and older who are up-to-date on recommended routine vaccines for influenza, tetanus and diphtheria (Td) or tetanus, diphtheria and acellular pertussis (Tdap), zoster and pneumococcal.

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

Adults age 19 years or older >> routine vaccines for influenza, Td/Tdap, herpes zoster and pneumococcal are given based on recommendations for age, timing and dosing >> prevent disease >> improved length and/or quality of life

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

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

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

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

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

☒ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

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

☐ Other

Table 1. Influenza Vaccine Recommendation

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• Title: Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices – United States, 2019-20 Influenza Season • Author: Lisa A. Grohskopf, Influenza Division, National Center for Immunization and Respiratory Diseases, CDC, et al. • Date: August 23, 2019 • Citation: Grohskopf LA, Alyanak E, Broder KR, Walter EB, Fry AM, Jernigan DB. Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices — United States, 2019–20 Influenza Season. MMWR Recomm Rep 2019;68(No. RR-3):1–21. DOI: http://dx.doi.org/10.15585/mmwr.rr6803a1 • URL: https://www.cdc.gov/mmwr/volumes/68/rr/pdfs/rr6803-H.pdf
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"Routine annual influenza vaccination is recommended for all persons aged ≥6 months who do not have contraindications. Balancing considerations regarding the unpredictability of timing of onset of the influenza season and concerns that vaccine-induced immunity might wane over the course of a season, it is recommended that vaccination should be offered by the end of October. Vaccination should continue to be offered as long as influenza viruses are circulating and unexpired vaccine is available. A licensed influenza vaccine that is appropriate for the recipient's age and health status should be used." Table 1 in guidelines: for the 2019-2020 influenza season, inactivated influenza vaccine (IIV4) are recommended for all adults, IIV3 are recommended for adults ages 65 and older, recombinant influenza vaccine (RIV4) are recommended for adults ages 18 and older and live attenuated influenza vaccine (LAIV) are recommended for people ages 2 through 49 years.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	ACIP did not provide a grade for the evidence underlying this recommendation. ACIP conducts a thorough review of peer-reviewed evidence on vaccine safety and effectiveness, discusses recommendations with professional organizations, and holds regular meetings for experts to vote on proposed recommendations.
Provide all other grades and definitions from the evidence grading system	N/A
Grade assigned to the **recommendation** with definition of the grade	ACIP did not provide a grade for this recommendation. CDC vaccine recommendations are developed using an explicit evidence-based method based on the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach. Key factors considered in development of recommendations include balance of benefits and harms, type or quality of evidence, values and

	preferences of the people affected, and health economic analyses. ACIP discusses recommendations with professional organizations and holds regular meetings for experts to vote on proposed recommendations.
Provide all other grades and definitions from the recommendation grading system	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The ACIP Influenza Work Group reviewed available data and evidence from 1979 to 2018 on immunogenicity, efficiency, effectiveness and safety of influenza vaccines. They also convene twice monthly to review "influenza surveillance, vaccine effectiveness and safety, vaccine coverage, program feasibility, cost-effectiveness, and vaccine supply" in order to provide annual recommendations for the use of influenza vaccines for the prevention and control of influenza. In total, the Work Group reviewed approximately 285 studies on the immunogenicity, efficacy and effectiveness of IIV, RIV and LAIV and 120 studies on influenza vaccine safety. This review included approximately 123 studies specifically assessing immunogenicity, efficacy and effectiveness of influenza vaccine for adults; and 36 studies assessing influenza vaccine safety for adults. These studies consist of randomized control trials, case control studies and observational studies, among others.
Estimates of benefit and consistency across studies	The Work Group reviewed vaccine safety—compared to the morbidity and mortality risks associated with influenza—and concluded that all persons aged ≥6 months without contraindications are recommended to receive routine influenza vaccinations. Vaccine type, product, and dose recommendations are based on age and pregnancy status. This includes vaccination of pregnant women with a licensed, recommended, and age-appropriate IIV or RIV4.
What harms were identified?	The Work Group examined the risk of adverse events for each vaccine type and concluded that the influenza vaccine is safe for routine administration, including administration to pregnant women. ACIP identified severe allergic reactions as a contraindication to all types of the influenza vaccine. LAIVs are contraindicated for pregnant women. For adults, the most common safety complaints were injection site pain (that did not interfere with daily activities) and systemic reactions, such as myalgia, headaches, and fatigue. Serious adverse events were uncommon. No specific safety concerns were identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	There have been no studies published since the guideline that would significantly affect the findings.

Table 2. Td/Tdap Vaccine Recommendation

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• Title: Prevention of Pertussis, Tetanus, and Diphtheria with Vaccines in the United States: Recommendations of the Advisory Committee on Immunization Practices (ACIP) • Author: Jennifer L. Liang, National Center for Immunization and Respiratory Diseases, CDC, et al. • Date: April 27, 2018

	• Citation: Liang J, Tiwari T, Moro P et al. Prevention of Pertussis, Tetanus, and Diphtheria with Vaccines in the United States: Recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep 2018;67(No. 2):1-44. • URL: https://www.cdc.gov/mmwr/volumes/67/rr/pdfs/rr6702a1-H.pdf
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"ACIP recommends routine vaccination for tetanus, diphtheria, and pertussis. Infants and young children are recommended to receive a 5-dose series of diphtheria and tetanus toxoids and acellular pertussis (DTaP) vaccines, with one adolescent booster dose of tetanus toxoid, reduced diphtheria toxoid, and acellular pertussis (Tdap) vaccine. Adults who have never received Tdap also are recommended to receive a booster dose of Tdap. After receipt of Tdap, adolescents and adults are recommended to receive a booster tetanus and diphtheria toxoids (Td) vaccine every 10 years to assure ongoing protection against tetanus and diphtheria."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	ACIP did not provide a grade for the evidence underlying this recommendation. ACIP conducts a thorough review of peer-reviewed evidence on vaccine safety and effectiveness, discusses recommendations with professional organizations, and holds regular meetings for experts to vote on proposed recommendations.
Provide all other grades and definitions from the evidence grading system	N/A
Grade assigned to the **recommendation** with definition of the grade	ACIP did not provide a grade for this recommendation. CDC vaccine recommendations are developed using an explicit evidence-based method based on the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach. Key factors considered in development of recommendations include balance of benefits and harms, type or quality of evidence, values and preferences of the people affected, and health economic analyses. ACIP discusses recommendations with professional organizations and holds regular meetings for experts to vote on proposed recommendations.
Provide all other grades and definitions from the recommendation grading system	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The ACIP Pertussis Vaccines Work Group reviewed available published and unpublished data and evidence from 2004 to 2017, covering topics such as tetanus, diphtheria and pertussis disease epidemiology in the United States, decision analyses, cost-effectiveness, programmatic considerations, vaccine immunogenicity, vaccine safety, and postlicensure vaccine effectiveness. In total, they reviewed 110 studies consisting of randomized control trials and other types of studies on Td and Tdap vaccination.
Estimates of benefit and consistency across studies	The Work Group reviewed vaccine safety—compared to the morbidity and mortality risks associated with pertussis, tetanus and diphtheria—and concluded that "All persons are recommended to receive routine pertussis, tetanus, and diphtheria vaccination. Vaccine type, product, number of doses and booster dose recommendations are based on age

	and pregnancy status." This includes the vaccination of adults with a single booster tetanus and diphtheria toxoids (Td) vaccine every 10 years.
What harms were identified?	The Work Group examined the risk of adverse events for each vaccine component and concluded that the Td and Tdap vaccines are safe for administration to adolescents and adults, including pregnant women. Any adverse reactions that were observed were limited to minor local reactions, including pain, erythema and swelling; no serious adverse events have been observed. Receipt of Tdap during pregnancy has not been found to be associated with an increased risk for frequency of major malformations, stillbirth, preterm birth, small for gestational age, or hypertensive disorders. ACIP identified two contraindications to the Tdap vaccine: severe allergic reactions or encephalopathy associated with administration of a prior dose of a DTP, DTaP, or Tdap vaccine.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	There have been no studies published since the guideline that would significantly affect the findings.

Table 3. Herpes Zoster Vaccination

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	**ACIP 2018 Guidelines for Recombinant Zoster Vaccine:** • Title: Recommendations of the Advisory Committee on Immunization Practices for Use of Herpes Zoster Vaccines • Author: Kathleen Dooling, Centers for Disease Control and Prevention, et al. • Date: January 2018 • Citation: Dooling KL, Guo A, Patel M, et al. Recommendations of the Advisory Committee on Immunization Practices for Use of Herpes Zoster Vaccines. MMWR Morb Mortal Wkly Rep 2018;67:103–108. DOI: http://dx.doi.org/10.15585/mmwr.mm6703a5 • URL: https://www.cdc.gov/mmwr/volumes/67/wr/pdfs/mm6703a5-H.pdf **ACIP 2008 Guidelines for Zoster Vaccine Live:** • Title: Prevention of Herpes Zoster: Recommendations of the Advisory Committee on Immunization Practices (ACIP) • Author: Rafael Harpaz, Centers for Disease Control and Prevention, et al. • Date: June 6, 2008 • Citation: Harpaz R, Ortega-Sanchez IR, Seward JF. Prevention of Herpes Zoster: Recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Early Release 2008;57[November 2019]:1-2. • URL: https://www.cdc.gov/mmwr/PDF/rr/rr5705.pdf

Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	ACIP 2018 Guidelines for Recombinant Zoster Vaccine: "Recombinant zoster vaccine (RZV) is recommended for the prevention of herpes zoster and related complications for immunocompetent adults aged ≥50 years. RZV is recommended for the prevention of herpes zoster and related complications for immunocompetent adults who previously received zoster vaccine live (ZVL). RZV is preferred over ZVL for the prevention of herpes zoster and related complications. These recommendations serve as a supplement to the existing recommendations for the use of ZVL in immunocompetent adults aged ≥60 years." ACIP 2008 Guidelines for Zoster Vaccine Live: "Licensed zoster vaccine is a lyophilized preparation of a live, attenuated strain of VZV, the same strain used in the varicella vaccines. Zoster vaccine is recommended for all persons aged ≥60 years who have no contraindications, including persons who report a previous episode of zoster or who have chronic medical conditions. The vaccine should be offered at the patient's first clinical encounter with his or her health-care provider."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	1: Randomized controlled trials (RCTs), or overwhelming evidence from observational studies.
Provide all other grades and definitions from the evidence grading system	2: RCTs with important limitations, or exceptionally strong evidence from observational studies. 3: Observational studies, or RCTs with notable limitations. 4: Clinical experience and observations, observational studies with important limitations, or RCTs with several major limitations.
Grade assigned to the **recommendation** with definition of the grade	ACIP did not provide a grade for this recommendation. CDC vaccine recommendations are developed using an explicit evidence-based method based on the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach. Key factors considered in development of recommendations include balance of benefits and harms, type or quality of evidence, values and preferences of the people affected, and health economic analyses. ACIP discusses recommendations with professional organizations and holds regular meetings for experts to vote on proposed recommendations.
Provide all other grades and definitions from the recommendation grading system	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The ACIP Herpes Zoster Vaccines Work Group evaluated studies published from 2015-2017 on the efficacy, cost-effectiveness, and safety of both RZV and ZVL. Their review included 10 studies of RZV, including seven randomized control trials (RCTs). They reviewed 40 studies of ZVL, including 16 high-quality RCTs, 13 RCTs with noted limitations, 10 cohort studies, and 1 case control study.
Estimates of benefit and consistency across studies	The Work Group reviewed vaccine safety—compared to the morbidity and mortality risks associated with herpes zoster—and concluded

	that "with high efficacy among adults aged ≥50 years, and modest waning of protection over 4 years following vaccination, RZV has the potential to prevent substantial herpes zoster disease burden. Vaccinating adults starting at age 50 will prevent disease incidence in midlife, and the vaccine will likely continue to provide substantial protection beyond 4 years as recipients age."
What harms were identified?	The Work Group examined the risk of adverse events for each vaccine component and concluded that RZV and ZVL are safe for administration to adults. Any adverse reactions that were observed were limited to minor local reactions, including pain, myalgia and fatigue. Overall, serious adverse events occurred at similar rates in vaccinated and placebo groups. ACIP identified severe allergic reactions as a contraindication to both RZV and ZVL vaccines. ACIP identified pregnancy and immunocompromising conditions as contraindications for ZVL.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	There have been no studies published since the guideline that would significantly affect the findings.

Table 4. Pneumococcal Vaccination

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• Title: Intervals Between PCV13 and PPSV23 Vaccines: Recommendations of the Advisory Committee on Immunization Practices (ACIP) • Author: Miwako Kobayashi, Centers for Disease Control and Prevention, et al. • Date: September 2015 • Citation: Kobayashi M, Bennett N, Gierke R, Almendares O et al. Intervals Between PCV13 and PPSV23 Vaccines: Recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep 2015;64(No. 34):944-947. • URL: https://www.cdc.gov/mmwr/pdf/wk/mm6434.pdf#page=16
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"For immunocompetent adults aged ≥65 years who have not previously received pneumococcal vaccine: a dose of PPSV23 should be given ≥1 year following a dose of PCV13. The two vaccines should not be co-administered. If a dose of PPSV23 is inadvertently given earlier than the recommended interval, the dose need not be repeated. For those for who previously received PPSV23 when aged <65 years and for whom an additional dose PPSV23 is indicated when aged ≥65 years, this subsequent PPSV23 dose should be given ≥1 year after PCV13 and ≥5 years after the most recent dose of PPSV23."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	2: RCTs with important limitations, or exceptionally strong evidence from observational studies.
Provide all other grades and definitions from the evidence grading system	1: Randomized controlled trials (RCTs), or overwhelming evidence from observational studies. 3: Observational studies, or RCTs with notable limitations.

	4: Clinical experience and observations, observational studies with important limitations, or RCTs with several major limitations.
Grade assigned to the **recommendation** with definition of the grade	ACIP did not provide a grade for this recommendation. CDC vaccine recommendations are developed using an explicit evidence-based method based on the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach. Key factors considered in development of recommendations include balance of benefits and harms, type or quality of evidence, values and preferences of the people affected, and health economic analyses. ACIP discusses recommendations with professional organizations and holds regular meetings for experts to vote on proposed recommendations.
Provide all other grades and definitions from the recommendation grading system	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The ACIP Pneumococcal Work Group evaluated studies published from 2004-2014 on benefits, harms, values and preferences, and cost-effectiveness on PCV13 for routine use among adults aged 65 years and older. Their review included 6 randomized control trials (RCTs) on immunogenicity, 3 RCTs on serious and systemic adverse events and 2 other RCTs that they determined were of high and moderate quality.
Estimates of benefit and consistency across studies	The Work Group reviewed vaccine safety—compared to the morbidity and mortality risks associated with pneumonia—and concluded that "benefits outweigh harms." Vaccine type, number of doses and interval between doses are based on age, prior vaccination history and presence of specific medical conditions.
What harms were identified?	The Work Group examined the risk of adverse events for each vaccine component and concluded that PCV13 and PPSV23 are safe for administration to adults. Any adverse reactions that were observed were limited to minor local reactions, including pain, myalgia and fatigue. Overall, serious adverse events occurred at similar rates in vaccinated and placebo groups. ACIP identified severe allergic reactions as a contraindication to both PCV13 and PPSV23 vaccines.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	There have been no studies published since the guideline that would significantly affect the findings.

1a.4 OTHER SOURCE OF EVIDENCE

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

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

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

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

1b. Performance Gap

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

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

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

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

See question 1c.3.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

The following data are extracted from HEDIS data collection reflecting the most recent year of measurement (2018) for this measure. Performance data is summarized at the health plan level and summarized by mean performance and performance at 10th, 25th, 50th, 75th, and 90th percentile. We also calculated the interquartile range (IQR), which can be interpreted as the difference between the 25th?and 75th?percentile. Data is stratified by product line (i.e. commercial, Medicaid and Medicare).

The following data demonstrate the variation in the rate of adult immunization across health plans. For the composite rate, there was an 11 point difference between plans in the 25th percentile and plans in the 75th percentile for Medicare plans, and 7 and 8 points for commercial and Medicaid plans, respectively. These gaps in performance underscore the opportunity for improvement.

Adult Immunization Status: Composite

Commercial, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

21.2 | 13.5 | 16.0| 18.1 | 22.9 | 30.2 | 6.9

Medicaid, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

14.0 | 2.7 | 10.0 | 13.7 | 17.5 | 20.7 | 7.5

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

19.5 | 6.6 | 9.5 | 14.4 | 20.6 | 43.8 | 11.1

Adult Immunization Status: Influenza

Commercial, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

18.7 | 11.4 | 14.9 | 18.1 | 20.6 | 26.4 | 5.7

Medicaid, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

11.6 | 2.8 | 7.8 | 11.7 | 15.3 | 20.7 | 26.0 | 7.5

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

18.3 | 5.3 | 8.3 | 12.5 | 21.7 | 30.1 | 13.4

Adult Immunization Status: Td/Tdap

Commercial, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

29.4 | 18.6 | 20.9 | 25.2 | 30.7 | 46.5 | 9.8

Medicaid, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

20.9 | 4.9 | 14.1 | 21.1 | 25.0 | 34.4 | 10.9

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

26.5 | 9.5 | 14.8 | 20.7 | 28.7 | 56.4 | 13.9

Adult Immunization Status: Herpes Zoster

Commercial, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

6.1 | 2.7 | 4.1 | 5.0 | 6.5 | 9.8 | 2.4

Medicaid, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

1.6 | 0.0 | 0.4 | 0.6 | 1.4 | 5.3 | 1.0

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

12.9 | 0.4 | 0.9 | 5.3 | 14.5 | 39.5 | 13.6

Adult Immunization Status: Pneumococcal

Medicare, 2018

Mean | 10th | 25th | 50th | 75th | 90th | Interquartile Range

20.3 | 5.4 | 8.1 | 10.8 | 22.5 | 55.5 | 14.4

The HEDIS performance data reflect the most recent year of measurement for this measure. Below is a description of the number of health plans that reported this measure and the median eligible population for the measure (stratified by commercial, Medicaid and Medicare).

Commercial, 2018

N Plans | Median Eligible Population

71 | 80,330

Medicaid, 2018

N Plans | Median Eligible Population

21 | 36,250

Medicare, 2018

N Plans | Median Eligible Population

44 | 11,648

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

Estimates of national vaccination coverage are available through the National Health Interview Survey (NHIS), in which a sample of adult's self-report receipt of vaccines. In 2015, 45 percent of adults 19 and older reported that they received the influenza vaccine during the 2014–2015 flu season, well below the Healthy People 2020 target of 70 percent (Williams et al. 2017). 64 percent of adults 65 and older reported having ever received the PPSV23 vaccine and/or the PCV13 vaccine, which is below the Healthy People 2020 target of 90 percent (Williams et al. 2017). Although there is no corresponding Healthy People 2020 goal for routine Tdap or Td vaccination among adults, only 23 percent of adults 19 and older responding to the 2015 NHIS reported receiving the Tdap vaccine within the past 10 years, and 62 percent reported receiving any tetanus toxoid-containing vaccination during the past 10 years (Williams et al. 2017).In 2015, 31 percent of adults ages 60 and older reported ever receiving the herpes zoster vaccine (Williams et al. 2017). Although zoster vaccination coverage meets the Healthy People 2020 target of 30 percent coverage, 70 percent of adults are not receiving this recommended vaccination due to factors that include vaccine shortages shortly after licensure (Hurley et al. 2010), complications in storing the vaccine and cost to consumers (Hurley et al. 2010). Barriers to adult vaccination in general include provider and patient lack of knowledge and awareness of the importance of vaccines, missed opportunities for vaccination and operational and systemic barriers (e.g., cost, lack of access to immunization records) (Ventola 2016; Tan 2015). Having health insurance coverage and a usual place for health care is associated with higher vaccination coverage (Williams et al. 2017).

Hurley, L.P., M.C. Lindley, R. Harpaz, S. Stokley, M.F. Daley, L.A. Crane, et al. 2010. "Barriers to the Use of Herpes Zoster Vaccine." Ann Intern Med. 152:555–60. doi: 10.7326/0003-4819-152-9-201005040-00005.

Ventola, C.L. 2016. "Immunization in the United States: Recommendations, Barriers, and Measures to Improve Compliance: Part 2: Adult Vaccinations." Pharmacy and Therapeutics. 41(8), 492–506.

Williams W.W., P. Lu, A. O'Halloran, et al. 2017. "Surveillance of Vaccination Coverage among Adult Populations—United States, 2015." MMWR Surveill Summ. 66(No. SS-11):1–28. DOI: http://dx.doi.org/10.15585/mmwr.ss6611a1.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

HEDIS data are stratified by type of insurance (e.g. Commercial, Medicaid, Medicare). While not specified in the measure, this measure can also be stratified by demographic variables, such as race/ethnicity or socioeconomic status, in order to assess the presence of health care disparities, if the data are available to a plan. The HEDIS Race/Ethnicity Diversity of Membership and the Language Diversity of Membership measures were designed to promote standardized methods for collecting these data and follow Office of Management and Budget and Institute of Medicine guidelines for collecting and categorizing race/ethnicity and language data. In addition, NCQA's Multicultural Health Care Distinction Program outlines standards for collecting, storing, and using race/ethnicity and language data to assess health care disparities.

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

There are racial and ethnic disparities in adult vaccination coverage. The 2015 NHIS found that White adults were more likely to have received the influenza vaccine (47 percent) than Blacks (37 percent) and Hispanics (33 percent) (Williams et al. 2017). Tdap and Td booster vaccination coverage was higher for White adults 19

and older than Black, Hispanic and Asian adults (Williams et al. 2017). Similarly, pneumococcal vaccination coverage and zoster vaccination coverage was higher for White older adults than for Black, Hispanic and Asian older adults (Williams et al. 2017). Racial and ethnic disparities in pneumococcal vaccination and herpes zoster vaccination coverage widened from 2014–2015 due to increases in vaccination coverage for older White adults (Williams et al. 2017). Vaccination coverage also varies by age for influenza and Tdap/Td. In the 2015 NHIS survey, older adults were more likely to report receiving the influenza vaccine; 32 percent of adults 19–49 reported receiving the flu vaccine, compared with 49 percent of adults 50–64 and 74 percent of adults 65 and older (Williams et al. 2017); however, adults 65 and older were less likely to report having received the Td or Tdap vaccine than adults 19–64 (Williams et al. 2017).

Williams W.W., P. Lu, A. O'Halloran, et al. 2017. "Surveillance of Vaccination Coverage among Adult Populations—United States, 2015." MMWR Surveill Summ. 66(No. SS-11):1–28. DOI: http://dx.doi.org/10.15585/mmwr.ss6611a1.

1c. Composite Quality Construct and Rationale

1c.1. A composite performance measure is a combination of two or more component measures, each of which individually reflects quality of care, into a single performance measure with a single score.

For purposes of NQF measure submission, evaluation, and endorsement, the following will be considered composites:

- Measures with two or more individual performance measure scores combined into one score for an accountable entity.
- Measures with two or more individual component measures assessed separately for each patient and then aggregated into one score for an accountable entity:
    - all-or-none measures (e.g., all essential care processes received, or outcomes experienced, by each patient);

1c.1. Please identify the composite measure construction: two or more individual performance measure scores combined into one score

1c.2. Describe the quality construct, including:

- the overall area of quality
- included component measures and
- the relationship of the component measures to the overall composite and to each other.

This measure assesses the provision of critical routine immunizations for adults 19 and older per clinical guidelines. The intent of the measure is to improve primary prevention of vaccine-preventable diseases including influenza, tetanus, diphtheria, pertussis, herpes zoster and pneumococcal disease. The measure calculates a rate for each specific vaccine type and a composite rate:

Influenza rate

Td/Tdap rate

Herpes Zoster rate

Pneumococcal rate

Composite rate: this rate is calculated by summing the number of immunizations that were administered to all adults across the plan's enrolled population (numerator) and dividing this sum by the total recommended number of immunizations, per clinical guidelines for the age group (denominator).

1c.3. Describe the rationale for constructing a composite measure, including how the composite provides a distinctive or additive value over the component measures individually.

The composite rate provides an overview of how many routine adult vaccines were received out of the total that were recommended for a health plan member population. The individual vaccine component rates are included to provide information on which types of adult vaccinations are being provided to members as recommended.

1c.4. Describe how the aggregation and weighting of the component measures are consistent with the stated quality construct and rationale.

The components are weighted equally in this composite to assess compliance with immunization guidelines for adults. We constructed the composite to assess the percentage of total vaccines received across the population because this approach more easily accounts for the different requirements across age groups and is more actionable than an all-or-nothing composite.

2.3 For maintenance of endorsement

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

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

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

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number
Composite Measure Title: Adult Immunization Status
Date of Submission: 8/1/2019
Composite Construction:
☒Two or more individual performance measure scores combined into one score
☐ All-or-none measures (e.g., all essential care processes received or outcomes experienced by each patient)

Instructions: Please contact NQF staff before you begin.

- If a component measure is submitted as an individual performance measure, the non-composite measure testing form must also be completed and attached to the individual measure submission.
- 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.
- **Sections 1, 2a2, 2b1, 2b2, and 2b4 must be completed.**
- **For composites with outcome and resource use measures**, section **2b3** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b5** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b1-2b6) and composites (2c) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 25 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment. and the 2017 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.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **instrument-based measures** (including **PRO-PMs**) **and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b1. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **instrument based measures (including PRO-PMs) and composite performance measures**, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidenceand are of sufficient frequency to warrant inclusion in the specifications of the measure; [12]
AND
If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). [13]

2b3. For outcome measures and other measures when indicated (e.g., resource use):
- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; [14],[15] and has demonstrated adequate discrimination and calibration

OR
- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [16] **differences in performance**;
OR
there is evidence of overall less-than-optimal performance.

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

2b6. Analyses identify the extent and distribution of **missing data** (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

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)*

Notes
10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal

consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).
11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.
12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.
13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.
14. Risk factors that influence outcomes should not be specified as exclusions.
15. With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE
Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

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

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.17*)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☒ claims	☒ claims
☒ registry	☒ registry
☒ abstracted from electronic health record	☒ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other:	☐ other:

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

1.3. What are the dates of the data used in testing? 01/01/2016-12/31/2018

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

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.20*)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☒ health plan	☒ health plan
☐ other: Click here to describe	☐ other: Click here to describe

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

This measure assesses whether adults enrolled in commercial, Medicaid and Medicare health plans are up-to-date on routine vaccines per clinical guidelines. The measure calculates a rate for each specific vaccine type and a composite rate.

Vaccine-Specific Indicators

Indicators	Ages Reported for Commercial & Medicaid Health Plans	Ages Reported for Medicare Health Plans
Influenza rate: Percentage of members who received an influenza vaccine on or between July 1 of the year prior to the measurement period and June 30 of the measurement period.	19-65	66 and older
Td/Tdap rate: Percentage of members who received a Td or Tdap vaccine on or between January 1 of the nine years prior to the measurement period and December 31 of the measurement period.	19-65	66 and older
Herpes Zoster rate: Percentage of members who received one dose of the herpes zoster live vaccine or two doses of the herpes zoster recombinant vaccine at least 28 days apart on or after the member's 50th birthday.	50-65	66 and older
Pneumococcal rate: Percentage of members who were administered the 13-valent pneumococcal conjugate vaccine and the 23-valent pneumococcal polysaccharide vaccine at least 12 months apart, with the first occurrence after the age of 60.	N/A	66 and older

Composite Indicator

The composite rate is calculated by summing the number of immunizations that were administered to all adults across the plan's enrolled population (numerator) and dividing this sum by the total recommended number of immunizations, per clinical guidelines for the age group (denominator).

- For commercial and Medicaid plan members age 19–65 years of age, the composite denominator is determined by summing the influenza, Td/Tdap and herpes zoster vaccines that should have been administered to each member based on age (e.g., members age 49 are eligible for two vaccines [influenza and Tdap], while members age 50 are eligible for three vaccines [influenza, Tdap and zoster]. The composite numerator is determined by summing the influenza, Td/Tdap and herpes zoster vaccines that were indicated as administered).

- For Medicare plan members 66 years of age and older, the composite denominator is determined by summing the influenza, Td/Tdap, herpes zoster and pneumococcal vaccines. The composite numerator is determined by summing the influenza, Td/Tdap, herpes zoster and pneumococcal vaccines that were indicated as administered.

The intended use of the measure is to assess the quality of care in health plans across an adult population. As required by the specified level of accountability, we assessed data from all health plans reporting the HEDIS measure to NCQA in 2018 and conducted a field test with 2016 data from health plans to assess scientific acceptability, usability and feasibility.

2018 HEDIS Data

Data used to assess measure score reliability, construct validity, distribution of exclusions and average percentage of members that were excluded, meaningful differences in performance, missing data and components of the composite were calculated from all commercial, Medicaid and Medicare health plans submitting data to NCQA for this HEDIS measure. Data came from 71 commercial health plans, 21 Medicaid health plans and 44 Medicare health plans that were geographically diverse and varied in size. Data from administrative claims, electronic health records, registries, health information exchanges and case management systems were eligible for use in the measure in line with NCQA's HEDIS Electronic Clinical Data Systems reporting method. The plans submitting HEDIS data used a range of data sources: administrative claims, immunization registry, electronic health record and case management data.

2016 Field Test Data

We also analyzed effect of exclusions on overall measure scores and composite aggregration/weighting using additional data from a field test of the measures. In the field test, three geographically-diverse health plans (each comprising commercial, Medicaid and Medicare product lines) were asked to submit electronic patient-level demographic, enrollment, diagnosis, procedure and medication data to NCQA. Data from administrative claims, electronic health records, registries, health information exchanges and case management systems were eligible for use in the measure in line with NCQA's HEDIS Electronic Clinical Data Systems reporting method. The plans participating in this field test used a range of data sources: administrative claims, immunization registry and electronic health record data.

Systematic Evaluation of Face Validity

The measure was tested for face validity with three independent panels of experts.

- The Adult Immunizations Measurement Advisory Panel included 7 experts in primary care, immunizations and measures development, as well as clinician, health-plan and state/federal representatives.
- The Technical Measurement Advisory Panel includes 12 members, including representation by health plan methodologists, clinicians, HEDIS auditors and state/federal users of measures.
- NCQA's Committee on Performance Measurement (CPM) oversees measures used in NCQA programs and includes representation by purchasers, consumers, health plans, health care providers, and policy makers. This panel is composed of 21 independent members that reflect the diversity of constituencies that performance measurement serves. The CPM's recommendations are reviewed and approved by NCQA's Board of Directors.

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

2018 HEDIS Data

Data are summarized at the health plan level and stratified by plan type (i.e. commercial, Medicaid, Medicare). Below is a description of the sample. It includes number of health plans submitting the measure for HEDIS and the median eligible population for the measure across plans.

Table 1. Median eligible population for *Adult Immunization Status* by plan type, 2018

Plan Type	Number of Plans	Median number of eligible patients per plan
Commercial	71	80,330
Medicaid	21	36,250
Medicare	44	11,648

2016 Field Test Data

We stratified the field test data by product line (i.e. commercial, Medicare, Medicaid). Below is a description of the sample. It includes the number of health plans; the minimum and maximum number of adults in the elgible population for the measure; and the median percentage of adults stratified by age.

Table 2. Description of field test sample for *Adult Immunization Status* by plan type, 2016

	Number of plans	Minimum and maximum number of adults 19 and older across plans	Median percentage of adults ages 19-49	Median percentage of adults ages 50-59	Median percentage of adults ages 60-64	Median percentage of adults ages 65 and older
Commercial	3	313,932 – 1,544,512	65%	22%	10%	2%
Medicaid	3	23,650 – 537,000	78%	16%	6%	1%
Medicare	3	83,719 – 3.3 million	2%	3%	6%	91%

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

The 2018 HEDIS data were used to assess measure score reliability, construct validity, distribution of exclusions and average percentage of members that were excluded, meaningful differences in performance, missing data and components of the composite, as described above. For empirical validity testing, NCQA explored whether the composite and component measure rates were correlated with other relevant HEDIS measures that the plans reported in 2018.

The 2016 field test data were used to assess effect of exclusions on overall measure scores and composite aggregration/weighting, as well as in our systematic assessment of face validity.

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

We examined measure rates by commercial, Medicaid and Medicare health plans, which serves as a proxy for socioeconomic status. We did not analyze additional social risk factors. Patient-reported data and patient community characteristics were not available in the testing data source. Moreover, in the case of

immunizations, studies have consistently demonstrated that the factor most commonly associated with low immunization rates in the adult population is lack of recommendations from the health care provider.[1,2]

2a2. RELIABILITY TESTING

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

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

***Note**: Current guidance for composite measure evaluation states that reliability must be demonstrated for the composite performance measure score.*

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. 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*)

Reliability Testing of Performance Measure Score: We used the Beta-binomial model[3] to assess how well one can confidently distinguish the performance of one accountable entity from another. Conceptually, the Beta-binomial model is the ratio of signal to noise. The signal is the proportion of the variability in measured performance that can be explained by real differences in performance. The Beta-binomial model is an appropriate model when estimating the reliability of simple pass/fail rate measures as is the case with most HEDIS measures. Reliability scores range from 0.0 to 1.0. A score of zero implies that all variation is attributed to measurement error (i.e., noise), whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance (across accountable entities). The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one plan from another. A reliability score greater than or equal to 0.7 is considered very good.

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

Table 3. Distribution of Beta-Binomial Statistics For Each Measure Rate, Commercial Plans - 2018

Rate	Overall Reliability	Min	10th	25th	50th	75th	90th	Max
Composite	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000
Influenza	1.000	0.961	0.997	0.999	1.000	1.000	1.000	1.000
Td/Tdap	1.000	0.988	0.999	1.000	1.000	1.000	1.000	1.000
Herpes Zoster	0.999	0.902	0.992	0.996	0.999	0.999	1.000	1.000

Table 4. Distribution of Beta-Binomial Statistics For Each Measure Rate, Medicaid Plans - 2018

---

[1] Zimmerman RK, Santibanez TA, Janosky JE, et al. What affects influenza vaccination rates among older adults? An analysis from inner-city, suburban, rural, and Veterans Affairs practices. Am J Med. 2003;114(1):31–38.

[2] Johnson DR, Nichol KL, Lipczynski K. Barriers to adult immunization. Am J Med. 2008 Jul; 121(7 Suppl 2) S28-35

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

Rate	Overall Reliability	Min	10th	25th	50th	75th	90th	Max
Composite	1.000	0.990	0.997	1.000	1.000	1.000	1.000	1.000
Influenza	1.000	0.976	0.991	0.999	1.000	1.000	1.000	1.000
Td/Tdap	1.000	0.988	0.996	0.999	1.000	1.000	1.000	1.000
Herpes Zoster	0.999	0.686	0.978	0.996	0.999	1.000	1.000	1.000

Table 5. Distribution of Beta-Binomial Statistics For Each Measure Rate, Medicare Plans - 2018

Rate	Overall Reliability	Min	10th	25th	50th	75th	90th	Max
Composite	1.000	0.966	0.999	1.000	1.000	1.000	1.000	1.000
Influenza	1.000	0.838	0.995	0.998	1.000	1.000	1.000	1.000
Td/Tdap	1.000	0.904	0.996	0.998	1.000	1.000	1.000	1.000
Herpes Zoster	1.000	0.928	0.998	1.000	1.000	1.000	1.000	1.000
Pneumococcal	1.000	0.951	0.997	0.999	1.000	1.000	1.000	1.000

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

The values for the overall beta-binomial statistic across all product lines and measure rates are all greater than 0.7, indicating the measure has very good reliability. The distribution of health plan level-reliability on this measure shows that all health plans (across all product lines) are above the threshold of 0.7 except the Medicaid minimum for the herpes zoster rate, which is just below 0.7. Good reliability is demonstrated since most variance is due to signal and not to noise.

_____

2b1. VALIDITY TESTING

Note*: Current guidance for composite measure evaluation states that validity should be demonstrated for the composite performance measure score. If not feasible for initial endorsement, acceptable alternatives include assessment of content or face validity of the composite OR demonstration of validity for each component. Empirical validity testing of the composite measure score is expected by the time of endorsement maintenance.*

2b1.1. What level of validity testing was conducted?

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

☒ **Composite performance measure score**

☒ **Empirical validity testing**

☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

☒ **Validity testing for component measures** *(check all that apply)*

Note*: applies to ALL component measures, unless already endorsed or are being submitted for individual endorsement.*

☐ **Endorsed (or submitted) as individual performance measures**

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

☒ **Empirical validity testing of the component measure score(s)**

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

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to*

authoritative source, relationship to another measure as expected; what statistical analysis was used)

Empiric Validity Testing of Composite and Component Performance Measure Score

We empirically evaluated composite performance measure score validity and component measure score validity.

Empiric validity of the results were assessed using Pearson correlation to demonstrate construct validity. This test estimates the strength of the linear association between two variables; the magnitude of correlation ranges from -1 to +1. A value of 1 indicates a perfect linear dependence in which increasing values on one variable is associated with increasing values of the second variable. A value of 0 indicates no linear association. A value of -1 indicates a perfect linear relationship in which increasing values of the first variable is associated with decreasing values of the second variable. Coefficients with absolute values of less than 0.3 are generally considered indicative of weak associations, whereas absolute values of 0.3 or higher denote moderate to strong associations. The significance of a correlation coefficient is evaluated by testing the hypothesis that an observed coefficient calculated for the sample is different from zero. The resulting p-value indicates the probability of obtaining a difference at least as large as the one observed due to chance alone. We used a threshold of 0.05, as p-values less than this threshold imply it is unlikely that a non-zero coefficient was observed due to chance alone.

Across all plan types, we explored whether the indicators within this measure were correlated with each other. We hypothesized that health plans that perform well on the component vaccine rates within the measure should perform well on the composite rate within the measure. All of the measure rates represent an underlying quality construct of administering recommended routine vaccines to adults.

We also assessed correlation of *Adult Immunization Status* with other relevant HEDIS vaccine measures that they reported in 2018. For commercial and Medicaid plans, we explored whether the composite and component measure rates were correlated with the HEDIS *Prenatal Immunization Status* and *Immunizations for Adolescents* measures. These measures assess receipt of recommended vaccines for pregnant women and for adolescents by age 13 years, respectively. We hypothesized that health plans that perform well on the *Adult Immunization Status* measure should perform well on vaccine measures for pregnant women and adolescents. For Medicare plans, we explored whether the composite and component measure rates were correleated with the HEDIS *Flu Vaccinations for Older Adults* and *Pneumococcal Vaccination Status for Older Adults* measures that assess receipt of immunizations using CAHPS health plan member survey data. We hypothesized that health plans that perform well on the *Adult Immunization Status* measure should perform well on similar measures that assess patient-reported vaccination status.

Systematic Assessment of Face Validity of Composite and Component Performance Measure Score

NCQA develops measures using a standardized process. For new measures, face validity is assessed at various steps as described below.

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

STEP 2: Development ensures that measures are fully defined and tested before the organization collects them. MAPs participate in this process by helping identify the best measures for assessing health care performance in clinical areas identified in the topic selection phase. Development includes the following tasks: (1) Prepare a detailed conceptual and operational work-up that includes a testing proposal and (2) Collaborate with health plans to conduct field-tests that assess the feasibility and validity of potential measures. At this step, face validity is systematically determined by the CPM, which uses testing results and proposed final specifications to determine if the measure will move forward to Public Comment. For this measure, the CPM voted to approve moving the proposed measure forward to public comment (15 CPM members approved, 0 members opposed and 0 abstained).

STEP 3: Public Comment is a 30-day period of review that allows interested parties to offer feedback to NCQA about proposed new measures. Public comment offers an opportunity to assess the validity, feasibility, importance and other attributes of a measure from a wider audience. For this measure, a majority of public comment respondents supported the measure. NCQA MAPs and the technical panels consider all comments and advise NCQA staff on appropriate recommendations brought to the CPM. Face validity is then again systematically assessed by the CPM. The CPM reviews all comments before making a final decision and votes to recommend approval of new measures for HEDIS. NCQA's Board of Directors then approves new measures. For this measure, the CPM voted to approve the measure for HEDIS health plan reporting (16 CPM members approved, 0 members opposed and 0 abstained). The Board of Directors approved the measure.

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

Table 6. Health-Plan Level Pearson Correlation Coefficients Among Adult Immunization Measure Performance Scores Within Measure– 2018

	Adult Immunization Status	Composite	Influenza	Td/Tdap	Herpes Zoster	Pneumococcal
Commercial Plans	Composite	1.00	0.93	0.96	0.80	N/A
	Influenza	0.93	1.00	0.82	0.70	N/A
	Td/Tdap	0.96	0.82	1.00	0.71	N/A
	Herpes Zoster	0.80	0.70	0.71	1.00	N/A
	Pneumococcal	N/A	N/A	N/A	N/A	N/A
Medicaid Plans	Composite	1.00	0.95	0.95	0.79	N/A
	Influenza	0.95	1.00	0.89	0.74	N/A
	Td/Tdap	0.95	0.89	1.00	0.68	N/A
	Herpes Zoster	0.79	0.74	0.68	1.00	N/A
	Pneumococcal	N/A	N/A	N/A	N/A	N/A
Medicare Plans	Composite	1.00	0.77	0.98	0.95	0.94
	Influenza	0.77	1.00	0.67	0.62	0.58
	Td/Tdap	0.98	0.67	1.00	0.95	0.95
	Herpes Zoster	0.95	0.62	0.95	1.00	0.89
	Pneumococcal	0.94	0.58	0.95	0.89	1.00

Note: all correlations significant at p<0.05

Table 7a. Health-Plan Level Pearson Correlation Coefficients Among Adult Immunization Status and Prenatal Immunization Status Measure Performance Scores in Commercial Plans – 2018

	Prenatal Immunization Status		
Adult Immunization Status	Influenza	Tdap	Receipt of both vaccines
Composite	0.79*	0.58*	0.78*
Influenza	0.79*	0.54*	0.75*
Td/Tdap	0.74*	0.59*	0.74*
Herpes Zoster	0.61*	0.40*	0.61*

**significant at p<0.05*

Table 7b. Health-Plan Level Pearson Correlation Coefficients Among Adult Immunization Status and Prenatal Immunization Status Measure Performance Scores in Medicaid Plans – 2018

	Prenatal Immunization Status		
Adult Immunization Status	Influenza	Tdap	Receipt of both vaccines

Composite	0.91*	0.66*	0.89*
Influenza	0.85*	0.67*	0.85*
Td/Tdap	0.87*	0.60*	0.83*
Herpes Zoster	0.78*	0.54*	0.77*

**significant at p<0.05*

Table 8a. Health-Plan Level Pearson Correlation Coefficients Among Adult Immunization and Adolescent Immunization Measure Performance Scores in Commercial Plans – 2018

	Immunizations for Adolescents			
Adult Immunization Status	Meningococcal	Tdap	Human Papillomavirus Vaccine	Receipt of all vaccines
Composite	0.35*	0.37*	0.67*	0.66*
Influenza	0.42*	0.41*	0.64*	0.63*
Td/Tdap	0.31*	0.35*	0.69*	0.69*
Herpes Zoster	0.29*	0.31*	0.54*	0.54*

**significant at p<0.05*

Table 8b. Health-Plan Level Pearson Correlation Coefficients Among Adult Immunization Status and Adolescent Immunization Measure Performance Scores in Medicaid Plans – 2018

	Immunizations for Adolescents			
Adult Immunization Status	Meningococcal	Tdap	Human Papillomavirus Vaccine	Receipt of all vaccines
Composite	0.43	0.69*	0.62*	0.64*
Influenza	0.57*	0.75*	0.63*	0.69*
Td/Tdap	0.30	0.61*	0.51*	0.54*
Herpes Zoster	0.30	0.45*	0.71*	0.66*

**significant at p<0.05*

9. Health-Plan Level Pearson Correlation Coefficients Among Adult Immunization Status and Flu and Pneumococcal Vaccinations for Older Adults Measure Performance Scores in Medicare Plans – 2018

Adult Immunization Status	*Flu Vaccinations for Older Adults*	*Pneumococcal Vaccinations for Older Adults*
Composite	0.45*	0.33*
Influenza	0.36*	0.06
Td/Tdap	0.46*	0.40*
Herpes Zoster	0.38*	0.31*
Pneumococcal	0.45*	0.41*

**significant at p<0.05*

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Composite and Component Measure Score Validity

For the purposes of this analysis and the intended use of this measure to evaluate the quality of care for members across health plans, correlation was considered high (strong) if the correlation coefficient is 0.75 to 1, moderate if 0.25 to 0.75, and low (weak) if 0 to 0.25.

Correlations were strong among the rates within the composite measure for commercial, Medicaid and Medicare plans. This suggests that plans that perform well on one rate are likely to perform well on other rates within the measure.

Beyond the within-measure correlations, we saw a moderate/strong relationship with benchmarks on other measures of quality for commercial, Medicaid and Medicare plans.

- For commercial and Medicaid plans, the correlation between the adult immunization composite and component measure rates with the prenatal immunization composite and component measure rates was mostly strong.
- For commercial and Medicaid plans, the correlation between the adult immunization composite and component measure rates with the adolescent immunization composite and component measure rates was moderate.
- For Medicare plans, the correlation between the adult immunization composite and component measure rates with the *Flu Vaccinations for Older Adults* and *Pneumococcal Vaccination Status for Older Adults* measure rates were mostly moderate.

_____

2b2. EXCLUSIONS ANALYSIS

Note: *Applies to the composite performance measure, as well all component measures unless they are already endorsed or are being submitted for individual endorsement.*

NA ☐ no exclusions — *skip to section 2b4*

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

The measure specifications for the initial population and the exclusions are as follows:

Definition	Commercial/Medicaid Plans	Medicare Plans
Initial Population	Members ages 19-65 as of January 1 of the measurement period who were continuously enrolled throughout the measurement period.	Members ages 66 and older as of January 1 of the measurement period who were continuously enrolled throughout the measurement period.
Exclusions	Members with any of the following: • Active chemotherapy any time during the Measurement Period. • Bone marrow transplant any time during the Measurement Period. • History of immunocompromising conditions, cochlear implants, anatomic or functional asplenia, sickle cell anemia and HB-S disease or cerebrospinal fluid leaks any time during the member's history through the end of the Measurement Period. • In hospice or using hospice services during the Measurement Period.	

We assessed the distribution of the initial population and exclusions for this measure and the average percentage of members that were excluded from the initial population. To understand the impact of exclusions, a sensitivity analysis was conducted using data from our field test to estimate the effect of the exclusions on the overall measure composite rate. The composite measure rate was calculated with and without the exclusions applied.

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

Table 10. Distribution of initial population and exclusions - 2018

Plan Type	No. of plans	Measure Component	Mean	Min	10th	25th	50th	75th	90th	Max
Com-mercial	71	Initial Population	162,817	929	10,989	30,306	80,330	213,646	307,159	2,140,744
		Exclusions	4,886	41	334	867	2,468	6,976	10,184	39,607
Medicaid	21	Initial Population	92,621	322	1,467	19,095	36,250	73,554	279,199	562,980
		Exclusions	4,181	40	80	571	1744	3511	11783	26075
Medicare	44	Initial Population	34,681	67	810	2,391	11,648	34,584	78,921	502,633
		Exclusions	3,371	11	74	264	1,376	3,386	7,278	38,086

Table 11. Percentage of members excluded from the initial population - 2018

Plan Type	No. of plans	Mean
Commercial	71	3%
Medicaid	21	6%
Medicare	44	11%

Table 12. Field test: composite performance rate with and without exclusions - 2016

Plan Type	Plan	Composite rate without exclusions	Composite rate with exclusions
Commercial Plans	Plan A	10%	10%
	Plan B	20%	20%
	Plan C	58%	58%
Medicaid Plans	Plan A	2%	2%
	Plan B	17%	17%
	Plan C	51%	51%
Medicare Plans	Plan A	8%	8%
	Plan B	11%	11%
	Plan C	79%	79%

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

Removing exclusions had a minimal impact on the number of members in the denominator and no impact on the performance rates. Removing exclusions reduced the initial population, on average, by 3% for commercial plans, 6% for Medicaid plans and by 11% for Medicare plans. We would expect more Medicare members to be excluded because it is a population which may be more likely to meet the specified exclusions.

Experts on our measurement advisory panels recommended specifying the exclusions in the measure based on the clinical rationale and from an accountability perspective, and because it is feasible to collect the data with minimal burden.

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

Note: *Applies to all outcome or resource use component measures, unless already endorsed or are being submitted for individual endorsement.*

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b4.

2b3.1. What method of controlling for differences in case mix is used? *(check all that apply)*

☐ **Endorsed (or submitted) as individual performance measures**

☐ **No risk adjustment or stratification**

☐ **Statistical risk model with** Click here to enter number of factors **risk factors**
☐ **Stratification by** Click here to enter number of categories **risk categories**
☐ **Other,** Click here to enter description

2b3.1.1 If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*) **Also discuss any "ordering" of risk factor inclusion;** for example, are social risk factors added after all clinical factors?

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:
☐ **Published literature**
☐ **Internal data analysis**
☐ **Other (please describe)**

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

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors (e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.) Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
If stratified, skip to 2b3.9

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

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

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

2b3.9. Results of Risk Stratification Analysis:

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

_____

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE
Note: *Applies to the composite performance measure.*

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

To demonstrate meaningful differences in performance, NCQA calculated an inter-quartile range (IQR) for each indicator. The IQR provides a measure of the dispersion of performance. The IQR can be interpreted as the difference between the 25th and 75th percentile on a measure.

To determine if this difference is statistically significant, NCQA calculated an independent sample t-test of the performance difference between two randomly selected plans at the 25th and 75th percentile. The t-test method calculates a testing statistic based on the sample size, performance rate, and standardized error of each plan. The test statistic is then compared against a normal distribution. If the p value of the test statistic is less than 0.05, then the two plans' performance is significantly different from each other. Using this method, we compared the performance rates of two randomly selected plans, one plan in the 25th percentile and another plan in the 75th percentile of performance. We used these two plans as examples of measures entities. However, the method can be used for comparison of any two measured entities.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

Table 13. Variation in Performance Across Commercial Plans, 2018

Rate	No. of plans	Mean denom-inator	Mean rate (%)	Min	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Max	IQR	p-value
Composite	71	373,600	21.2	8.8	13.5	16.0	18.1	22.9	30.2	58.2	6.9	<0.001
Influenza	71	157,931	18.7	7.7	11.4	14.9	18.1	20.6	26.4	53.6	5.7	<0.001
Td/Tdap	71	157,931	29.4	11.1	18.6	20.9	25.2	30.7	46.5	78.0	9.8	<0.001
Herpes Zoster	71	57,739	6.1	0.7	2.7	4.1	5.0	6.5	9.8	25.2	2.4	<0.001

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

Table 14. Variation in Performance Across Medicaid Plans, 2018

Rate	No. of plans	Mean denom-inator	Mean rate (%)	Min	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Max	IQR	p-value
Composite	21	221,556	14.0	1.8	2.7	10.0	13.7	17.5	20.7	36.0	7.5	<0.001
Influenza	21	88,440	11.6	1.1	2.8	7.8	11.7	15.3	20.7	26.0	7.5	<0.001
Td/Tdap	21	88,440	20.9	2.8	4.9	14.1	21.1	25.0	34.4	52.5	10.9	<0.001
Herpes Zoster	21	44,676	1.6	0.0	0.0	0.4	0.6	1.4	5.3	6.7	1.0	<0.001

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

Table 15. Variation in Performance Across Medicare Plans

Rate	No. of plans	Mean denom-inator	Mean rate (%)	Min	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Max	IQR	p-value
Composite	44	122,965	19.5	1.3	6.6	9.5	14.4	20.6	43.8	79.8	11.1	<0.001
Influenza	44	31,310	18.3	0.5	5.3	8.3	12.5	21.7	30.1	80.0	13.4	<0.001
Td/Tdap	44	31,310	26.5	3.3	9.5	14.8	20.7	28.7	56.4	89.2	13.9	<0.001
Herpes Zoster	44	31,099	12.9	0.0	0.4	0.9	5.3	14.5	39.5	81.0	13.6	<0.001
Pneumo-coccal	44	29,246	20.3	0.5	5.4	8.1	10.8	22.5	55.5	84.2	14.4	<0.001

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

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

The IQRs ranged from 2%-9% across the commercial plan rates, 1%-11% across the Medicaid plan rates and 11%-14% across the Medicare plan rates. For example, in commercial plans, the IQR for the influenza rate was 5.7%. This gap represents an average of 9,000 additional patients being up-to-date on the influenza vaccine in high-performing commercial plans compared to low-performing commercial plans.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
Note: *Applies to all component measures, unless already endorsed or are being submitted for individual endorsement.*
If only one set of specifications, this section can be skipped.

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

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted?*)

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS
Note: *Applies to the overall composite measure.*

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

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

The HEDIS Compliance Audit addresses the following functions:

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

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

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

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

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

2c. EMPIRICAL ANALYSIS TO SUPPORT COMPOSITE CONSTRUCTION APPROACH
Note: *If empirical analyses do not provide adequate results—or are not conducted—justification must be provided and accepted in order to meet the must-pass criterion of Scientific Acceptability of Measure Properties. Each of the following questions has instructions if there is no empirical analysis.*

2d1. Empirical analysis demonstrating that the component measures fit the quality construct, add value to the overall composite, and achieve the object of parsimony to the extent possible.

2d1.1 Describe the method used (*describe the steps—do not just name a method; what statistical analysis was used; if no empirical analysis, provide justification)*

Internal Consistency: We used Cronbach's alpha statistic (internal consistency coefficients) to measure the extent to which the components (e.g., individual vaccine rates) represent a single quality construct.

2d1.2. What were the statistical results obtained from the analysis of the components? (e.g., *correlations, contribution of each component to the composite score, etc.*; *if no empirical analysis, identify the components that were considered and the pros and cons of each*)

Internal Consistency Results: Table 16 shows the results of the calculation of the Cronbach's alpha for the composite rate in commercial, Medicaid and Medicare plans. The Cronbach's alpha statistic was >0.94 across the plan types.

Table 16. Cronbach's Alpha for Composite Rate, 2018

	No. of plans	Cronbach's alpha
Commercial Plans	71	0.948
Medicaid Plans	21	0.952
Medicare Plans	44	0.961

2d1.3. What is your interpretation of the results in terms of demonstrating that the components included in the composite are consistent with the described quality construct and add value to the overall composite? (i.*e., what do the results mean in terms of supporting inclusion of the components; if no empirical analysis, provide rationale for the components that were selected)*

Our results suggest there is good internal consistency within the composite rate across all three plan types.

2d2. Empirical analysis demonstrating that the aggregations and weighting rules are consistent with the quality construct and achieve the objective of simplicity to the extent possible

2d2.1 Describe the method used (*describe the steps—do not just name a method; what statistical analysis was used; if no empirical analysis, provide justification)*

The measure composite rate assesses the percentage of total recommended routine vaccines received across adult members enrolled in health plans (see section 1.5 for more detail). The team also explored an alternative construction of the measure composite rate: an all-or-nothing composite construction that assesses the percentage of adult members who received all of the recommended vaccines for their age. To determine the appropriate method of aggregation for the measure, we constructed these two approaches to the composite rate and calculated measure performance (alpha testing). We also explored the impact of including and excluding specific vaccines in the composite rate (beta testing). We used feedback from the Adult Immunization Measurement Advisory Panel, the Technical Measurement Advisory Panel and the Committee on Performance Measurement to determine the appropriate construction for the measure.

2d2.2. What were the statistical results obtained from the analysis of the aggregation and weighting rules? (e.g., *results of sensitivity analysis of effect of different aggregations and/or weighting rules; if no empirical analysis, identify the aggregation and weighting rules that were considered and the pros and cons of each*)

Alpha Testing Results

Table 17. Composite Performance Rates Based on Two Measure Construction Approaches

Plan Type	Plan	Percentage of total recommended vaccines received across adult population (current composite measure specifications)	Percentage of members who received all recommended vaccines (alternative all-or-nothing approach)
Commercial Plans	Plan A	10%	1%
	Plan B	20%	6%
	Plan C	58%	31%
Medicaid Plans	Plan A	2%	0.2%
	Plan B	17%	6%
	Plan C	51%	26%
Medicare Plans	Plan A	8%	0%
	Plan B	11%	0.2%
	Plan C	79%	50%

Beta Testing Results

Table 18. Composite Performance Rates With and Without Individual Vaccines, Commercial & Medicaid Plans

Plan Type	Plan	Percentage of total recommended vaccines received			
		Includes - Influenza - Td/Tdap -Herpes Zoster (ages 60-64)	Includes - Td/Tdap - Herpes Zoster (ages 60-64)	Includes - Influenza - Herpes Zoster (ages 60-64)	Includes - Influenza - Td/Tdap
Commercial Plans	Plan A	10%	6%	13%	10%
	Plan B	20%	23%	18%	20%
	Plan C	58%	81%	37%	58%
Medicaid Plans	Plan A	2%	1%	3%	2%
	Plan B	17%	23%	10%	17%
	Plan C	51%	73%	30%	51%

Table 19. Composite Performance Rates With and Without Individual Vaccines, Medicare Plans

Plan Type	Plan	Percentage of total recommended vaccines received				
		Includes - Influenza - Td/Tdap - Herpes Zoster - Pneumococcal	Includes - Td/Tdap - Herpes Zoster - Pneumococcal	Includes - Influenza - Herpes Zoster - Pneumococcal	Includes - Influenza - Td/Tdap - Pneumococcal	Includes: - Influenza - Td/Tdap -Herpes Zoster
Medicare Plans	Plan A	8%	1%	9%	10%	10%
	Plan B	11%	12%	8%	12%	11%
	Plan C	79%	80%	73%	76%	84%

2d2.3. What is your interpretation of the results in terms of demonstrating the aggregation and weighting rules are consistent with the described quality construct? (i.*e., what do the results mean in terms of supporting the selected rules for aggregation and weighting; if no empirical analysis, provide rationale for the selected rules for aggregation and weighting*)

As expected, the alpha testing results showed that composite performance rates for commercial, Medicaid and Medicare plans were higher using the "percentage of total recommended vaccines received" composite construction compared to the "all-or-nothing" composite construction.

The beta testing results showed that certain vaccines can have a large impact on the composite, but which vaccines drove the composite rates was inconsistent. For example, for commercial and Medicaid plans, Plan A, which had a high influenza vaccinination rate, had a much lower composite rate when this vaccine was removed (as expected). The other plans' composites improved when removing the influenza vaccine. For Medicare plans, removal of either herpes zoster or pneumoccocal vaccine improved composite rates for Plan A; removal of either herpes zoster or influenza vaccine improved rates for Plan B; and removal of pneumoccocal vaccine improved rates for Plan C.

The expert panels determined that the most appropriate composite construction was the percentage of total vaccines received composite rate. Because herpes zoster and pneumococcal vaccines are only recommended for older adults while influenza and Td/Tdap are recommended for all adults, they concluded that the total vaccine approach is more simple and actionable than the all-or-nothing composite. There were differences among plans with respect to which vaccines contributed the most or least to the composite performance rate. Qualitative feedback from the field test sites suggest that this may be a due to health plans focusing quality improvement efforts on specific vaccinations (for instance, one plan had an iniative in place to improve influenza vaccination rates among Medicare members in 2016). In addition, the expert panels determined there was no conceptual basis for differential weighting per vaccine given all vaccines are of equal clinical importance.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

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

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

De.5. Subject/Topic Area *(check all the areas that apply):*

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

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

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

Not applicable.

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

This is not an eMeasure **Attachment:**

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

Attachment: 3483_AIS_Value_Sets_Fall_2019-637093357011416352.xlsx

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

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

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

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

No

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

N/A

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Adults who are up-to-date on influenza, Td or Tdap, herpes zoster and pneumococcal vaccinations based on age and recommendations.

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The measure calculates a numerator for each vaccine type and a composite numerator.

Numerator 1 (influenza): adults 19 and older who received an influenza vaccine on or between July 1 of the year prior to the measurement period and June 30 of the measurement period, or who had a prior influenza virus vaccine adverse reaction any time before or during the Measurement Period.

Numerator 2 (Td/Tdap): adults 19 and older who received at least one Td or one Tdap vaccine between nine years prior to the measurement period and the end of the measurement period, or with a history of at least one of the following contraindications any time before or during the Measurement Period: anaphylaxis due to Tdap vaccine, anaphylaxis due to Td vaccine or its components, or encephalopathy due to Tdap or Td vaccination (post-tetanus vaccination encephalitis, post-diphtheria vaccination encephalitis, post-pertussis vaccination encephalitis).

Numerator 3 (herpes zoster): adults 50 and older who received at least one dose of the herpes zoster live vaccine or two doses of the herpes zoster recombinant vaccine at least 28 days apart, any time on or after the member's 50th birthday and before or during the Measurement Period, or who had a prior adverse reaction caused by zoster vaccine or its components any time before or during the Measurement Period.

Numerator 4 (pneumococcal): adults 66 and older who received both the 13-valent pneumococcal conjugate vaccine and the 23-valent pneumococcal polysaccharide vaccine at least 12 months apart, with the first occurrence after the age of 60, before or during the Measurement Period, or prior pneumococcal vaccine adverse reaction any time before or during the Measurement Period.

Numerator 5 (composite): The total number of immunizations administered to members across the plan's adult population, per clinical guideline recommendations for the age group (sum of numerators 1-4).

See attached code value sets.

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

Adults ages 19 years and older.

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Adults ages 19 years and older at the start of the measurement period (January 1). The measure calculates a denominator for each vaccine type and a composite denominator.

Denominator 1 (influenza): adults 19 and older by the start of the measurement period.

Denominator 2 (Td/Tdap): adults 19 and older by the start of the measurement period.

Denominator 3 (herpes zoster): adults 50 and older by the start of the measurement period.

Denominator 4 (pneumococcal): adults 66 and older by the start of the measurement period.

Denominator 5 (composite): the total number of immunizations recommended for members, determined by their age at the start of the measurement period, per clinical guideline recommendations (sum of denominators 1-4).

Note: Commercial and Medicaid plans report denominators for members 19–65; Medicare plans report denominators for members 66 and older.

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

Adults who received chemotherapy, had a bone marrow transplant or were in hospice during the measurement year or those with a history of immunocompromising conditions.

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

Exclude adults with any of the following:

Active chemotherapy any time during the measurement period.

Bone marrow transplant any time during the measurement period.

History of immunocompromising conditions, cochlear implants, anatomic or functional asplenia, sickle cell anemia & HB-S disease or cerebrospinal fluid leaks any time during the member's history through the end of the measurement period.

In hospice or using hospice services during the measurement period.

See attached code value sets.

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

N/A

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

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

Better quality = Higher score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

Step 1: Determine the eligible population. Identify all adults ages 19 and older at the start of the measurement period who were continuously enrolled in the plan during the measurement period (January 1-December 31).

Step 2: Remove adults with any of the following exclusions from the eligible population: active chemotherapy during the measurement period; bone marrow transplant during the measurement period; history of immunocompromising conditions, cochlear implants, anatomic or functional asplenia, sickle cell anemia and HB-S disease or cerebrospinal fluid leaks any time during the member's history through the end of the measurement period; in hospice or using hospice services during the measurement period.

Step 3: Determine denominators 1-5 based on the age of the members at the start of the measurement period:

Commercial and Medicaid plans:

-Denominator 1 (influenza): ages 19-65

-Denominator 2 (Td/Tdap): ages 19-65

-Denominator 3 (herpes zoster): ages 50-65

-Denominator 4 (pneumococcal): N/A

-Denominator 5 (composite): sum of denominators 1-3

Medicare plans:

-Denominator 1 (influenza): ages 66 and older

-Denominator 2 (Td/Tdap): ages 66 and older

-Denominator 3 (herpes zoster): ages 66 and older

-Denominator 4 (pneumococcal): ages 66 and older

-Denominator 5 (composite): sum of denominators 1-4

Step 4: Determine numerators 1-5:

Commercial and Medicaid plans:

-Numerator 1 (influenza): received an influenza vaccine on or between July 1 of the year prior to the measurement period and June 30 of the measurement period, or who had a prior influenza virus vaccine adverse reaction any time before or during the Measurement Period.

-Numerator 2 (Td/Tdap): received at least one Td or one Tdap vaccine between nine years prior to the measurement period and the end of the measurement period, or with a history of at least one of the following contraindications any time before or during the Measurement Period: anaphylaxis due to Tdap vaccine,

anaphylaxis due to Td vaccine or its components, or encephalopathy due to Tdap or Td vaccination (post-tetanus vaccination encephalitis, post-diphtheria vaccination encephalitis, post-pertussis vaccination encephalitis).

-Numerator 3 (herpes zoster): received at least one dose of the herpes zoster live vaccine or two doses of the herpes zoster recombinant vaccine at least 28 days apart, any time on or after the member's 50th birthday and before or during the Measurement Period, or who had a prior adverse reaction caused by zoster vaccine or its components any time before or during the Measurement Period.

-Numerator 4 (pneumococcal): N/A

-Numerator 5 (composite): sum of numerators 1-3

Medicare plans:

-Numerator 1 (influenza): received an influenza vaccine on or between July 1 of the year prior to the measurement period and June 30 of the measurement period, or who had a prior influenza virus vaccine adverse reaction any time before or during the Measurement Period.

-Numerator 2 (Td/Tdap): received at least one Td or one Tdap vaccine between nine years prior to the measurement period and the end of the measurement period, or with a history of at least one of the following contraindications any time before or during the Measurement Period: anaphylaxis due to Tdap vaccine, anaphylaxis due to Td vaccine or its components, or encephalopathy due to Tdap or Td vaccination (post-tetanus vaccination encephalitis, post-diphtheria vaccination encephalitis, post-pertussis vaccination encephalitis).

-Numerator 3 (herpes zoster): received at least one dose of the herpes zoster live vaccine or two doses of the herpes zoster recombinant vaccine at least 28 days apart, any time on or after the member's 50th birthday and before or during the Measurement Period, or who had a prior adverse reaction caused by zoster vaccine or its components any time before or during the Measurement Period.

-Numerator 4 (pneumococcal): received both the 13-valent pneumococcal conjugate vaccine and the 23-valent pneumococcal polysaccharide vaccine at least 12 months apart, with the first occurrence after the age of 60, before or during the Measurement Period, or prior pneumococcal vaccine adverse reaction any time before or during the Measurement Period.

-Numerator 5 (composite): sum of numerators 1-4

Step 5: Calculate the measure rates:

-Numerator 1 / Denominator 1

-Numerator 2 / Denominator 2

-Numerator 3 / Denominator 3

-Numerator 4 / Denominator 4 (N/A for commercial and Medicaid plans)

-Numerator 5 / Denominator 5

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

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

N/A

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

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

N/A

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

If other, please describe in S.18.

Claims, Electronic Health Data, Electronic Health Records, Enrollment Data, Management Data, Other, Registry Data

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

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

This measure is specified for administrative claims, electronic health record, registry, health information exchange or case management data collected in the course of providing care to health plan members. NCQA collects the Healthcare Effectiveness Data and Information Set (HEDIS) data for this measure directly from Health Management Organizations and Preferred Provider Organizations via NCQA's online data submission system.

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

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

Health Plan

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

Outpatient Services

If other:

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

The components are weighted equally in the composite to assess compliance with immunization guidelines for adults.

2. Validity – See attached Measure Testing Submission Form

AIS_Composite_Testing_Form.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

No

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

No

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated

even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

3. Feasibility

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

3a. Byproduct of Care Processes

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

3a.1. Data Elements Generated as Byproduct of Care Processes.

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score), Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

If other:

3b. Electronic Sources

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

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

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

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

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

Attachment:

3c. Data Collection Strategy

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

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

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

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

The HEDIS Compliance Audit addresses the following functions:

1)Information practices and control procedures

2)Sampling methods and procedures

3)Data integrity

4)Compliance with HEDIS specifications

5)Analytic file production

6)Reporting and documentation

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

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

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

4. Usability and Use

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

4a. Accountability and Transparency

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

4.1. Current and Planned Use

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

Specific Plan for Use	Current Use (for current use provide URL)

Public Reporting Quality Improvement (external benchmarking to organizations)	Quality Improvement (Internal to the specific organization) Healthcare Effectiveness Data and Information Set (HEDIS) https://www.ncqa.org/hedis/using-hedis-measures/

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

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

HEDIS: The Healthcare Effectiveness Data and Information Set (HEDIS) is one of health care's most widely used performance improvement tools.190 million people are enrolled in health plans across the nation that report HEDIS results. HEDIS measures are used by health plans and other various levels of the health care system for quality improvement initiatives.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

This measure was a new HEDIS measure in 2018. NCQA's standard process is to evaluate data for all new measures prior to use for public reporting, benchmarking and/or other programs.

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

As part of new measure evaluation, NCQA works with multi-stakeholder advisory panels to assess the number of plans that have shown they can report the measure; whether measure results match what we expect; whether results seem indicative of true performance; and whether performance indicates an opportunity for improvement for the industry overall. Because this measure uses the newer HEDIS Electronic Clinical Data Systems Reporting Method, NCQA's timeline and plan are to assess these issues after each year of the measure's reporting. We anticipate that the measure will be approved for public reporting and eligible for use in programs within the next several years, but this is pending our continued assessment.

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

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

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

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

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

Describe how feedback was obtained.

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

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

During a recent public comment posting held during the measure development process, most of the comments from measured entities supported the new measure. In general, respondents found the measures to be relevant and clearly specified.

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

This measure has been deemed a priority measure by NCQA and other entities such as the Centers for Disease Control and Prevention, the federal National Vaccine Program Office and the American Immunization Registry Association. During a recent public comment posting conducted during the measure development process, commenters were supportive of the measure and specifically highlighted the need for measures assessing routine adult immunizations. Commenters noted that many adults still do not receive these important vaccines, despite Advisory Committee on Immunization Practices recommendations and national efforts to improve adult immunization rates in the US.

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

During measure development, feedback obtained through the mechanisms described in 4a2.2.1 informed how we specified the measure to align with immunization guidelines from the Advisory Committee on Immunization Practices.

Improvement

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

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

This is a new measure; therefore, we do not yet have data on improvement over time. Adoption of this measure has the potential to improve the immunization rates for adults.

4b2. Unintended Consequences

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

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

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

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

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

5. Comparison to Related or Competing Measures

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

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

0041 : Preventive Care and Screening: Influenza Immunization

0043 : Pneumococcal Vaccination Status for Older Adults (PNU)

0431 : INFLUENZA VACCINATION COVERAGE AMONG HEALTHCARE PERSONNEL

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

0681 : Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)

0682 : Percent of Residents or Patients Assessed and Appropriately Given the Pneumococcal Vaccine (Short-Stay)

0683 : Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long-Stay)

1653 : Pneumococcal Immunization

1659 : Influenza Immunization

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

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):

Are the measure specifications harmonized to the extent possible?

No

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

This measure assesses influenza, Td/Tdap, herpes zoster and pneumococcal vaccination for a general adult population. It is a population-based measure that assesses vaccines provided in the outpatient setting at the health plan level. Most of the other NQF-endorsed vaccination measures focus only on either pneumococcal or influenza vaccination. These measures specifically apply to inpatient populations, residents in long-term care/skilled nursing facilities or healthcare personnel or are specified at the provider-level. Moreover, our proposed measure is specified to use electronic clinical data, while other related measures (e.g., NQF 0039) are specified to use survey data in which patients must recall whether they had received a vaccine.

5b. Competing Measures

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

OR

Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Our proposed measure is more specific than several of the other adult vaccination measures because it assesses whether health plan members received the appropriate type and doses of vaccines at the right time according to clinical guidelines. Other vaccine measures that require the use of survey data are less specific because they rely on patient recall of whether they had received a vaccine. In addition, our proposed measure combines all recommended routine vaccines in one measure, which provides a more complete picture of routine adult vaccinations at the health plan level.

Appendix

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

No appendix Attachment:

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): National Committee for Quality Assurance

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

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

Co.4 Point of Contact: Bob, Rehm, nqf@ncqa.org, 202-955-1728-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

ADULT IMMUNIZATION MEASUREMENT ADVISORY PANEL

Alison Chi, American Immunization Registry Association

Nicole Johnson, CareSource Management Group

Sarah Royce, California Department of Health

Krista Ventrone, Excellus Blue Cross/Blue Shield

Kimberly Wildes, ThedaCare Physicians

Walter Williams, Centers for Disease Control and Prevention

Jane Zucker, New York City Department of Health and Mental Hygiene

TECHNICAL MEASUREMENT ADVISORY PANEL?

Andy Amster, MSPH, Kaiser Permanente?

Sarah Bezeredi, MBA, MSHL, UnitedHealth Group

Jennifer Brudnicki, MBA, Inovalon Inc.

Lindsay Cogan, MS, PhD, New York State Department of Health

Mike Farina, MBA, R.Ph, Capital District Physicians' Health Plan

Marissa Finn, MBA, CIGNA?

Scott Fox, MS, Med, FAMIA, The MITRE Corporation?

Carlos Hernandez,?CenCal?Health?

Harmon Jordan, ScD, Westat??

Gigi Raney, LCSW, Center for Medicaid and CHIP Services

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

Laurie Spoll, Aetna

?

COMMITTEE ON PERFORMANCE MEASUREMENT??

Andrew Baskin, MD, Aetna??

Elizabeth Drye, MD, SM, Yale School of Medicine

Andrea Gelzer, MD, MS, FACP, AmeriHealth Caritas

Kate Goodrich, MD, MHS, Centers for Medicare & Medicaid Services

David Grossman, MD, MPH, Washington Permanente Medical Group

Christine Hunter, (Co-Chair), MD, WPS Health Solutions?

David Kelley, MD, MPA, Pennsylvania Department of Human Services

Jeffrey Kelman,?MMSc, MD, Department of Health and Human Services

Nancy Lane, PhD, Independent Consultant

Bernadette Loftus, MD, Freelance

Adrienne Mims, MD, MPH, AGSF, FAAFP, Alliant Health Solutions

Amanda Parsons, MD, MBA, Metroplus

Wayne Rawlins, MD, MBA, ConnectiCare

Misty Roberts, MSN, RN, CPHQ, PMP, Humana

Rudy Saenz, MD, MMM, FACOG, Riverside Medical Clinic

Marcus?Thygeson, (Co-Chair), MD, MPH, Blind On-Demand

JoAnn Volk, MA, Georgetown University

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

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2018

Ad.3 Month and Year of most recent revision:

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

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

Ad.6 Copyright statement: © 2018 by the National Committee for Quality Assurance

1100 13th Street, NW, 3rd floor

Washington, DC 20005

Ad.7 Disclaimers: These performance measures are not clinical guidelines and do not establish a standard of medical care, and have not been tested for all potential applications. THE MEASURSE AND SPECIFICATIONS ARE PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND.

Ad.8 Additional Information/Comments: NCQA Notice of Use. Broad public use and dissemination of these measures is encouraged and NCQA has agreed with NQF that noncommercial uses do not require the consent of the measure developer. Use by health care physicians in connection with their own practices is not commercial use. Commercial use of a measure requires the prior written consent of NCQA. As used herein, “commercial use” refers to any sale, license, or distribution of a measure for commercial gain, or incorporation of a measure into any product or service that is sold, licensed, or distributed for commercial gain, even if there is no actual charge for inclusion of the measure.

These performance measures were developed and are owned by NCQA. They are not clinical guidelines and do not establish a standard of medical care. 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. NCQA holds a copyright in these measures and can rescind or alter these measures at any time. Users of the measures shall not have the right to alter, enhance, or otherwise modify the measures, and shall not disassemble, recompile, or reverse engineer the source code or object code relating to the measures. Anyone desiring to use or reproduce the measures without modification for a noncommercial purpose may do so without obtaining approval from NCQA. All commercial uses must be approved by NCQA and are subject to a license at the discretion of NCQA. © 2018 by the National Committee for Quality Assurance.