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

Purple text represents the responses from measure developers.

Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 2903

Corresponding Measures:

De.2. Measure Title: Contraceptive Care – Most & Moderately Effective Methods

Co.1.1. Measure Steward: HHS Office of Population Affairs

De.3. Brief Description of Measure: The percentage of women aged 15-44 years at risk of unintended pregnancy that is provided a most effective (i.e., sterilization, implants, intrauterine devices or systems (IUD/IUS)) or moderately effective (i.e., injectables, oral pills, patch, or ring) method of contraception.

The measure is an intermediate outcome measure because it represents a decision that is made at the end of a clinical encounter about the type of contraceptive method a woman will use, and because of the strong association between type of contraceptive method used and risk of unintended pregnancy.

1b.1. Developer Rationale: Unintended pregnancies and interpregnancy intervals of less than 18 months have been associated with poor perinatal outcomes such as preterm birth, low birth weight, small size for gestational age, as well as adverse maternal outcomes [1, 2]. Studies among U.S. women report that women at younger maternal age are at higher risk for unintended pregnancy [14] and older maternal age is associated with closely spaced pregnancies [15]. Contraception is a highly effective clinical preventive service that can assist women in reaching their reproductive health goals, like reducing unintended pregnancies and the percentage of births occurring within 18 months of a previous birth [3, 4]. The type of contraceptive method used by a woman is strongly associated with her risk of unintended pregnancy. The most effective methods (LARC and sterilization) have a failure rate that is less than 1% per year under typical use [4]. The moderately effective methods (injectable, pill, patch, ring) have a typical failure rate of 4-7% per year, while the less effective methods have a typical failure rate of 13-27% [4]. One recent study also indicates that the most used contraceptive methods in the United States have

experienced reductions in their typical use failure rates [16]. Not using any method at all has a typical failure rate of 85% [4].

After NQF endorsed #2903 in 2016, OPA published multiple articles in peer-reviewed journals to inform health care providers in public and private settings (e.g., commercial health plans, Medicaid, community health centers, free-standing reproductive health clinics) about the new measure. These publications outline our conceptual framework for developing #2903 alongside its two complementary measures (NQF #2902 and #2904) and describe appropriate measure implementation and use. Furthermore, OPA highlighted systematic reviews which indicate that effective contraceptive method use increases the interbirth interval and reduces adolescent and unintended pregnancies. This association between use of most and moderately effective methods and positive reproductive health outcomes demonstrates the importance of contraceptive care measures to health care quality [17-19].

While NQF #2903 and the contraceptive care measures reflect that some contraceptive methods are more effective than others at preventing pregnancy, these measures and their guidelines for use are designed to encourage providers to offer those clients seeking contraception the full range of methods. The goal of providing contraception should never be to recommend any one method or class of methods over women's individual choices. Women who want to delay or prevent pregnancy should have access to a broad range of contraceptive methods, preferably on a same-day, on-site basis. Furthermore, it is important that these contraceptive services are provided in a client-centered manner that treats each person as a unique individual with respect, empathy, and understanding, providing accurate, easy-to-understand information based on the client's self-identified needs, goals, preferences, and values [11]. Patients receiving client-centered care may feel motivated to continue seeking reproductive health care for contraception and if they become pregnant, prenatal care and birth [13]. Thus, efforts to provide client-centered contraceptive services aligned with the American Academy of Pediatrics (AAP), the American College of Obstetricians and Gynecologists (ACOG), the Centers for Disease Control and Prevention (CDC), and Office of Population Affairs (OPA) recommendations [7-12] may be strengthened by quality improvement processes based on standardized metrics of contraceptive care provision.

References

[1] Conde-Agudelo, A., Rosas-Bermúdez, A., & Kafury-Goeta, A. C. (2006). Birth spacing and risk of adverse perinatal outcomes: a meta-analysis. JAMA, 295(15), 1809–1823. https://doi.org/10.1001/jama.295.15.1809

[2] Conde-Agudelo A, Rosas-Bermúdez A, Kafury-Goeta AC. Effects of birth spacing on maternal health: a systematic review. Am J Obstet Gynecol. 2007 Apr;196(4):297-308. doi: 10.1016/j.ajog.2006.05.055. PMID: 17403398.

[3] Mansour, D., Inki, P., & Gemzell-Danielsson, K. (2010). Efficacy of contraceptive methods: A review of the literature. The European journal of contraception & reproductive health care: the official journal of the European Society of Contraception, 15(1), 4–16. https://doi.org/10.3109/13625180903427675

[4] Trussell, J., Aiken, A.R.A., Micks, E., Guthrie, K.A. (2018). Efficacy, safety, and personal considerations. In R.A. Hatcher, A.L. Nelson, J. Trussell, C. Cwiak, P. Cason, M.S. Policar, A. Edelman, A.R.A. Aiken, J. Marrazzo, D. Kowal (Eds.). Contraceptive technology (21st ed., pp. 95–128). Ayer Company Publishers, Inc.

[5] Winner, B., Peipert, J. F., Zhao, Q., Buckel, C., Madden, T., Allsworth, J. E., & Secura, G. M. (2012). Effectiveness of long-acting reversible contraception. The New England journal of medicine, 366(21), 1998–2007. https://doi.org/10.1056/NEJMoa1110855

[6] Harper, C. C., Rocca, C. H., Thompson, K. M., Morfesis, J., Goodman, S., Darney, P. D., Westhoff, C. L., & Speidel, J. J. (2015). Reductions in pregnancy rates in the USA with long-acting reversible contraception: a cluster randomised trial. Lancet (London, England), 386(9993), 562–568. https://doi.org/10.1016/S0140-6736(14)62460-0

[7] Committee on Gynecologic Practice Long-Acting Reversible Contraception Working Group (2015). Committee Opinion No. 642: Increasing Access to Contraceptive Implants and Intrauterine Devices to Reduce Unintended Pregnancy. Obstetrics and gynecology, 126(4), e44–e48. https://doi.org/10.1097/AOG.0000000000001106

[8] ACOG Committee Opinion No. 735: Adolescents and Long-Acting Reversible Contraception: Implants and Intrauterine Devices. (2018). Obstetrics and gynecology, 131(5), e130–e139. https://doi.org/10.1097/AOG.0000000000002632

[9] Menon, S., & COMMITTEE ON ADOLESCENCE (2020). Long-Acting Reversible Contraception: Specific Issues for Adolescents. Pediatrics, 146(2), e2020007252. https://doi.org/10.1542/peds.2020-007252

[10] Committee on Adolescence (2014). Contraception for adolescents. Pediatrics, 134(4), e1244–e1256. https://doi.org/10.1542/peds.2014-2299

[11] Gavin, L., Moskosky, S., Carter, M., Curtis, K., Glass, E., Godfrey, E., Marcell, A., Mautone-Smith, N., Pazol, K., Tepper, N., Zapata, L., & Centers for Disease Control and Prevention (CDC) (2014). Providing quality family planning services: Recommendations of CDC and the U.S. Office of Population Affairs. MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 63(RR-04), 1–54.

[12] Curtis, K. M., Tepper, N. K., Jatlaoui, T. C., Berry-Bibee, E., Horton, L. G., Zapata, L. B., Simmons, K. B., Pagano, H. P., Jamieson, D. J., & Whiteman, M. K. (2016). U.S. Medical Eligibility Criteria for Contraceptive Use, 2016. MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 65(3), 1–103. https://doi.org/10.15585/mmwr.rr6503a1

[13] Gomez, A. M., & Wapman, M. (2017). Under (implicit) pressure: young Black and Latina women´s perceptions of contraceptive care. Contraception, 96(4), 221–226. https://doi.org/10.1016/j.contraception.2017.07.007

[14] Finer, L.B. and Zolna, M.R. (2016). Declines in unintended pregnancy in the United States, 2008-2011. New Engl J Med, 374(9), 843-52

[15] Thoma, M. E., Copen, C. E., & Kirmeyer, S. E. (2016). Short Interpregnancy Intervals in 2014: Differences by Maternal Demographic Characteristics. NCHS data brief, (240), 1–8.

[16] Sundaram, A., Vaughan, B., Kost, K., Bankole, A., Finer, L., Singh, S., & Trussell, J. (2017). Contraceptive Failure in the United States: Estimates from the 2006-2010 National Survey of Family Growth. Perspectives on sexual and reproductive health, 49(1), 7–16. https://doi.org/10.1363/psrh.12017

[17] Gavin, L., Frederiksen, B., Robbins, C., Pazol, K., & Moskosky, S. (2017). New clinical performance measures for contraceptive care: their importance to healthcare quality. Contraception, 96(3), 149–157. https://doi.org/10.1016/j.contraception.2017.05.013

[18] Gavin, L. E., Ahrens, K. A., Dehlendorf, C., Frederiksen, B. N., Decker, E., & Moskosky, S. (2017). Future directions in performance measures for contraceptive care: a proposed framework. Contraception, 96(3), 138–144. https://doi.org/10.1016/j.contraception.2017.06.001

[19] Moniz, M. H., Gavin, L. E., & Dalton, V. K. (2017). Performance Measures for Contraceptive Care: A New Tool to Enhance Access to Contraception. Obstetrics and gynecology, 130(5), 1121–1125. https://doi.org/10.1097/AOG.0000000000002314

S.4. Numerator Statement: Women ages 15-44 at risk of unintended pregnancy who are provided a most (sterilization, intrauterine device, implant) or moderately (injectable, pill, patch, ring) effective method of contraception.

S.6. Denominator Statement: Women ages 15-44 who are at risk of unintended pregnancy.

S.8. Denominator Exclusions: The following categories of women are excluded from the denominator: (1) those who are infecund for non-contraceptive reasons; (2) those who had a live birth in the last 2 months of the measurement year; or (3) those who were still pregnant or their pregnancy outcome was unknown at the end of the measurement year.

De.1. Measure Type: Outcome: Intermediate Clinical Outcome

S.17. Data Source: Claims

S.20. Level of Analysis: Clinician : Group/Practice, Facility, Health Plan, Population : Regional and State

IF Endorsement Maintenance – Original Endorsement Date: Oct 25, 2016 **Most Recent Endorsement Date:** Oct 25, 2016

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? Although not a requirement, two other measures have been submitted for maintenance endorsement in separate applications that are complementary to this measure and – if reported together – would provide a broad perspective on the quality of contraceptive services. The two other measures are focused on:

- Postpartum women – this is a very important sub-population of all women at risk of unintended pregnancy. It has been proposed as a separate measure because of the unique need of this population for birth spacing, and the need to raise awareness so that opportunities are not missed to provide contraceptive services during pregnancy, at delivery and in the postpartum period.
- Long-acting reversible contraceptive methods (LARC) – the LARC methods of intrauterine devices (IUD) and implants are a very important sub-set of all contraceptive methods that have extremely low failure rates. The primary goal of this measure is to monitor whether women have access to LARC methods as determined by whether any health facilities or other reporting units report very low levels of LARC use (e.g., less than 1-2 percent) or at a level that is substantially below the median when compared to other reporting units.

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measure still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality.

Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

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

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

The developer provides the following evidence for this measure:

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

Evidence Summary or **Summary of prior review in [2016]**

- The developer provided robust summaries of clinical practice guideline recommendations and other SRs. This evidence included data developed through randomized control trials (RCTs) and meta-analyses for the most effective (i.e., sterilization, implants, intrauterine devices or systems (IUD/IUS)) or moderately effective (i.e., injectables, oral pills, patch, or ring) method of contraception. The developer reported that the evidence showed support for contraceptive effectiveness and its impact on unintended pregnancies.

Changes to evidence from last review

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

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

Updates:

- The developer cited a robust number of guidelines and a conceptual framework in support of the measure. These included guidelines from the Centers for Disease Control (CDC), the U.S. Office of Population Affairs (OPA), American College of Obstetricians and Gynecologists (ACOG), and the Health Resources and Services Administration (HRSA).
    - To support the division of the measure into the two sub-measure rates, most and moderate effective methods of contraceptives for women aged 15-44 years at risk of unintended pregnancy.
    - The use of a diaphragm was removed from the moderate effective contraceptive list.

Exception to evidence

- Does the Committee want to discuss how patient choice for no, over the counter (OTC), or lower effective contraceptives are captured in the measure?
- Does the evidence support excluding deliveries that did not end in a live birth (i.e., miscarriage, ectopic, stillbirth or induced abortion) for #2902 and not #2903, or patients with live or not live births in the last two months of the measurement period where contraceptives may be applicable?

Questions for the Committee:

- The evidence provided by the developer is updated, directionally the same, and stronger compared to that for the previous NQF review. Does the Committee agree there is no need for repeat discussion and vote on Evidence?
- Is the evidence directly applicable to the process of care being measured?
- Does the Committee agree with removing diaphragm from the list of moderate contraceptive methods?
- If derived from patient report, does the target population value the measured process or structure and find it meaningful?

Guidance from the Evidence Algorithm

Measure does not assess a health outcome or PRO (Box 1) → Measure assesses an intermediate clinical outcome based on an SR and grading of the evidence (Box 3) → A summary of the quantity, quality, and consistency (QQC) of the body of evidence is provided (Box 4) → The summary includes high quality, quantity, and consistency of evidence are high and the net benefit is substantial and outweighs undesirable effects (Box 5a) → High

The highest possible rating is high.

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

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

Maintenance measures – increased emphasis on gap and variation

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

- Performance data was provided for the following levels of analysis: Clinician group/practice, Facility, Health Plan, Public Health Region, and State from nine different programs. For example:
    - Centers for Medicaid & Medicare Services (CMS): Maternal and Infant Health Initiative, Core Measure Set
        - FFY 2016 Median Measure Scores: Ages 15-20: 30.5 and Ages 21-44: 26.3
        - FFY 2017 Median Measure Scores: Ages 15-20: 30.8 and Ages 21-44: 25.6
        - FFY 2018 Measure Scores Ages 15-20 Median: 28.1, Range: 7.6 – 39.0
        - FFY 2019 Measure Scores Ages 15-20 Median: 29.5, Range: 1.4 – 98.0
- Performance scores are not reported by moderate and most, rather as overall median or mean performance. Although #2903 has been adopted into CMS' Adult and Child Core Set, the measure performance for adult women ages 21-44 have not yet been reported because fewer than 25 states have reported the measure. In FFY 2018, #2903 were reported for the first time in the Child Core Set for women ages 15-20 and then again in FFY 2019.

- See the testing attachment for other performance gap data. Depending on the sample size, significant differences are noted in overall median and mean performance, as well as larger standard deviations and ranges.

Disparities

- A 2015-2017 National Survey of Family Growth (NSFG) study examined contraceptive use among women who were at risk of unintended pregnancy because they had ever had sex, were fecund, and were neither pregnant nor seeking pregnancy found that 51.7% of adolescents and 60.8% of adult women used a most or moderately effective method. 2015-2017.
- The Planned Parenthood Federation of America (PPFA) final dataset analyzed included 123,978 female patients aged 15-44 years, who received services from two PPFA affiliates between January 1 and December 31, 2019. Performance by race and ethnicity included African American: 53.50, Alaskan Native: 64.87, Asian: 68.23, Hispanic: 66.27, Multi-racial: 64.64, Native American: 59.83, Pacific Islander: 65.18, White: 66.53, and Other race: 58.49.
- The 2014-2018 Washington State Health Care Authority (WA HCA) reported for female clients ages 15-44 by age group and race/ethnicity (https://www.hca.wa.gov/assets/program/ccw-contraceptive-care.pdf). The percentages of 2018 women aged 15-20 were provided most and moderately effective methods by race/ethnicity remained stable over these five years: Hispanic: 24.4, White: 37.2, Asian: 19.4, Black: 24.5, American Indian/Alaska Native: 33.7, Hawaiian/Pacific Islander: 18.9, More than One Race: 34.9, and Other/Unknown: 23.7. For women aged 21-44 years, race and ethnicity findings were Hispanic: 33.1, White: 27.0, Asian: 26.0, Black: 26.1, American Indian/Alaska Native: 24.6, Hawaiian/Pacific Islander: 23.6, More than One Race: 29.9, and Other/Unknown: 26.9.

Questions for the Committee:

- Was the performance gaps and disparities data available by moderate and most effective contraceptive method to more clearly identify and target quality improvement activities?
- Is there a gap in care that warrants a national performance measure?

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: For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures–are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure.

- strong evidence
- High evidence

- Is there evidence to support excluding pregnancies ending in a live birth but not pregnancies that ended in other ways? It seems reasonable to treat all pregnancies the same, either include or exclude all, regardless of how they ended. 2) I agree with removing diaphragm from the list of moderately effective methods.
- The evidence showed support for contraceptive effectiveness. Not aware of additional studies.
- Yes

1b. Performance Gap: Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?

- significant gaps with racial disparities
- large gaps in outcomes, high opportunity for improvement
- I am not clear that the measure is able to adequately differentiate performance gaps and disparities vs. differences in patient choice due to cultural and other factors.
- The data demonstrates there is a gap in care. Disparity data is available.
- yes

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.

Composite measures only:

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: NQF Scientific Methods Panel Subgroup 3

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

- The developer states that *#2902* Contraceptive Care Postpartum and *#2904* Contraceptive Care - Access to LARC are complementary measures to this measure. The developer excludes patients with a pregnancy that did not end with a live birth in *#2902*, but not *#2903* and *#2904*. The developer emphasizes the measure is not to be used in pay for performance programs.
- Reliability testing was conducted at the measure score level. Data element validity testing was conducted; therefore, additional data element reliability testing is not required.
- The measure level of analysis includes the following levels: Clinician: Group/Practice, Facility, Health Plan, Population: Regional and State. Reliability testing is provided in state-level payer programs, although not all-payer state programming.
- Several reviewers had concerns regarding performance not being measured in the last two months of the year and could disincentivize positive performance.
- Using the beta-binomial model and the parametric empirical Bayes methods (which is appropriate for the measure), measure score reliability was calculated in signal-to-noise analyses for all four levels: Clinician: Group/Practice, Facility, Health Plan, Population: Regional and State.
- Claims data from seven organizations were utilized for testing: Iowa Medicaid Enterprise (2018), Iowa Department of Public Health (IDPH) (2019), NewYork Presbyterian Hospital/Columbia University Irving Medical Center (2018), Washington State Health Care Authority (2019), Massachusetts Mass Health (2019), Oregon Medicaid (2015) and Louisiana Medicaid Program (2019).
- Planned Parenthood Federation of America (2019) and Title X Family Planning Program (2019) were also included using different calculations and interpretations as the patient population is women seeking reproductive care.
- Reliability scores were very high at all testing levels, except the group level. Many reviewers prefer case limits, such as the 75 case counts obtained at group level, especially in high stakes program use. Targets greater than 0.90 may be used for high-stake purposes and greater than 0.70 used for reporting and monitoring. The developer emphasizes the measure is not to be used in pay for performance programs.

Validity

- Validity testing was conducted at the measures score and data element levels. Measure score validity testing was not conducted for health plans as populations as the limited numbers of units for these levels were not sufficient for correlation testing.
- The developer performed construct validity testing of the measure to (1) Cervical Cancer Screening, (2) Chlamydia Screening, (3) Encounter for Contraceptive Counseling, and (4) Encounter for Gynecological Exam Measures, hypothesizing measured entities performing well on contraceptive care should perform well on the other measures, and correlation magnitudes may be weak for cervical cancer and chlamydia screenings with screening frequency differences.
- Pearson correlations and a novel multilevel correlation estimation method (due to low volume events in high volume populations) were used with thresholds of 25, 50, and 75 eligible patients. The novel approach generally showed slightly higher or similar correlations to Pearson's for Contraceptive Counseling and Gynecological Examination measures in group reporting with moderate reliability. The Cervical Cancer Screening and Chlamydia Screening measures generally showed slightly higher or the same correlations to Pearson's than the novel approach, except 21-44 in Chlamydia Screening. The submitted measure showed "just" to poor reliability for these two measures. As predicted, the correlations were weak to none in the Planned Parenthood Federation of America in Cervical Cancer Screening and Chlamydia Screening measures possibly due to screening frequency differences.
- Data element validity testing was conducted with 423 patients, compared claims vs. patient record for 10 critical data elements in calculated sensitivity, specificity, PPV, NPV, Cohen's Kappa statistics with 95 percent CIs, and percent agreement for each data element. Sensitivity was above 0.5 for most data elements, except the contraceptive patch, in which specificity, PPV, and NPV were greater than 0.8 for all data elements. Percent agreement was greater than 80 percent for all data elements.
- Reviewers were concerned about sensitivity results being "less than desirable," specifically with the contraceptive patch of 0.25 used to define numerator.
- Face validity was conducted with nine independent panel experts to assess whether the measure will reflect quality of contraceptive care. The mean rating measure was 4.67 with a median of 5 (Strongly Agree), range 4-5. One reviewer was "unclear on patient-centeredness of this overall (face validity)".
- The measure is not risk-adjusted, yet it is stratified by adolescents and adults. Multiple reviewers had concern with the lack of social risk stratification. The developer stated, "statistically significant differences by age group (for ages 20-29 compared to ages 30-44) and among women who have never been married (compared to women of other marital status", were identified, yet "no significant differences occur between race/ethnicity, most categories of marital status, and poverty level" were seen. These findings contrast the identified disparities from measure #2902 with overlapping populations.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure about the lack of minimum sample size (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 construct validity testing of the measure?
- Do you have any concerns regarding the exclusions in the measure?
- The Scientific Methods Panel is satisfied with the validity analyses for the measure. Does the Committee think there is a need to re-vote on validity?

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

Committee Pre-evaluation Comments:

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

2a1. Reliability-Specifications: Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?

- no concerns
- None
- As noted above, I have concerns regarding the exclusions in this measure. I have no other concerns.
- No concerns
- None

2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure?

- no concerns
- No concerns
- No concerns.
- No concerns
- None

2b1. Validity - Testing: Do you have any concerns with the testing results?

- no concerns
- No concerns
- No concerns.
- no, validity testing showed >80%
- none

2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)

2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do

analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?

- no concerns
- Concerns with data not showing differences among race and ethnicity
- The measure is able to distinguish between facilities with higher and lower scores. Interpretation of these results needs to be in context and in conjunction with other measures, especially NQF #3543 (Person-Centered Contraceptive Counseling).
- Claims data so no concerns
- None

2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment) 2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure? 2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?

no concerns-

Concerns with data not showing differences among race and ethnicity

Concerns regarding exclusions/inclusions of pregnancies regardless of how they end as noted above.

the developer provider rationale for no risk adjustment, however, it may be beneficial to test and learn.

Not sure why, with such big gaps, we are splitting hairs with the type of BC prescribed. Would combine LARC and moderate methods. Allows for more patient choice and takes away some potential downsides to the measure.

Criterion 3. Feasibility

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

3. **Feasibility** is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.
    - The developer reports that the measure is coded by someone other than the person obtaining the original information.
    - The developer reports that all data elements are in defined fields in electronic administrative claims. The developer also reports that there is ongoing work with UCSF to develop an eCQM version of this measure.
    - The measure developer participated in a MIHI grant program to develop the measure and identified several important lessons from this collaborative work:

- The co-design process for measure development increased feasibility of the measure.
- Measure users found calculation of the measure time-consuming. Technical assistance is available from OPA for measure users, and OPA is exploring ways to improve efficiency.

Questions for the Committee:

- Do you have concerns about the measure users' experiences with calculating the measure?

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

Committee Pre-evaluation Comments:

Criteria 3: Feasibility

3. **Feasibility: Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?**

- highly feasible
- Claims based data, so no concerns
- Measure users found calculation of the measure time-consuming, which could lead to low use of the measure, despite TA from OPA available.
- No concerns
- none

Criterion 4: Usability and Use

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

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

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

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

Current uses of the measure

Publicly reported? ☒ Yes ☐ No

Public Reporting

- CMCS Core Set of Adult and Child Health Care Quality Measures https://www.medicaid.gov/medicaid/quality-of-care/performance-measurement/adult-and-child-health-care-quality-measures/index.html
- Iowa Medicaid Enterprise, https://dhs.iowa.gov/ime/members/medicaid-a-to-z

- Louisiana Medicaid, https://qualitydashboard.ldh.la.gov/
- MassHealth, https://www.mass.gov/orgs/masshealth
- New York-Presbyterian Hospital/Columbia University Irving Medical Center Ambulatory Care Network, https://www.nyp.org/acn
- Washington State Health Care Authority, https://www.hca.wa.gov/about-hca/reproductive-health
- Title X Family Planning Program, https://rhntc.org/resources/contraceptive-access-change-package
- Title X Family Planning Program, https://opa.hhs.gov/evaluation-research/title-x-services-research/family-planning-annual-report

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

Accountability program details

See Public Reporting details for more Accountability information.

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 is presented for current use in eight programs: federal Medicaid efforts to publicly report and support state use of the measures; four state Medicaid programs (i.e., the Iowa Medicaid Enterprise, the Washington State Health Care Authority, Louisiana Medicaid, and MassHealth); and one outpatient clinic network within an academic health system (NewYork-Presbyterian Hospital/Columbia University).
- The developer also provides program data from two national organizations that focus on the delivery of reproductive health services (i.e., the Planned Parenthood Federation of America and the Title X program). Feedback from these programs have significantly contributed to updates for the measure.
- OPA has published multiple peer-reviewed articles on the appropriate implementation and use of the measure.
- OPA publishes information on its website to help implementors appropriately use and understand the limitations of the measure.
- OPA manages two email addresses to field questions from measure users. CMS and NCQA also forward questions that they receive to these addresses. As a contractor, Mathematica Policy Research also collects feedback and answers user questions.
    - Questions have included input on various unexpected issues with certain coding systems, how to deal with states' differences in coding systems, and recommendations for stratification of the measure.

Additional Feedback:

- The measure has been included in the Core Quality Measures Collaborative (CQMC) Consensus Core Set: Obstetrics and Gynecology

***Questions for the Committee**:*

- How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?

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

- Since 2015, OPA has been the recipient of on-going feedback on NQF #2903 through CMS. CMS has a contract with Mathematica Policy Research to provide technical assistance (TA) on states reporting NQF #2902, NQF #2903, and NQF #2904 for the CMS Adult and Child Core sets.
- Performance improvements have found the provision of most or moderately effective methods to be about 24% in states with Medicaid expansion and 20% in non-expansion states, and an approximate 35-percentage point opportunity for improvement. A more realistic improvement opportunity is reported between 15-20 percentage points as 100% performance should never be anticipated for this measure concept.

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

- No unexpected findings have been reported since initial endorsement.

Potential harms

- The developer reports that they remind measure users of the potential for coercive care practices in response to this measure. Measure users should not strive for a particular benchmark.
- Although not yet tested in pregnant patients, the developer believes that use of balancing measure #3543 will promote person-centered contraceptive care and post-partum LARC utilization. The developer reports that research in the pregnant population is warranted.

Additional Feedback

- The measure has been included in the Core Quality Measures Collaborative (CQMC) Consensus Core Set: Obstetrics and Gynecology

Questions for the Committee:

- 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: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided? 4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?

- accountable- yes!
- Publicly reported and in accountability programs
- The measure developer has provided significant opportunities for feedback from users.
- feedback has been considered and incorporated.
- same questions re overzealous prescribing

4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations? 4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.

- highly usable
- Some concerns about targeted populations
- Additional guidance may be needed to learn how this measure can be used to set performance improvement targets.
- Benefits outweigh the unintended negative consequences.
- need a better balancing measure

Criterion 5: Related and Competing Measures

Related or competing measures

- 1517: Prenatal & Postpartum Care (PPC)
- 2902: Contraceptive Care - Postpartum
- 2904: Contraceptive Care - Access to LARC
- 3543: Person-Centered Contraceptive Counseling (PCCC) measure

Harmonization

- The developer reports that these related measures are harmonized to the extent possible.
- Namely, measures #2902 and #2904 are complementary to this measure.
    - #2902 focuses on most or moderately effective contraceptive provision in all women of who had a live birth
    - #2904 focuses on LARC provision only in all women of reproductive age

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

5. Related and Competing: Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?

- 1517, 2903, 2904
- No concerns
- Are there ways that this measure can be harmonized with the #3543: Person-Centered Contraceptive Counseling (PCCC) measure?
- No concerns, related measures are harmonized.
- yes other BC measures being reviewed

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 06/29/2021

- No NQF Members have submitted support/non-support choices as of this date.
- No Public or NQF Member comments submitted as of this date.

Combined Methods Panel Scientific Acceptability Evaluation

Measure Number: 2903

Measure Title: Contraceptive Care – Most & Moderately Effective Methods

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: No concerns.

   Panel Member 3: No concerns except that the following attachment was not found: NQF_2903_Codes_2021-637453719019907247.xlsx

   Panel Member 4: No code spreadsheet provided.

   Panel Member 5: No concerns

   Panel Member 6: None

   Panel Member 7: When does specification overlap with adequate demonstration of harmonizing with related measures?

   Panel Member 8: None

RELIABILITY: TESTING

Type of measure:

☐ **Outcome (including PRO-PM)** ☒ **Intermediate Clinical Outcome** ☒ **Process**

☐ **Structure** ☐ **Composite** ☐ **Cost/Resource Use** ☐ **Efficiency**

Data Source:

☐ **Abstracted from Paper Records** ☒ **Claims** ☐ **Registry**
☐ **Abstracted from Electronic Health Record (EHR)** ☐ **eMeasure (HQMF) implemented in EHRs**
☐ **Instrument-Based Data** ☐ **Enrollment Data** ☐ **Other (please specify)**

Panel Member 7: Chart abstracts from clinical records for data element validity testing

Level of Analysis:

☐ **Individual Clinician** ☒ **Group/Practice** ☒ **Hospital/Facility/Agency** ☒ **Health Plan**
☒ **Population: Regional, State, Community, County or City** ☐ **Accountable Care Organization**
☐ **Integrated Delivery System** ☐ **Other (please specify)**

Panel Member 3: Public health region
Panel Member 4: Public Health Region

Measure is:

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

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

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

   ☐ **Yes** ☐ **No**
6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

   Panel Member 1: Used appropriate methods. Estimated a signal-to-noise ratio statistic from a Beta-binomial model using parametric empirical Bayes methods. Statistic was calculated for each level of analysis.

   Panel Member 2: The developer evaluated measure score reliability at measure entity level via a beta-binomial model using empirical Bayes methods, which is appropriate for this measure.

   Panel Member 3: No concerns

   Panel Member 4: Beta-binomial model using parametric empirical Bayes methods.

Panel Member 5: Reliability was estimated from a Beta-binomial model using parametric empirical Bayes methods. Two distributional shape parameters (alpha and beta) were estimated from the observed quality scores, and reliability was then calculated as a function of alpha, beta, and total patient count for each unit of analysis. Overall reliability in this context represents the ability of the proposed measure to confidently distinguish the performance of one entity (e.g., facility) from another.

Panel Member 6: Claims from seven organizations were utilized for testing: Planned Parenthood Federation of America, Iowa Medicaid Enterprise, Iowa Department of Public Health, NewYork Presbyterian Hospital/Columbia University Irving Medical Center, Washington State Health Care Authority, Massachusetts Mass Health, and the Louisiana Medicaid Program. This covered the period between January 1, 2019 to December 31, 2019 for most of them, and January 1, 2018 to December 31, 2018 for two. Reliability of the measure is assessed at facility, group billing provider, public health region, and health plan levels. A signal to noise approach was utilized from a beta binomial model. A cutoff reliability of greater than .90 for making high-stakes decisions and greater than .70 for general reporting/monitoring was targeted.

Panel Member 7: OK: , reliability was estimated from a Beta-binomial model using parametric empirical Bayes methods. Two distributional shape parameters (alpha and beta) were estimated from the observed quality scores, and reliability was then calculated as a function of alpha, beta, and total patient count for each unit of analysis.

Panel Member 8: A signal-to-noise (SNR) method was used to assess reliability at the facility level. It would have been good perhaps to also include a split-sample or stability of classification (e.g., deciles) analysis.

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

   Submission document: Testing attachment, section 2a2.3

   Panel Member 1: The reliability statistics are consistently greater than .70 at the facility, public health region, and health plan levels, showing adequate to high reliability at these levels.

   Panel Member 2: Reliability scores were very high at all testing levels with the exception of group billing provider level. With a 75 case counts limit, similar reliability scores were obtained at group billing provider level. If this measure is to be used for measure entities with low case counts, it will be useful to establish a volume threshold for reporting.

   Panel Member 3: No concerns

   Panel Member 4: Appropriate

   Panel Member 5: Tested reliability is consistently greater than .70 at the facility, public health region, and health plan levels, showing adequate to high reliability at these levels.

   Panel Member 6: Beta-binomial reliability estimates were greater than .75 for all levels with the exception of the group billing provider for all ages (N single digit). The requirement of a unit size greater than 75 resulted in reliability of greater than .75 for all levels.

   Panel Member 7: OK - depending on N, level of analysis

   Panel Member 8: In general, and especially for entities with>75 women, reliabilities were high. It would have been helpful to have a fuller description of the distribution of reliabilities.

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

Submission document: Testing attachment, section 2a2.2

☒ **Yes**

☐ **No**

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

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

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

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

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

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

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

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

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

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

    Panel Member 1: Used appropriate methods for testing. Calculated reliability statistics all demonstrated adequate to high reliability.

    Panel Member 2: Most reliability scores were above 0.9 for different measure levels and both age strata.

    Panel Member 4: Seventy-five patients needed when measuring the facility and provider level.

    Panel Member 5: Tested reliability is consistently greater than .70 at the facility, public health region, and health plan levels, showing adequate to high reliability at these levels.

    Panel Member 6: For almost all levels of analysis, the signal to noise from a beta binomial model was at least .70

    Panel Member 7: Unclear to me if theoretical construct makes sense (regardless of empiric result). Mod-Low.

    Panel Member 8: Especially for entities with>75 women, reliabilities were high

VALIDITY: TESTING

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

13. **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)

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

☒ **Face validity**

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

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

15. **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)

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

Submission document: Testing attachment, section 2b2.2

Panel Member 1: Data element: For 423 patients, compared claims vs. patient record for 10 critical data elements Face validity for score: Used a panel of 9 independent individuals to assess whether the measure will reflect quality of contraceptive care Empirical validity for score: Compared performance on the contraceptive care measure to other measures of women's health services; hypothesis was performance would move in the same direction.

Panel Member 2: For measure score validity, the developer correlated this measures with four other related measures and provided conceptual reasons for expected findings. The developer also assessed data element validity for a few critical data elements following a commonly used approach.

Panel Member 3: No concerns

Panel Member 4: Appropriate.

Panel Member 5: For score level convergent validity of the most or moderately effective contraceptive measure by exploring whether it was correlated with other similar quality measures. For data element data elements used for contraceptive care measure calculations were compared between the claims records and the patient charts, and agreement numbers were summarized in a 2 by 2 table (yes/yes, yes/no, no/yes, and no/no) for each element. We compared 10 data elements in total, including 7 most or moderately effective methods (Female sterilization, Implantable, IUD, Injectables, Contraceptive pills, Contraceptive patch, and Vaginal ring) and 3 exclusion criteria elements (Infecund, Currently pregnant or unknown pregnancy outcome, and Live births in the last 2 months of the year). Using the patient chart as the authoritative source, we calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), Cohen's Kappa statistics

Panel Member 6: Empirical face validity testing was performed by correlation with other quality measures pertinent to this group, specifically cervical cancer screening, chlamydia screening, an encounter for contraceptive counseling, and encounter for a gynecologic exam. It was hypothesized that providers who performed well on these measures would also perform well on the current measure. First, a Pearsons correlation test was performed. Secondly, to reduce the impact of non-linearity on the Pearson, a multilevel correlation estimation methodology was employed which

involved the use of a logit transformation of the binomial model framework. Data element validity testing was assessed by sensitivity, specificity, PPV, NPV, %agreement, and kappa.

Panel Member 7: Correlations between the contraceptive care measure with contraceptive counseling and gynecological exam measures at both facility and group billing provider levels among the 15-44 age group

Panel Member 8: Correlation analyses (both standard and improved) of the measure with similar measures. They hypothesized that facilities/providers that perform well on this measure should perform well on other contraceptive care measures, and less correlated with cervical cancer screening and chlamydia screening.

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

Submission document: Testing attachment, section 2b2.3

Panel Member 1: Data element: Sensitivity was above 0.5 for the majority of the data elements, except for contraceptive patch, whereas specificity, PPV, and NPV were above 0.8 for all data elements. Percent agreement was consistently over 80% for all data elements. Face validity: The mean rating from the face validity assessment for this measure was 4.67 with a median of 5 (Strongly Agree), range 4-5. Empirical score validity: Saw moderate correlations with performance on the contraceptive counseling and gynecological exam measures.

Panel Member 2: The developer found positive correlations between this measures and three other measures as expected. Lack of significant correlation with Chlamydia screening was not unexpected. It might be informative to include measures 2902 and 2904 for this analysis. Sensitive results were less than desirable. For example, for contraceptive patch, it was 0.25. This is concerning that it is used to define numerator of this measure.

Panel Member 3: Data element validity results were satisfactory. Empirical validity results were as expected, i.e., with weak to moderate correlations with related measures and in the expected direction.

Panel Member 4: Acceptable.

Panel Member 5: Empirical validity testing 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. Using the multilevel correlation estimation method, we observed mostly moderate to strong positive correlations between the contraceptive care measure with contraceptive counseling and gynecological exam measures at both facility and group billing provider levels among the 15-44 age group. Pearson's correlation test showed similar positive correlations except for a non-significant correlation with contraceptive counseling. We also found positive associations among the sub-age groups with contraceptive counseling and gynecological exam, although some of the associations were not statistically significant, likely due to smaller number of units in the analysis. For cervical cancer screening, both methods showed positive correlations, although the correlation was not statistically significant at the facility level when using the multilevel correlation estimation. For chlamydia screening, we did not observe any statistically significant associations at either facility or group billing provider levels. Critical data elements Sensitivity was above 0.5 for the majority of the data elements, except for contraceptive patch, whereas specificity, PPV, and NPV were above 0.8 for all data elements. Percent agreement was consistently over 80% for all data elements. We also observed statistically significant Kappa above 0.6 for all data elements except for contraceptive patch, indicating moderate to almost perfect agreement between the claims records and the patient charts (Watson and Petrie, 2010). Overall,

our data provide fairly strong evidence for validity of the contraceptive care measure at the data element level.

Panel Member 6: At the facility level, Pearson correlation with other quality measures was .09 to .40 (the latter was best for chlamydia screening) with a multilevel correlation estimation of -.13 to .55). At the group provider billing level, Pearson correlation with other quality measures was .23 to .59 (the latter was best for contraceptive counseling) with a multilevel correlation estimation of .21 to .63). At the group provider billing level, Pearson correlation with other quality measures was .23 to .59 (the latter was best for contraceptive counseling) with a multilevel correlation estimation of .21 to .63). It is noted that while the correlation improved for some usage of the multilevel estimate, for others it decreased and in a couple of circumstances, actually flipped from positive to negative or vice-versa. Kappa's for the data element validity testing ranged from . 398 (contraceptive patch) to the mid-70's to mid-80's for most of the data elements.

Panel Member 7: OK

Panel Member 8: The results generally support the hypotheses, in some cases quite strongly.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

Submission document: Testing attachment, section 2b2.

Panel Member 1: No concerns. The frequency of exclusions for the datasets analyzed is low.

Panel Member 2: The sensitivity for live birth was low. Because this variable is used to define an exclusion criterion, this raises the concern of failing to exclude cases that should have been excluded. Different levels of exclusion based on this variable were found among different data sources were also concerning.

Panel Member 3: I have the same concern raised for measure 2902 related to the exclusion of those who had a live birth in the last 2 months of the measurement year. This could potentially cause a lower incentive to achieve a successful score for these women. A simple date adjustment could be considered to avoid the exclusion of 2/12 months of data, as proposed for measure 2902. Additionally, no testing was conducted to assess how this exclusion criteria impacted the group level scores. It would be helpful to add such analysis to this submission.

Panel Member 4: Frequency of exclusions for the datasets analyzed is low.

Panel Member 5: None

Panel Member 6: Exclusions are clearly defined and the rationale provided. The most common exclusion was pregnant or pregnancy outcome was unknown at the end of the measurement period (4.4%).

Panel Member 8: "Women with live births that occurred in the last 2 months of the measurement year might not have had a chance to receive postpartum contraceptive care in the 60-day time frame and were therefore excluded." This exclusion contradicts and makes this measure distinct from the post-partum measure, and builds in a dependence on that measure to get the full population.

19. **Risk Adjustment**

Submission Document: Testing attachment, section 2b3

19a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

19b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☒ Yes ☒ No ☒ Not applicable

19c. **Social risk adjustment:**

19c.1 Are social risk factors included in risk model? ☒ Yes ☒ No ☒ Not applicable

19c.2 Conceptual rationale for social risk factors included? ☒ Yes ☐ No

19c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☒ Yes ☐ No

19d. **Risk adjustment summary:**

19d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

19d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

19d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

19d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

19d.5. Appropriate risk-adjustment strategy included in the measure? ☒ Yes ☒ No

19e**. Assess the risk-adjustment approach**

Panel Member 1: The developers offer a rationale for not adjusting performance (i.e., variation exists due to modifiable clinical and programmatic considerations, not patient-level factors).

Panel Member 2: The developer provided rationale why this measure should not be risk adjusted.

Panel Member 3: I have the same concerns about lack of risk adjustment as mentioned form measure 2902.

Panel Member 4: Acceptable justification and no contrary evidence to developer's rationale.

Panel Member 5: No risk adjustment but authors recommend stratifying by age group so that measure scores for adolescent and adult women can be calculated separately for quality improvement (QI) purposes.

Panel Member 6: Risk adjustment was felt to be not justified, not because socio-economic differences do not exist, but are system driven, not biologically drive.

Panel Member 7: "Correlations between the contraceptive care measure with contraceptive counseling and gynecological exam measures at both facility and group billing provider levels among the 15-44 age group" __> May be worth trialing risk adjustment/SES adjustment to demonstrate no difference.

Panel Member 8: This is a process measure and not risk adjusted. The results are stratified by age on a well-justified conceptual basis.

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

Submission document: Testing attachment, section 2b4.

Panel Member 1: No concerns. Measure rates vary considerably across levels, with reduced variation in the "larger" levels (health plans, population).

Panel Member 2: The range of measure scores was quite wide, indicating substantial variation among measure entities.

Panel Member 3: No concerns

Panel Member 4: No concerns.

Panel Member 5: No concerns

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

    Submission document: Testing attachment, section 2b5.

    Panel Member 1: N/A

    Panel Member 2: When data element reliability may vary across data sources, we should assess how this may affect the measure scores before conducting any comparisons across data sources.

    Panel Member 4: n/a

    Panel Member 6: Not applicable

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

    Submission document: Testing attachment, section 2b6.

    Panel Member 1: No concerns. The measure is based on claims data, which has low rates of missing data.

    Panel Member 3: No concerns

    Panel Member 4: No concerns.

    Panel Member 5: No concerns

    Panel Member 6: None

For cost/resource use measures ONLY:

23. **Are the specifications in alignment with the stated measure intent?**

    ☐ **Yes** ☐ **Somewhat** ☐ **No (If "Somewhat" or "No", please explain)**

24. **Describe any concerns of threats to validity related to attribution, the costing approach, carve outs, or truncation (approach to outliers):**

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

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

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

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

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

26. **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: The developers tested data element validity and two types of score-level validity. All of the testing results support this is a valid measure of performance (data elements are valid + score correlated with "like" measures).

Panel Member 2: Data element validity for a couple of critical data elements (e.g., contraceptive patch and live birth in the last 2 months) is concerning.

Panel Member 4: Reasonably can reliably distinguish facilities performance. Used an expert panel to discuss appropriate measure use and interpretation

Panel Member 5: Adequate agreement at the data level. Empirical score level was weaker but acceptable at most levels

Panel Member 6: Validity testing for correlation with other quality measures was generally positive in the .15 to .35 range. The use of an alternative approach to estimation of the correlation did not improve the ability to demonstrate a significant correlation.

Panel Member 7: Could be Moderate. Unclear on patient-centeredness of this overall (face validity). This may not be a matter for the SMP. I defer.

Panel Member 8: This appears to be a population access measure more than a typical process measure or intermediate clinical outcome measure. The extent to which the numerator is accurately captured is probably approximate and may vary to an unknown degree between entities. But overall, the validity and validity testing appear sound

FOR COMPOSITE MEASURES ONLY: Empirical analyses to support composite construction

27. **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**

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

ADDITIONAL RECOMMENDATIONS

29. **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: None.

Developer Submission

NQF #: 2903

Corresponding Measures:

De.2. Measure Title: Contraceptive Care – Most & Moderately Effective Methods

Co.1.1. Measure Steward: HHS Office of Population Affairs

De.3. Brief Description of Measure: The percentage of women aged 15-44 years at risk of unintended pregnancy that is provided a most effective (i.e., sterilization, implants, intrauterine devices or systems (IUD/IUS)) or moderately effective (i.e., injectables, oral pills, patch, or ring) method of contraception.

The measure is an intermediate outcome measure because it represents a decision that is made at the end of a clinical encounter about the type of contraceptive method a woman will use, and because of the strong association between type of contraceptive method used and risk of unintended pregnancy.

1b.1. Developer Rationale: Unintended pregnancies and interpregnancy intervals of less than 18 months have been associated with poor perinatal outcomes such as preterm birth, low birth weight, small size for gestational age, as well as adverse maternal outcomes [1, 2]. Studies among U.S. women report that women at younger maternal age are at higher risk for unintended pregnancy [14] and older maternal age is associated with closely spaced pregnancies [15]. Contraception is a highly effective clinical preventive service that can assist women in reaching their reproductive health goals, like reducing unintended pregnancies and the percentage of births occurring within 18 months of a previous birth [3, 4]. The type of contraceptive method used by a woman is strongly associated with her risk of unintended pregnancy. The most effective methods (LARC and sterilization) have a failure rate that is less than 1% per year under typical use [4]. The moderately effective methods (injectable, pill, patch, ring) have a typical failure rate of 4-7% per year, while the less effective methods have a typical failure rate of 13-27% [4]. One recent study also indicates that the most used contraceptive methods in the United States have experienced reductions in their typical use failure rates [16]. Not using any method at all has a typical failure rate of 85% [4].

After NQF endorsed #2903 in 2016, OPA published multiple articles in peer-reviewed journals to inform health care providers in public and private settings (e.g., commercial health plans, Medicaid, community health centers, free-standing reproductive health clinics) about the new measure. These publications outline our conceptual framework for developing #2903 alongside its two complementary measures (NQF #2902 and #2904) and describe appropriate measure implementation and use. Furthermore, OPA highlighted systematic reviews which indicate that effective contraceptive method use increases the interbirth interval and reduces adolescent and unintended pregnancies. This association between use of most and moderately effective methods and positive reproductive health outcomes demonstrates the importance of contraceptive care measures to health care quality [17-19].

While NQF #2903 and the contraceptive care measures reflect that some contraceptive methods are more effective than others at preventing pregnancy, these measures and their guidelines for use are designed to encourage providers to offer those clients seeking contraception the full range of methods. The goal of providing contraception should never be to recommend any one method or class of methods over women's individual choices. Women who want to delay or prevent pregnancy should have access to a broad range of contraceptive methods, preferably on a same-day, on-site basis. Furthermore, it is important that these contraceptive services are provided in a client-centered manner that treats each person as a unique individual with respect, empathy, and understanding, providing accurate, easy-to-understand information based on the client's self-identified needs, goals, preferences, and values [11].

Patients receiving client-centered care may feel motivated to continue seeking reproductive health care for contraception and if they become pregnant, prenatal care and birth [13]. Thus, efforts to provide client-centered contraceptive services aligned with the American Academy of Pediatrics (AAP), the American College of Obstetricians and Gynecologists (ACOG), the Centers for Disease Control and Prevention (CDC), and Office of Population Affairs (OPA) recommendations [7-12] may be strengthened by quality improvement processes based on standardized metrics of contraceptive care provision.

References

[1] Conde-Agudelo, A., Rosas-Bermúdez, A., & Kafury-Goeta, A. C. (2006). Birth spacing and risk of adverse perinatal outcomes: a meta-analysis. JAMA, 295(15), 1809–1823. https://doi.org/10.1001/jama.295.15.1809

[2] Conde-Agudelo A, Rosas-Bermúdez A, Kafury-Goeta AC. Effects of birth spacing on maternal health: a systematic review. Am J Obstet Gynecol. 2007 Apr;196(4):297-308. doi: 10.1016/j.ajog.2006.05.055. PMID: 17403398.

[3] Mansour, D., Inki, P., & Gemzell-Danielsson, K. (2010). Efficacy of contraceptive methods: A review of the literature. The European journal of contraception & reproductive health care: the official journal of the European Society of Contraception, 15(1), 4–16. https://doi.org/10.3109/13625180903427675

[4] Trussell, J., Aiken, A.R.A., Micks, E., Guthrie, K.A. (2018). Efficacy, safety, and personal considerations. In R.A. Hatcher, A.L. Nelson, J. Trussell, C. Cwiak, P. Cason, M.S. Policar, A. Edelman, A.R.A. Aiken, J. Marrazzo, D. Kowal (Eds.). Contraceptive technology (21st ed., pp. 95–128). Ayer Company Publishers, Inc.

[5] Winner, B., Peipert, J. F., Zhao, Q., Buckel, C., Madden, T., Allsworth, J. E., & Secura, G. M. (2012). Effectiveness of long-acting reversible contraception. The New England journal of medicine, 366(21), 1998–2007. https://doi.org/10.1056/NEJMoa1110855

[6] Harper, C. C., Rocca, C. H., Thompson, K. M., Morfesis, J., Goodman, S., Darney, P. D., Westhoff, C. L., & Speidel, J. J. (2015). Reductions in pregnancy rates in the USA with long-acting reversible contraception: a cluster randomised trial. Lancet (London, England), 386(9993), 562–568. https://doi.org/10.1016/S0140-6736(14)62460-0

[7] Committee on Gynecologic Practice Long-Acting Reversible Contraception Working Group (2015). Committee Opinion No. 642: Increasing Access to Contraceptive Implants and Intrauterine Devices to Reduce Unintended Pregnancy. Obstetrics and gynecology, 126(4), e44–e48. https://doi.org/10.1097/AOG.0000000000001106

[8] ACOG Committee Opinion No. 735: Adolescents and Long-Acting Reversible Contraception: Implants and Intrauterine Devices. (2018). Obstetrics and gynecology, 131(5), e130–e139. https://doi.org/10.1097/AOG.0000000000002632

[9] Menon, S., & COMMITTEE ON ADOLESCENCE (2020). Long-Acting Reversible Contraception: Specific Issues for Adolescents. Pediatrics, 146(2), e2020007252. https://doi.org/10.1542/peds.2020-007252

[10] Committee on Adolescence (2014). Contraception for adolescents. Pediatrics, 134(4), e1244–e1256. https://doi.org/10.1542/peds.2014-2299

[11] Gavin, L., Moskosky, S., Carter, M., Curtis, K., Glass, E., Godfrey, E., Marcell, A., Mautone-Smith, N., Pazol, K., Tepper, N., Zapata, L., & Centers for Disease Control and Prevention (CDC) (2014). Providing quality family planning services: Recommendations of CDC and the U.S. Office of Population Affairs.

MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 63(RR-04), 1–54.

[12] Curtis, K. M., Tepper, N. K., Jatlaoui, T. C., Berry-Bibee, E., Horton, L. G., Zapata, L. B., Simmons, K. B., Pagano, H. P., Jamieson, D. J., & Whiteman, M. K. (2016). U.S. Medical Eligibility Criteria for Contraceptive Use, 2016. MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 65(3), 1–103. https://doi.org/10.15585/mmwr.rr6503a1

[13] Gomez, A. M., & Wapman, M. (2017). Under (implicit) pressure: young Black and Latina women´s perceptions of contraceptive care. Contraception, 96(4), 221–226. https://doi.org/10.1016/j.contraception.2017.07.007

[14] Finer, L.B. and Zolna, M.R. (2016). Declines in unintended pregnancy in the United States, 2008-2011. New Engl J Med, 374(9), 843-52

[15] Thoma, M. E., Copen, C. E., & Kirmeyer, S. E. (2016). Short Interpregnancy Intervals in 2014: Differences by Maternal Demographic Characteristics. NCHS data brief, (240), 1–8.

[16] Sundaram, A., Vaughan, B., Kost, K., Bankole, A., Finer, L., Singh, S., & Trussell, J. (2017). Contraceptive Failure in the United States: Estimates from the 2006-2010 National Survey of Family Growth. Perspectives on sexual and reproductive health, 49(1), 7–16. https://doi.org/10.1363/psrh.12017

[17] Gavin, L., Frederiksen, B., Robbins, C., Pazol, K., & Moskosky, S. (2017). New clinical performance measures for contraceptive care: their importance to healthcare quality. Contraception, 96(3), 149–157. https://doi.org/10.1016/j.contraception.2017.05.013

[18] Gavin, L. E., Ahrens, K. A., Dehlendorf, C., Frederiksen, B. N., Decker, E., & Moskosky, S. (2017). Future directions in performance measures for contraceptive care: a proposed framework. Contraception, 96(3), 138–144. https://doi.org/10.1016/j.contraception.2017.06.001

[19] Moniz, M. H., Gavin, L. E., & Dalton, V. K. (2017). Performance Measures for Contraceptive Care: A New Tool to Enhance Access to Contraception. Obstetrics and gynecology, 130(5), 1121–1125. https://doi.org/10.1097/AOG.0000000000002314

S.4. Numerator Statement: Women ages 15-44 at risk of unintended pregnancy who are provided a most (sterilization, intrauterine device, implant) or moderately (injectable, pill, patch, ring) effective method of contraception.

S.6. Denominator Statement: Women ages 15-44 who are at risk of unintended pregnancy.

S.8. Denominator Exclusions: The following categories of women are excluded from the denominator: (1) those who are infecund for non-contraceptive reasons; (2) those who had a live birth in the last 2 months of the measurement year; or (3) those who were still pregnant or their pregnancy outcome was unknown at the end of the measurement year.

De.1. Measure Type: Outcome: Intermediate Clinical Outcome

S.17. Data Source: Claims

S.20. Level of Analysis: Clinician : Group/Practice, Facility, Health Plan, Population : Regional and State

IF Endorsement Maintenance – Original Endorsement Date: Oct 25, 2016 **Most Recent Endorsement Date:** Oct 25, 2016

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? Although not a requirement, two other measures have been submitted for maintenance endorsement in separate applications that are complementary to this measure and – if reported together – would provide a broad perspective on the quality of contraceptive services. The two other measures are focused on:

- Postpartum women – this is a very important sub-population of all women at risk of unintended pregnancy. It has been proposed as a separate measure because of the unique need of this population for birth spacing, and the need to raise awareness so that opportunities are not missed to provide contraceptive services during pregnancy, at delivery and in the postpartum period.
- Long-acting reversible contraceptive methods (LARC) – the LARC methods of intrauterine devices (IUD) and implants are a very important sub-set of all contraceptive methods that have extremely low failure rates. The primary goal of this measure is to monitor whether women have access to LARC methods as determined by whether any health facilities or other reporting units report very low levels of LARC use (e.g., less than 1-2 percent) or at a level that is substantially below the median when compared to other reporting units.

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

MostMod_2903_NQF_Evidence_attachment_2021-04-27.docx

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

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

Yes

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** #2903

Measure Title: Contraceptive Care – Most & Moderately Effective Methods

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

Date of Submission: **4/19/2021**

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

Outcome

☐ Outcome:

☐ Patient-reported outcome (PRO):

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

☒ Intermediate clinical outcome (*e.g., lab value*): **Contraceptive provision**

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

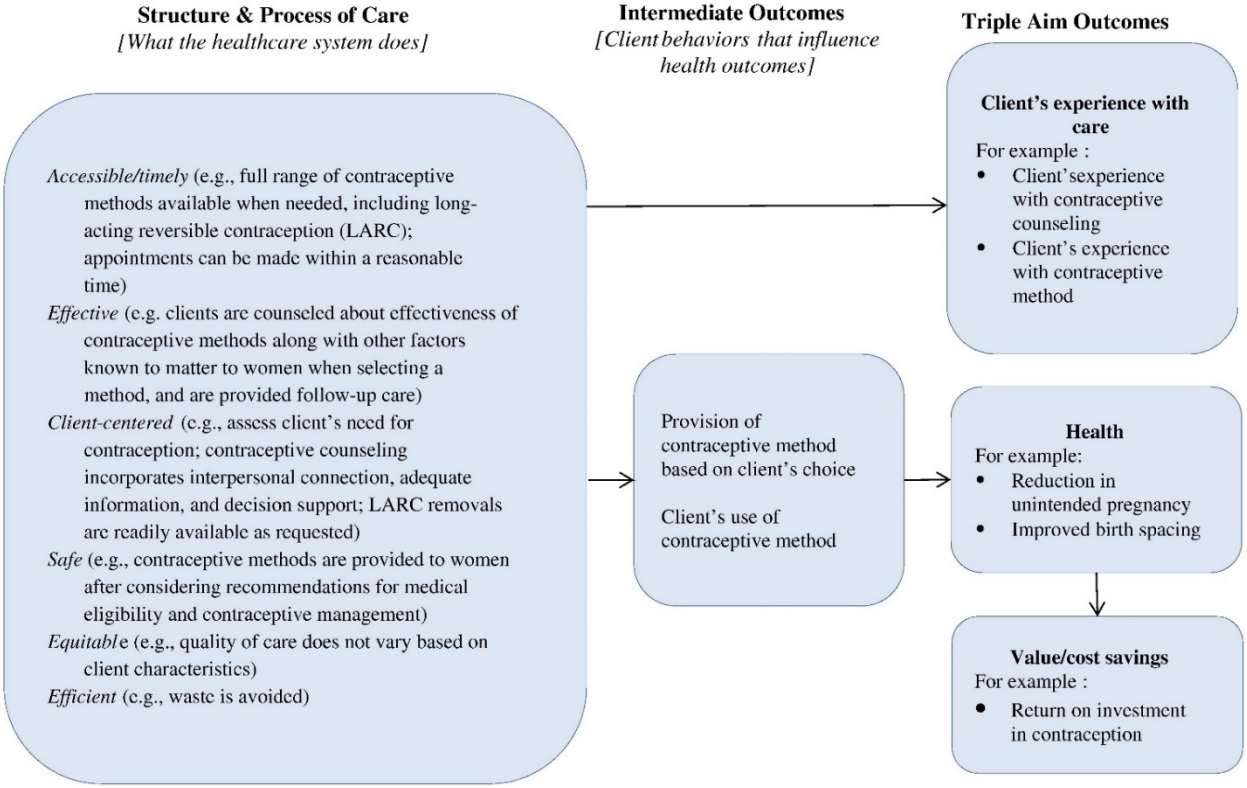
1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process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

2021 Submission

The diagram in Figure 1 below describes the relationship between the structures and processes of quality contraceptive care, including patient- (or client-) centered care, and improved outcomes, including the intermediate clinical outcome of relevance for this application: contraceptive provision. This diagram was developed in 2017 by the U.S. Department of Health and Human Services (HHS), Office of Population Affairs (OPA), in collaboration with the University of California San Francisco (UCSF) Person-Centered Reproductive Health Program, measure steward for the NQF-endorsed Person-Centered Contraceptive Counseling (PCCC) measure (NQF #3543). The diagram was created in the context of describing OPA's work to develop claims-based measures of contraceptive provision (NQF #2902, #2903, and #2904, endorsed in 2016), and the need for the development of the PCCC (which is a Patient-Reported Outcome Performance Measure or PRO-PM) to help provide a more robust picture of contraceptive care quality (Gavin 2017). NQF endorsed the PCCC in November 2020.

OPA's conceptual framework for contraceptive care incorporates essential components of the Institute of Medicine's six dimensions of quality care, Donabedian's quality of care model *structure* and *process* categories, and the Institute for Healthcare Improvement's concept of the "Triple Aim". Several evidence-based clinical family planning recommendations of CDC and OPA serve as examples of health systems' structure and process components in contraceptive care (Gavin and CDC, 2014). These components affect two intermediate clinical outcomes: provision of contraceptive methods based on client's choice, and client's use of contraception. The intermediate outcomes signify a client's decision at the end of a clinical encounter that will influence their probability of having an unintended pregnancy. The structure and process also directly affect the client's experience with care. Health outcomes are influenced through the intermediate outcomes of client behavior; and cost-savings result in reductions in unintended pregnancy and improvements in birth spacing. (Gavin 2017).

Figure 1: Office of Population Affairs' conceptual framework for clinical performance measures for contraceptive care.



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.)

2021 Submission

Not applicable; measure is not derived from patient report.

****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.

2021 Submission

Not applicable; measure is not derived from patient report.

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

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	Clinical Practice Guideline recommendation • Providing Quality Family Planning Services: Recommendations of CDC and the U.S. Office of Population Affairs. • Gavin L, Moskosky S, Carter M, Curtis K, Glass E, Godfrey E, Marcell A, Mautone-Smith N, Pazol K, Tepper N, Zapata L; Centers for Disease Control and Prevention (CDC) • 2014 Apr 25 • Gavin L, Moskosky S, Carter M, Curtis K, Glass E, Godfrey E, Marcell A, Mautone-Smith N, Pazol K, Tepper N, Zapata L; Centers for Disease Control and Prevention (CDC). Providing quality family planning services: Recommendations of CDC and the U.S. Office of Population Affairs. MMWR Recomm Rep. 2014 Apr 25;63(RR-04):1-54. PMID: 24759690. • https://www.cdc.gov/mmwr/pdf/rr/rr6304.pdf

Systematic Review	Evidence
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.	"Providers are encouraged to present information on potential reversible methods of contraception by using a tiered approach (i.e., presenting information on the most effective methods first, before presenting information on less effective methods). This information should include an explanation that long-acting reversible contraceptive methods are safe and effective for most women, including those who have never given birth and adolescents. Information should be tailored and presented to ensure a client-centered approach. It is not appropriate to omit presenting information on a method solely because the method is not available at the service site. If not all methods are available at the service site, it is important to have strong referral links in place to other providers to maximize opportunities for clients to obtain their preferred method that is medically appropriate." Source: CDC/OPA (2014). Providing Quality Family Planning Services (QFP), page 8 and Appendix B Generally, the QFP recommendations outline how to provide family planning services by: • defining a core set of family planning services for women and men, • describing how to provide contraceptive and other clinical services, serve adolescents, and perform quality improvements, and • encouraging the use of the family planning visit to provide selected preventive health services for women, in accordance with the recommendations for women issued by the Institute of Medicine (IOM) and adopted by HHS • support offering a full range of Food and Drug Administration (FDA)-approved contraceptive methods as well as counseling that highlights the effectiveness of contraceptive methods overall

Systematic Review	Evidence
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Of 132 studies, 41 are graded level I and the rest are graded II-1 to II-3 using the USPSTF system. The authors described their method to assess the internal and external validity of included studies below: "The quality, or internal validity, of each individual study was assessed to consider the risk that the findings may be confounded by a systematic bias. We used the schema developed by the USPSTF for describing a study's level of risk for bias. A rating of risk for bias was determined through the presence or absence of several characteristics that are known to protect a study from the confounding influence of bias. We developed criteria by which the risk for bias of individual studies could be evaluated, based on recommendations from several sources, including the USPSTF; the Grading of Recommendations Assessment, Development and Evaluation (GRADE); and Community Guide for Preventive Services." Further details can be found in Appendix A of QFP (p. 30-32). In addition, CDC published its methodology for the systematic reviews describing the evidence and their grading in the following paper: Tregear, S. J., Gavin, L. E., & Williams, J. R. (2015). Systematic Review Evidence Methodology: Providing Quality Family Planning Services. *American journal of preventive medicine*, *49*(2 Suppl 1), S23–S30. https://doi.org/10.1016/j.amepre.2015.03.033 The SRs contained in the body of evidence are provided in a supplement of *American Journal of Preventive Medicine*: American Journal of Preventive Medicine, Volume 49, Issue 2, Supplement 1, Pages S1-S123 (August 2015). Available online at: https://www.ajpmonline.org/issue/S0749-3797(15)X0002-X

Systematic Review	Evidence
Provide all other grades and definitions from the recommendation grading system	A: There is good evidence to support the recommendation that the condition be considered specifically in a preconception care evaluation. B: There is fair evidence to support the recommendation that the condition be considered specifically in a preconception care evaluation. C: There is insufficient evidence to recommend for or against the inclusion of the condition in a preconception care evaluation, but recommendation to include or exclude may be made on other grounds. D: There is fair evidence to support the recommendation that the condition be excluded in a preconception care evaluation. E: There is good evidence to support the recommendation that the condition be excluded in a preconception care evaluation.

Systematic Review	Evidence
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Quantity: Summaries of the evidence used to prepare these recommendations are published in 9 separate systematic reviews in the original 2014 version of the guideline, and a total of 132 studies are included in the 9 systematic reviews. Quality: CDC and the Office of Population Affairs developed QFP recommendations by conducting an extensive review of published evidence, seeking expert opinion, and synthesizing existing clinical recommendations from CDC, agencies such as the U.S. Preventive Services Task Force (USPSTF), and professional medical associations such as the American College of Obstetricians and Gynecologists and the American Academy of Pediatrics. Types of studies included in the systematic reviews included: randomized controlled trials (41 studies), non-randomized controlled trials, national survey data, prospective cohorts, case-control cohort, cross-sectional studies, pre-post studies, ecological evaluation, and descriptive studies. Summary can be found in Appendix B of the 2014 QFP (p. 35-44). In addition, CDC published its methodology for the systematic reviews describing the evidence and their grading in the following paper: Tregear, S. J., Gavin, L. E., & Williams, J. R. (2015). Systematic Review Evidence Methodology: Providing Quality Family Planning Services. *American journal of preventive medicine*, *49*(2 Suppl 1), S23–S30. https://doi.org/10.1016/j.amepre.2015.03.033
Estimates of benefit and consistency across studies	QFP provides guidelines to provide family planning services, including the provision of contraception, to help women plan and space births, prevent unintended pregnancies, and reduce the number of abortions.

Systematic Review	Evidence
What harms were identified?	The harms were not listed in these guidelines. However, CDC clinical recommendations on contraceptive safety address this question. CDC's "US Medical Eligibility Criteria for Contraceptive Use" (USMEC) describe what contraceptive methods are safe for women with a range of characteristics (e.g., age, postpartum) and medical conditions (e.g., infectious, or chronic diseases). The citation for the USMEC recommendations is: Curtis, K. M., Tepper, N. K., Jatlaoui, T. C., Berry-Bibee, E., Horton, L. G., Zapata, L. B., Simmons, K. B., Pagano, H. P., Jamieson, D. J., & Whiteman, M. K. (2016). U.S. Medical Eligibility Criteria for Contraceptive Use, 2016. *MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports*, *65*(3), 1–103. https://doi.org/10.15585/mmwr.rr6503a1

Systematic Review	Evidence
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Gavin L, Pazol K, Ahrens K. Update: Providing Quality Family Planning Services — Recommendations from CDC and the U.S. Office of Population Affairs, 2017. MMWR Morb Mortal Wkly Rep 2017;66:1383–1385. DOI: http://dx.doi.org/10.15585/mmwr.mm6650a4External Gavin L, Pazol K. Update: Providing Quality Family Planning Services — Recommendations from CDC and the U.S. Office of Population Affairs, 2015. MMWR Morb Mortal Wkly Rep 2016;65:231–234. DOI: http://dx.doi.org/10.15585/mmwr.mm6509a3 These two reviews revised and updated the 2014 version based on new scientific findings. They did not make a substantial shift in how family planning care should be provided. The American Academy of Family Physicians issued a clinical practice guideline recommendation in support of and advocating use for use of QFP, which did not change conclusions of original SR. This AAFP guideline is available online at: https://www.aafp.org/afp/2015/0501/p625.html In 2018, OPA updated and expanded several systematic reviews on the following topics addressed in the 2014 QFP: counseling and education (three updated and one new systematic review), serving adolescents (one updated and one new systematic review), and community education and engagement (one paper updating two previous systematic reviews). These articles did not change conclusions of the original SR and were published in a theme issue of *American Journal of Preventative Medicine*: American Journal of Preventative Medicine, Volume 55, Issue 5, Pages 677-690, (November 01, 2018). Available online at: https://www.ajpmonline.org/issue/S0749-3797(17)X0016-0#

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	Clinical Practice Guideline recommendation • Long-Acting Reversible Contraception: Implants and Intrauterine Devices • American College of Obstetricians and Gynecologists (ACOG) • 2017 November, reaffirmed in 2019 • Long-acting reversible contraception: implants and intrauterine devices. Practice Bulletin No. 186. American College of Obstetricians and Gynecologists. Obstet Gynecol 2017; 130:e251-69 • https://doi.org/10.1097/AOG.0000000000002400
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.	In summary, intrauterine devices (IUDs) and contraceptive implants, also called long-acting reversible contraceptives (LARC), are the most effective reversible contraceptive methods that can be provided to a broad range of patients wishing to prevent pregnancy, including postpartum women. Below is the Summary of Recommendations, by grade: **"The following recommendations are based on good and consistent scientific evidence (Level A):** Insertion of an IUD immediately after first-trimester uterine aspiration should be offered routinely as a safe and effective contraceptive option. Insertion of the contraceptive implant on the same day as first-trimester or second-trimester induced or spontaneous abortion should be offered routinely as a safe and effective contraceptive option. Routine antibiotic prophylaxis is not recommended before IUD insertion. **The following recommendations are based on limited or inconsistent scientific evidence (Level B):**

Systematic Review	Evidence
	Intrauterine devices and the contraceptive implant should be offered routinely as safe and effective contraceptive options for nulliparous women and adolescents. Insertion of an IUD or an implant may occur at any time during the menstrual cycle as long as pregnancy may be reasonably excluded. Insertion of an IUD immediately after confirmed completion of first-trimester medication-induced abortion should be offered routinely as a safe and effective contraceptive option. Immediate postpartum IUD insertion (i.e., within 10 minutes after placental delivery in vaginal and cesarean births) should be offered routinely as a safe and effective option for postpartum contraception. Immediate postpartum initiation of the contraceptive implant (i.e., insertion before hospital discharge after a hospital stay for birth) should be offered routinely as a safe and effective option for post-partum contraception, regardless of breastfeeding status. Women who have not undergone routine screening for STIs or who are identified to be at increased risk of STIs based on patient history should receive CDC-recommended STI screening at the time of a single visit for IUD insertion. Intrauterine device insertion should not be delayed while awaiting test results. Treatment for a positive test result may occur without removal of the IUD. Intrauterine devices may be offered to women with a history of ectopic pregnancies. **The following recommendations are based primarily on consensus and expert opinion (Level C):** Long-acting reversible contraceptives have few contraindications and should be offered routinely as safe and effective contraceptive options for most women.

Systematic Review	Evidence
	The copper IUD should be offered routinely to women who request emergency contraception and are eligible for IUD placement. To improve LARC method satisfaction and continuation, patient counseling should include information on expected bleeding changes and reassurance that these changes are not harmful. Endometrial biopsy, colposcopy, cervical ablation or excision, and endocervical sampling may all be performed with an IUD in place. Actinomyces on cytology is considered an incidental finding. In the absence of symptoms, no antimicrobial treatment is needed, and the IUD may be left in place. Intrauterine device removal is recommended in pregnant women when the strings are visible or can be removed safely from the cervical canal. There is no compelling evidence for the removal of an IUD or implant before its expiration date in menopausal women." (p. e262)
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Grades assigned to the evidence followed the method outlined by the U.S. Preventive Services Task Force (USPSTF). The evidence associated with the recommendations included 132 graded studies. The evidence was graded as follows: • 30 studies were graded I (Evidence obtained from at least one properly designed randomized controlled trial.) • 13 studies were graded II-2 (Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.) • 43 studies were graded II-3 (Evidence obtained from multiple time series with or without the intervention.

Systematic Review	Evidence
	Dramatic results in uncontrolled experiments also could be regarded as this type of evidence.) • 46 studies were graded III (Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.)
Provide all other grades and definitions from the evidence grading system	Studies were reviewed and evaluated for quality according to the method outlined by the USPSTF. All grades in the USPSTF grading system for research studies were assigned to the analyses comprising the evidence, except for the following grade: II-1 Evidence obtained from well-designed controlled trials without randomization.
Grade assigned to the **recommendation** with definition of the grade	The USPSTF grading system for recommendations was used to assign grades. A total of 17 recommendations were provided in this clinical practice guideline recommendation with evidence review. 3 recommendations were assigned the grade Level A (Recommendations are based on good and consistent scientific evidence) 7 recommendations were assigned the grade Level B (Recommendations are based on limited or inconsistent scientific evidence) 7 recommendations were assigned the grade Level C (Recommendations are based primarily on consensus and expert opinion)
Provide all other grades and definitions from the recommendation grading system	Not applicable. All grades are included in the box above.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• This SR counted 151 studies in its body of evidence. About one-third of these studies were randomized controlled trials, case-control studies, or cohort studies. • 30 randomized controlled trials • 13 cohort or case-control analytic studies

Systematic Review	Evidence
	• 43 studies from multiple time series with or without intervention, uncontrolled experiments • 46 descriptive studies, expert committee reports, expert opinions based on clinical experience • 15 systematic reviews • 2 cost-benefit studies • 2 meta-analyses
Estimates of benefit and consistency across studies	ACOG's review indicated that LARC methods are safe, highly effective forms of contraception for most women, including subpopulations of women like adolescent females, nulliparous women, and women post-abortion. An increase in LARC use may have partially contributed to the decline in the rate of unintended pregnancies in the United States from 51% to 45% between 2008-2011. Citing Trussell's 2011 review of contraceptive failure rates, this review reported that the LARC methods have a typical failure rate less than 1%. ACOG found good and consistent evidence that LARC methods can be inserted immediately after induced or spontaneous abortion, providing safe and effective contraception to prevent pregnancy. One RCT reported that among women receiving immediate insertion post-abortion, six-month IUD use rates were higher than in the delayed-insertion group (92.3% vs. 76.6%; $p<0.001$) with no difference for expulsion risk between groups. No pregnancies occurred in the immediate insertion group. For post-abortion implant insertion, one RCT found that risk of medication abortion failure was low and similar between the immediate placement (i.e., same day as mifepristone administration) and after medication-induced abortion (3.9% vs. 3.8%). Another prospective cohort study indicated that continuation rates were similar among women with immediate and delayed post-abortion implant placement (82% for immediate and interval placement). ACOG determined that adequate scientific evidence exists that IUDs and implants should be offered to adolescents and nulliparous women routinely as safe and effective contraceptive options with a prevent pregnancy. One retrospective cohort study in IUD users reported that serious complications (i.e., ectopic pregnancy, pelvic inflammatory disease) were rare regardless of age or IUD type. Although

Systematic Review	Evidence
	adolescent women (ages 15-19) were more likely to have a claim for menstrual bleeding changes or normal pregnancy than women ages 25-44, early discontinuation rates were similar in both groups (13% vs. 11%). The Contraceptive CHOICE project, a prospective cohort study, reported high uptake for LARC methods by adolescents when these methods were offered for free. Young women ages 14-17 years selecting a LARC method were more likely to use the implant (63%) while those ages 18-20 chose an IUD (71%). Another study reported that continuation rates for postpartum adolescents using the implant were higher than those using contraceptive injection or combined oral contraceptive pills; this difference was statistically significant ($p<0.001$).
What harms were identified?	ACOG described the following harms for LARC methods in this review. **Harms identified with IUDs** In two studies (prospective and retrospective cohorts), users of copper and levonorgestrel-releasing (LNG) IUDs had similar mean weight gain. Commonly reported adverse effects with the copper IUD are heavy menstrual bleeding and pain. Some LNG IUD users reported the following hormone-related side effects: headaches, nausea, breast tenderness, mood changes, and ovarian cyst formation. Expulsion, method failure, and perforation are complications with IUDs that appear to rarely occur. A large, prospective, noninterventional 2015 study surveilling over 61,000 women for seven years reported 1.4 per 1000 LNG IUD insertions and 1.1 per 1000 copper IUD insertions. **Harms identified with Implants** Changes in menstrual bleeding patterns is a common side effect of implant use. One randomized, multicenter comparative study noted that the median number of bleeding/spotting days decreased from the first 90 days to the last year of the study period (Implanon: 33.5 to 19-21.5 days; Norplant: 34.5 to 18.0-23.0). The mean overall incidence decreased during the study (Implanon: 66.0% to 27.3%; Norplant: 69.0% to 21.7%).

Systematic Review	Evidence
	Additional adverse events reported by implant users are gastrointestinal difficulties, headaches, breast pain, vaginitis, acne, and weight gain. Another RCT reported that 1-year cumulative discontinuation rates due to menstrual bleeding disturbances was 2.1% for implants, but weight gain was cited as the main reason for 7% of users to discontinue the implant. About 83% of participants in this study continued using the implant for the project duration. One integrated analysis of international clinical trials reported that complications were rare during implant insertion and removal (1.0% for insertion, 1.7% for removal). Women experiencing insertion complications reported pain, slight bleeding, hematoma formation, deep or incorrect insertion and unrecognized insertion. Complications with removal include breakage of the implant and failure to palpate or locate the implant due to deep insertion.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	This clinical guidance was reaffirmed in 2019 without changing the SR's conclusions.

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	Clinical Practice Guideline recommendation • US Medical Eligibility Criteria (US MEC) for Contraceptive Use, 2016. • CDC • 2016 • Curtis KM, Tepper NK, Jatlaoui TC, et al. U.S. Medical Eligibility Criteria for Contraceptive Use, 2016. MMWR Recomm Rep 2016;65(No. RR-3):1–104. DOI: http://dx.doi.org/10.15585/mmwr.rr6503a1 • http://dx.doi.org/10.15585/mmwr.rr6503a1
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline,	The United States Medical Eligibility Criteria for Contraceptive Use, 2016 (US MEC) includes recommendations for using specific contraceptive methods by women and men who have certain characteristics or medical conditions. The recommendations in this report are intended to assist health care providers when they counsel women, men, and couples

Systematic Review	Evidence
summarize the conclusions from the SR.	about contraceptive method choice. This report serves as an update to the 2010 US MEC, which was adapted from the fourth edition of World Health Organization's *Medical Eligibility Criteria for Contraceptive Use* (WHO MEC). The SR concludes that most women, including those with certain characteristics (e.g., adolescents, postpartum) and medical conditions (e.g., infectious or chronic diseases), can use most contraceptive methods safely to prevent pregnancy. Recommendations related to IUDs and implants are reported in this review. Women who have health conditions associated with increased risk for adverse health events as a result of pregnancy should consider long-acting, highly effective contraception. The 2016 US MEC recommendations are summarized in the following tables: https://www.cdc.gov/reproductivehealth/contraception/pdf/summary-chart-us-medical-eligibility-criteria_508tagged.pdf Safety of contraceptive methods is a component of the structure and process of the health care system, which affects the provision of contraceptive methods, including LARC. The recommendations aim to eliminate unneeded medical barriers to accessing and using contraception, which in turn may decrease the number of unintended pregnancies.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were followed for reporting systematic reviews. The level of evidence from the systematic reviews for each evidence summary are provided based on the U.S. Preventive Services Task Force (USPSTF) system, which includes ratings for study design (I: randomized controlled trials; II-1: controlled trials without randomization; II-2: observational studies; and II-3: multiple time series or descriptive studies), ratings for internal validity (good, fair, or poor), and categorization of the evidence as direct or indirect for the specific review topic.

Systematic Review	Evidence
	Evidence in this guideline ranges from I to II-3, good to poor, direct to indirect, depending on the condition and contraceptive method evaluated. For the 2016 US MEC update, CDC published 13 systematic reviews describing the evidence and their grading related to new recommendations not previously included in the 2010 US MEC. These reviews are provided in a supplement of *Contraception*: Contraception, Volume 96, Issue 6, Pages 579-760 (December 2016). Available online at: https://www.sciencedirect.com/journal/contraception/vol/94/issue/6
Provide all other grades and definitions from the evidence grading system	The following grade from the USPSTF system was not assigned to evidence in this SR: III: Opinions of respected authorities, based on clinical experience, descriptive studies and case reports, or reports of expert committees.
Grade assigned to the **recommendation** with definition of the grade	Appendices B – J provide summaries of classifications for whether women with certain medical conditions or characteristics can use contraceptive methods. The following methods are included: IUDs, progestin-only contraceptives (including etonogestrel implants), combined hormonal contraceptives, barrier contraceptive methods, fertility awareness-based methods, lactational amenorrhea method, coitus interruptus, female and male sterilization, and emergency contraceptive pills. The four categories utilized to classify the use of contraceptive methods, including LARC methods, for women with certain medical conditions or characteristics are as follows: U.S. MEC 1 = A condition for which there is no restriction for the use of the contraceptive method. U.S. MEC 2 = A condition for which the advantages of using the method generally outweigh the theoretical or proven risks.

Systematic Review	Evidence
	U.S. MEC 3 = A condition for which the theoretical or proven risks usually outweigh the advantages of using the method. U.S. MEC 4 = A condition that represents an unacceptable health risk if the contraceptive method is used. Depending on the contraceptive methods and conditions, the grading ranges from U.S. MEC 1 – 4. The recommendations were developed as part of a multi-stage process. First, CDC reviewed the existing recommendations in the US MEC 2010 for new evidence identified through the WHO/CDC CIRE system that might result in a changed recommendation. CDC then held an initial expert panel meeting to obtain input and draft a list of topics to consider for the update, including new recommendations. Next CDC staff and other invited authors conducted independent systematic reviews for topics under consideration. These reviews were conducted to identify direct evidence about the safety of contraceptive methods use by women with selected conditions. At a second expert meeting, participants were asked to provide their input using the scientific evidence presented from the systematic reviews to develop potential recommendations. Feedback also was received from three external reviewers, composed of health care providers and researchers who had not participated in the update meetings. These reviewers were asked to provide comments on the accuracy, feasibility, and clarity of the recommendations. During the second expert meeting, areas of research that need additional investigation also were considered. Afterwards CDC chose and documented the recommendations in this report, taking into account the perspectives offered by expert meeting participants.
Provide all other grades and definitions from the recommendation grading system	Not applicable. All grades are included in the box above.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Quantity – Summaries of the evidence used to prepare the new recommendations issued in 2016 are published in 13 separate systematic reviews. These summaries included a total of 108 articles.

Systematic Review	Evidence
	Quality – The 108 articles described the following types of studies: randomized controlled trials, non-randomized controlled trials, prospective and retrospective cohort studies, case-control studies, pharmacokinetic studies, cross-sectional studies, and pooled analyses.
Estimates of benefit and consistency across studies	A broad range of contraceptive methods are safe for women with a range of characteristics (e.g., age, postpartum) and medical conditions (e.g., infectious, or chronic diseases). The goal of these recommendations is to remove unnecessary medical barriers to accessing and using contraception, thereby decreasing the number of unintended pregnancies.
What harms were identified?	Some harms were noted in the clarification column in each appendix. However, the individual studies comprising the body of evidence may have specifically identified potential adverse events related to contraceptive method use among women with certain health conditions and characteristics. CDC published 13 systematic reviews describing the evidence and their grading for this update in a supplement of *Contraception*: Contraception, Volume 96, Issue 6, Pages 579-760 (December 2016). Available online at: https://www.sciencedirect.com/journal/contraception/vol/94/issue/6
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	• American Academy of Family Physicians issued the following practice guidelines which support and advocate for the use of US MEC: ○ https://www.aafp.org/afp/2017/0115/afp20170115p125.pdf ○ https://www.aafp.org/afp/2016/1201/afp20161201p942.pdf ○ https://www.aafp.org/afp/2015/0501/afp20150501p625.pdf ○ https://www.aafp.org/afp/2012/0215/afp20120215p403.pdf These new guidelines did not change the SR's conclusions.

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	Clinical Practice Guideline recommendation • U.S. Selected Practice Recommendations for Contraceptive Use • Curtis KM, Jatlaoui TC, Tepper NK, et al. • 2016 • Curtis KM, Jatlaoui TC, Tepper NK, et al. U.S. Selected Practice Recommendations for Contraceptive Use, 2016. MMWR Recomm Rep 2016;65(No. RR-4):1–66. DOI: http://dx.doi.org/10.15585/mmwr.rr6504a1 • http://dx.doi.org/10.15585/mmwr.rr6504a1
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.	Most women can start most contraceptive methods at any time, and few examinations or tests, if any, are needed before starting a contraceptive method. Routine follow-up for most women includes assessment of her satisfaction with the contraceptive method, concerns about method use, and changes in health status or medications that could affect medical eligibility for continued use of the method.

Systematic Review	Evidence
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Evidence in this guideline ranges from I to II-3, good to poor, direct to indirect, depending on the contraceptive methods provision and related services. The level of evidence from the systematic reviews for each evidence summary are provided based on the U.S. Preventive Services Task Force (USPSTF) system, which includes ratings for study design (I: randomized controlled trials; II-1: controlled trials without randomization; II-2: observational studies; and II-3: multiple time series or descriptive studies), ratings for internal validity (good, fair, or poor), and categorization of the evidence as direct or indirect for the specific review question. Evidence considered for use in SPR was considered in a multi-tiered approach. For the 2013 version of SPR, CDC initiated a process to adapt WHO SPR for the United States. This adaptation process included four steps: 1) determining the scope of and process for the adaptation, including an October 2010 meeting in which individual feedback was solicited from a small group of partners and experts; 2) preparing the systematic reviews of the evidence during October 2010–September 2011 to be used for the adaptation, including peer review; 3) convening a larger meeting of experts in October 2011 to examine the evidence and receive input on the recommendations; and 4) finalizing recommendations by CDC. Additional evidence was similarly garnered and considered for the 2016 update to SPR.
Provide all other grades and definitions from the evidence grading system	USPSTF I Evidence obtained from at least one properly randomized controlled trial. II–1 Evidence obtained from well-designed controlled trials without randomization. II–2 Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group. II–3 Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of

Systematic Review	Evidence
	penicillin treatment in the 1940s) could also be regarded as this type of evidence. III Opinions of respected authorities, based on clinical experience, descriptive studies and case reports, or reports of expert committees
Grade assigned to the **recommendation** with definition of the grade	Appendix A of SPR provides a summary of classifications for hormonal contraceptive methods and intrauterine devices by condition, pregnancy, and age. (https://www.cdc.gov/mmwr/volumes/65/rr/pdfs/rr6504.pdf), pages -53-61 Depending on the contraceptive methods and conditions, the grading ranges from U.S. MEC 1 – 4. Categories of medical eligibility criteria for contraceptive use: U.S. MEC 1 = A condition for which there is no restriction for the use of the contraceptive method. U.S. MEC 2 = A condition for which the advantages of using the method generally outweigh the theoretical or proven risks. U.S. MEC 3 = A condition for which the theoretical or proven risks usually outweigh the advantages of using the method. U.S. MEC 4 = A condition that represents an unacceptable health risk if the contraceptive method is used.
Provide all other grades and definitions from the recommendation grading system	Not applicable. All grades are included in the box above.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity – 353 Studies • Quality – study types included systematic reviews, meta-analyses, randomized controlled trials, clinical trials, diagnostic accuracy studies, and case series.
Estimates of benefit and consistency across studies	Most women can start most contraceptive methods at any time, and few examinations or tests, if any, are needed before starting a contraceptive method.
What harms were identified?	Because changes in bleeding patterns are one of the major reasons for discontinuation of contraception,

Systematic Review	Evidence
	recommendations are provided for the management of bleeding irregularities with various contraceptive methods. In addition, because women and health care providers can be confused about the procedures for missed pills and dosing errors with the contraceptive patch and ring, the instructions are streamlined for easier use.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Not applicable.

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	Clinical Practice Guideline recommendation • Women's Preventive Services Guidelines • Health Resources and Services Administration (HRSA) and ACOG • 2019 December 17 • Health Resources and Services Administration. (2019, December). *Women's Preventive Services Guidelines*. U.S. Department of Health and Human Services, Health Resources and Services Administration. https://www.hrsa.gov/womens-guidelines/index.html • https://www.hrsa.gov/womens-guidelines/index.html
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.	The Women's Preventive Services Initiative recommends that adolescent and adult women have access to the full range of female-controlled contraceptives to prevent unintended pregnancy and improve birth outcomes. Contraceptive care should include contraceptive counseling, initiation of contraceptive use, and follow-up care (e.g., management, and evaluation as well as changes to and removal or discontinuation of the contraceptive method). The Women's Preventive Services Initiative recommends that the full range of female-controlled U.S. Food and Drug Administration-approved contraceptive methods, effective family planning practices, and sterilization procedures be available as part of contraceptive care. The full range of contraceptive methods for women currently identified by the U.S. Food and Drug Administration include:

Systematic Review	Evidence
	(1) sterilization surgery for women, (2) surgical sterilization via implant for women, (3) implantable rods, (4) copper intrauterine devices, (5) intrauterine devices with progestin (all durations and doses), (6) the shot or injection, (7) oral contraceptives (combined pill), 8) oral contraceptives (progestin only, and), (9) oral contraceptives (extended or continuous use), (10) the contraceptive patch, (11) vaginal contraceptive rings, (12) diaphragms, (13) contraceptive sponges, (14) cervical caps, (15) female condoms, (16) spermicides, and (17) emergency contraception (levonorgestrel), and (18) emergency contraception (ulipristal acetate), and additional methods as identified by the FDA. Additionally, instruction in fertility awareness-based methods, including the lactation amenorrhea method, although less effective, should be provided for women desiring an alternative method.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	While grades of evidence is not presented in the guideline, below is how the recommendations were developed: The WPSI has contracted with physician scientists with extensive experience in systematic review methodology and clinical guideline development from the Pacific Northwest Evidence-based Practice Center (EPC) at Oregon Health & Science University to conduct reviews and updates of the evidence for each topic under consideration. Focused updates of evidence reviewed for the nine topics considered for revision include overviews of recent systematic reviews for the U.S. Preventive Services Task Force (USPSTF) published since the last recommendations were issued by the Institute of Medicine (IOM) Committee in 2011, as well as systematic reviews and key studies published since the most recent systematic reviews for the USPSTF. A research librarian conducted searches in Ovid MEDLINE, Cochrane Central Register of Controlled Trials, and Cochrane Database of Systematic Reviews through July 2016 for all topics.

Systematic Review	Evidence
	A best evidence approach was applied when reviewing abstracts and selecting studies to include for the updates that involves using the most relevant studies with the strongest methodologies. For well-woman visits and contraceptive methods and counseling, there are no USPSTF reviews or recommendations, therefore, other systematic reviews and studies published since the 2011 IOM recommendations for these topics were included. Randomized controlled trials and large (>100) prospective cohort studies were included if they provided relevant information for each topic. Other study designs, such as case-control and modeling studies, were included when evidence was lacking or when they demonstrated new findings. Studies conducted in settings applicable to the United States were targeted. The focus of each review was on gaps identified in the 2011 IOM recommendations and any new evidence that could change or additionally inform the recommendations where evidence was not previously available. Selection criteria specific to each topic were developed to address issues specific to the WPSI. Applicability is defined as the extent to which the effects observed in published studies are likely to reflect the expected results when a specific intervention is applied to the population of interest under "real-world" conditions. It is an indicator of the extent to which research included in a review might be useful for informing clinical decisions in specific situations. Factors important for understanding the applicability of studies were considered including differences in the interventions and comparators, populations, and settings. No new or revised statistical meta-analyses were conducted. Studies were qualitatively synthesized according to interventions, populations, and outcomes measured. Studies and their findings were summarized in a narrative, descriptive format to provide an overview of the new evidence for each topic.

Systematic Review	Evidence
	MSC members interact with the EPC to identify topics and scope. Updates to previous recommendations were evaluated using established methodology. In 2019, HRSA published updated guidelines online.
Provide all other grades and definitions from the evidence grading system	Not applicable.
Grade assigned to the **recommendation** with definition of the grade	While grades of recommendations are not presented in the guideline, below is how the recommendations were developed: In addition to current systematic reviews and randomized controlled trials, other supporting evidence is considered including organization guidelines and policies, epidemiologic data, and other relevant sources. Physician investigators from the EPC attend in-person and teleconference MSC meetings to assist with interpretation of evidence, including addressing queries about individual studies included in the literature search. Investigators work closely with the MSC, and each of the subcommittees, to provide expert perspective on the quality and strength of the supporting evidence. In addition, like the 2011 IOM Panel, the MSC panel considered multiple levels of evidence when developing the recommendations and permitted recommendations to be based on varying levels of evidence, expert consensus, or standard best practices.
Provide all other grades and definitions from the recommendation grading system	Preventive services recommended by the committee followed the criteria of the 2011 IOM Panel: • The condition to be prevented affects a broad population

Systematic Review	Evidence
	• The condition to be prevented has a large potential impact on health and well being • The quality and strength of evidence is supportive.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• 2 systematic reviews • 1 randomized controlled trial • 2 observational studies • 1 clustered randomized trial • 1 book chapter
Estimates of benefit and consistency across studies	The effectiveness of the full range of FDA-approved contraceptive methods for preventing or delaying pregnancy is well established. Effective comprehensive contraceptive care includes counseling, initiation, and follow-up. Contraceptive counseling and access to contraceptive methods is associated with increased contraceptive use and decreased unintended pregnancy rates. Long-acting reversible contraceptive (LARC) methods are the most effective reversible contraceptive option for most women, including nulliparous women and adolescents who are sexually active. Counseling on LARC methods is associated with lower pregnancy rates and lower rates of abortion and repeat abortion. Providing an increased supply of oral contraceptives at initiation is associated with higher continuation rates and lower unintended pregnancy rates.
What harms were identified?	The harms related to contraceptive method use were not listed in these guidelines. However, CDC clinical recommendations on contraceptive safety explicitly address this question. CDC's "US Medical Eligibility Criteria for Contraceptive Use" (USMEC) describe what contraceptive methods are safe for women with a range of characteristics (e.g., age, postpartum) and medical conditions (e.g., infectious, or chronic diseases). The citation for the USMEC recommendations is: Curtis, K. M., Tepper, N. K., Jatlaoui, T. C., Berry-Bibee, E., Horton, L. G., Zapata, L. B., Simmons, K. B., Pagano, H. P., Jamieson, D. J., & Whiteman, M. K. (2016). U.S. Medical Eligibility Criteria for Contraceptive Use, 2016. *MMWR. Recommendations and reports : Morbidity and mortality*

Systematic Review	Evidence
	weekly report. Recommendations and reports, *65*(3), 1–103. https://doi.org/10.15585/mmwr.rr6503a1
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Not applicable.

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	• Contraceptive Technology. 21st Ed • Hatcher RA, Nelson AL, Trussell J, Cwiak C, Cason P, Policar MS, Edelman A, Aiken ARA, Marrazzo J, Kowal D, eds. • 2018 • Hatcher RA, Nelson AL, Trussell J, Cwiak C, Cason P, Policar MS, Edelman A, Aiken ARA, Marrazzo J, Kowal D, eds. Contraceptive technology. 21st ed. New York, NY: Ayer Company Publishers, INC., 2018. • http://www.contraceptivetechnology.org/the-book/
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.	• Use of the top-tier reversible contraceptives – the intrauterine devices (IUDs) and the contraceptive implant – entails the lowest risk of pregnancy. • Correct and consistent use of most contraceptive methods results in a low risk of pregnancy • Most contraceptives pose little risk to most users' health, although personal risk factors should influence personal choice.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Grade not assigned, but *Contraceptive Technology* serves as the primary source of information about contraceptive failure rates and is cited by the World Health Organization, CDC, and leading health professional associations in the US and other countries.
Provide all other grades and definitions from the evidence grading system	Not applicable.

Systematic Review	Evidence
Grade assigned to the **recommendation** with definition of the grade	Grade not assigned, but *Contraceptive Technology* serves as the primary source of information about contraceptive failure rates and is cited by the World Health Organization, CDC, health care service delivery organizations, and leading health professional associations in the US and other countries.
Provide all other grades and definitions from the recommendation grading system	Not applicable.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Quantity – 3,136 total studies in book, 103 in the chapter on Efficacy, Safety, and Personal Considerations (p. 95-129) Quality – *Contraceptive Technology* serves as the primary source of information about contraceptive failure rates and is cited by the World Health Organization, CDC, and leading professional associations in the US and other countries. Two sources of data are used to estimate contraceptive failure. The first is published research comprised of results from clinical trials and surveys. The second source is CDC's National Survey of Family Growth (NSFG) is used to estimate typical use rates using data from a nationally representative sample of users.
Estimates of benefit and consistency across studies	Key findings of this review are estimated failure rates for a wide range of contraceptive methods under "perfect" and "typical" use. The most recent findings, published in 2018 are that the most effective methods, (LARC and sterilization) have a failure rate less than 1% per year under typical use; the moderately effective methods (shot/Depo, pills/patch/ring (PPR)) have a typical failure fate of 4-7%. PPR typical use failure rates have slightly (6 to 7%) increased from 2011 to 2018 while shot typical use failure rate has dropped from 6% to 4%. Diaphragm typical use failure rates have increased since the 2011 study and are no longer considered moderately effective.
What harms were identified?	Authors state that, "In general, contraceptives pose few serious health risks to users. Moreover, the use of contraceptive methods is generally far safer than pregnancy." (p. 111). The authors state that the absolute level of risk for death is very low for most people and that other major health risks from contraceptive use are uncommon and are most likely to occur in individuals with underlying medical conditions (p. 111).

Systematic Review	Evidence
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Not applicable.

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	• Contraceptive Counseling in Clinical Settings: An Updated Systematic Review • Lauren B Zapata, Karen Pazol, Christine Dehlendorf, Kathryn M. Curtis, Nikita M. Malcolm, Rachel B. Rosmarin, Brittni N. Frederiksen • 2018 November 1 • Lauren B. Zapata, Karen Pazol, Christine Dehlendorf, Kathryn M. Curtis, Nikita M. Malcolm, Rachel B. Rosmarin, Brittni N. Frederiksen, Contraceptive Counseling in Clinical Settings: An Updated Systematic Review, American Journal of Preventive Medicine, Volume 55, Issue 5, 2018, Pages 677-690. • https://doi.org/10.1016/j.amepre.2018.07.006
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.	Overall, evidence supports the utility of contraceptive counseling, in general, and specific interventions or aspects of counseling. Promising components of contraceptive counseling were identified.

Systematic Review	Evidence
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Evidence in this guideline ranges from I to II-3, low to high risk of bias, depending on the age group of study participants (e.g., adolescents, young adults, adults, and mixed populations) and outcome type (e.g., long-, medium-, and short-term outcomes and client experiences). This SR is reported according to the PRISMA checklist. The strength and quality of the evidence in this SR are graded using on the U.S. Preventive Services Task Force (USPSTF) system, which includes ratings for study design (I: randomized controlled trials; II-1: controlled trials without randomization; II-2: observational studies; and II-3: multiple time series or descriptive studies) and risk of bias (low, moderate, high). The studies included in the SR were graded as follows: I: 12 studies (2 high risk, 10 moderate risk) II-1: 3 studies (2 high risk, 1 moderate risk) II-2: 11 studies (9 high risk, 2 moderate risk) II-3: 6 studies (6 high risk) Six key questions (KQs) were developed, and an analytic framework was utilized to describe the relationships between the population of interest; the intervention of interest; and the outcomes of interest.
Provide all other grades and definitions from the evidence grading system	Not applicable. All grades and definitions are included in the box above.
Grade assigned to the **recommendation** with definition of the grade	Not applicable.
Provide all other grades and definitions from the recommendation grading system	Not applicable.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Quantity – 35 articles; 32 studies Quality – 14 RCTs, 2 non-randomized trials, 5 cohort studies, 5 cross-sectional studies, and 6 pre-post studies

Systematic Review	Evidence
Estimates of benefit and consistency across studies	Overall, findings support the provision of contraceptive counseling, compared with no counseling, on contraceptive use behaviors. Six of nine studies among adolescents and young adults and 16 of 23 studies among adults or mixed populations found a statistically significant positive impact of counseling on at least one outcome of interest. Promising components of contraceptive counseling include an emphasis on the quality of interaction between counselor and client (e.g., developing rapport); personalizing discussions to meet clients' individual needs; and addressing psychosocial determine of contraceptive use behaviors (e.g., perceived benefits and barriers, outcome expectations. New components that resulted in some statistically significant positive effects include an emphasis on shared decision making, asking about the patient's reproductive life plan/pregnancy intentions, and discussion of contraceptive methods by level of effectiveness.
What harms were identified?	While the article did not identify any harms of contraceptive counseling, authors stated that following would strengthen the evidence base: improved documentation of counseling content and processes, increased attention to the relationships between client experiences and behavioral outcomes and examining the comparative effectiveness of different counseling approaches to identify those that are most effective.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Not applicable.

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	Clinical Practice Guideline recommendation • Committee Opinion No. 710: Counseling Adolescents About Contraception • ACOG • 2017, reaffirmed 2019 • Committee Opinion No. 710 Summary: Counseling Adolescents About Contraception. (2017). Obstetrics and gynecology, 130(2), 486–487. https://doi.org/10.1097/AOG.0000000000002228 • https://doi.org/10.1097/AOG.0000000000002228
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.	*Regardless of a patient's age or previous sexual activity, the obstetrician-gynecologist routinely should address her contraceptive needs, expectations, and concerns. * Statutes on the rights of minors to consent to health care services vary by state, and obstetrician-gynecologists should be familiar with the regulations that apply to their practice. * Emergency contraception routinely should be included in discussions about contraception, including access issues. The American College of Obstetricians and Gynecologists recommends that obstetrician-gynecologists write advance prescriptions for oral emergency contraception for their patients. * Long-acting reversible contraceptive (LARC) methods have higher efficacy, higher continuation rates, and higher satisfaction rates compared with short-acting contraceptives. Because LARC methods are safe, they are excellent contraceptive choices for adolescents. * Discussions about contraception should begin with information on the most effective methods first. * Obstetrician-gynecologists should be aware of and be prepared to address the most common misperceptions about contraceptive methods in a way that is age appropriate and compatible with the patient's health literacy. * The initial encounter and follow-up visits should include continual reassessment of sexual concerns, behavior, relationships, prevention strategies, and testing and treatment for sexually transmitted infections (STIs) per the Centers for Disease Control and Prevention's (CDC) guidelines.

Systematic Review	Evidence
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Not applicable.
Provide all other grades and definitions from the evidence grading system	Not applicable.
Grade assigned to the **recommendation** with definition of the grade	Not applicable.
Provide all other grades and definitions from the recommendation grading system	Not applicable.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Not applicable.
Estimates of benefit and consistency across studies	Modern contraceptives are very effective when used correctly and, thus, effective counseling regarding contraceptive options and provision of resources to increase access are key components of adolescent health care. Regardless of a patient's age or previous sexual activity, the obstetrician-gynecologist routinely should address her contraceptive needs, expectations, and concerns. Obstetrician-gynecologists should be aware of and be prepared to address the most common misperceptions about contraceptive methods in a way that is age appropriate and compatible with the patient's health literacy. The American College of Obstetricians and Gynecologists recommends that discussions about contraception begin with information on the most effective methods first.
What harms were identified?	At no time should an adolescent patient be forced to use a method chosen by someone other than herself, including a parent, guardian, partner, or health care provider.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Not applicable.

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	Clinical Practice Guideline recommendation • "Committee Opinion No. 642: Increasing Access to Contraceptive Implants and Intrauterine Devices to Reduce Unintended Pregnancy." • ACOG • 2015, reaffirmed 2018 • Committee Opinion No. 642: Increasing Access to Contraceptive Implants and Intrauterine Devices to Reduce Unintended Pregnancy. (2015). Obstetrics and gynecology, 126(4), e44–e48. https://doi.org/10.1097/AOG.0000000000001106 • https://doi.org/10.1097/AOG.0000000000001106
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 all women at risk of unintended pregnancy, obstetrician-gynecologists should provide counseling on all contraceptive options, including implants and IUDs. * Encourage consideration of implants and IUDs for all appropriate candidates, including nulliparous women and adolescents. * Adopt best practices for LARC insertion. * Advocate for coverage and appropriate payment and reimbursement for every contraceptive method by all payers in all clinically appropriate circumstances. * Become familiar with and support local, state (including Medicaid), federal, and private programs that improve affordability of all contraceptive methods.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Not applicable.
Provide all other grades and definitions from the evidence grading system	Not applicable.
Grade assigned to the **recommendation** with definition of the grade	Not applicable.
Provide all other grades and definitions from the recommendation grading system	Not applicable.

Systematic Review	Evidence
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Not applicable.
Estimates of benefit and consistency across studies	Unintended pregnancy persists as a major public health problem in the United States. Although lowering unintended pregnancy rates requires multiple approaches, individual obstetrician-gynecologists may contribute by increasing access to contraceptive implants and intrauterine devices. Obstetrician-gynecologists should encourage consideration of implants and intrauterine devices for all appropriate candidates, including nulliparous women and adolescents. Obstetrician-gynecologists should adopt best practices for long-acting reversible contraception insertion. Obstetrician-gynecologists are encouraged to advocate for coverage and appropriate payment and reimbursement for every contraceptive method by all payers in all clinically appropriate circumstances.
What harms were identified?	Not applicable.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Not applicable.

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.

2021 Submission

Not applicable.

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

2021 Submission

Not applicable.

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

2021 Submission

Not applicable.

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

2021 Submission

Not applicable.

References

2021 References

Gavin, L., Moskosky, S., Carter, M., Curtis, K., Glass, E., Godfrey, E., Marcell, A., Mautone-Smith, N., Pazol, K., Tepper, N., Zapata, L., & Centers for Disease Control and Prevention (CDC) (2014). Providing quality family planning services: Recommendations of CDC and the U.S. Office of Population Affairs. *MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 63*(RR-04), 1–54.

Gavin, L. E., Ahrens, K. A., Dehlendorf, C., Frederiksen, B. N., Decker, E., & Moskosky, S. (2017). Future directions in performance measures for contraceptive care: a proposed framework. *Contraception*, *96*(3), 138–144. https://doi.org/10.1016/j.contraception.2017.06.001

2016 Submission below

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

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

Measure Title: Contraceptive Care – Postpartum

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

Date of Submission: **2/15/2016**

Instructions

- *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.*
- Respond to all questions as instructed with answers immediately following the question. 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.
- Maximum of 10 pages (*incudes 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.

Note: The information provided in this form is intended to aid the Steering 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:

- Health outcome:[3] a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- 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.

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 U.S. Preventive Services Task Force (USPSTF) grading definitions and methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.
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

☐ Health outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors

☒ Intermediate clinical outcome (*e.g., lab value*): **Contraceptive use**

☐ Process:

☐ Structure:

☐ Other:

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

Not a health outcome or PRO.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

***Note**: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.*

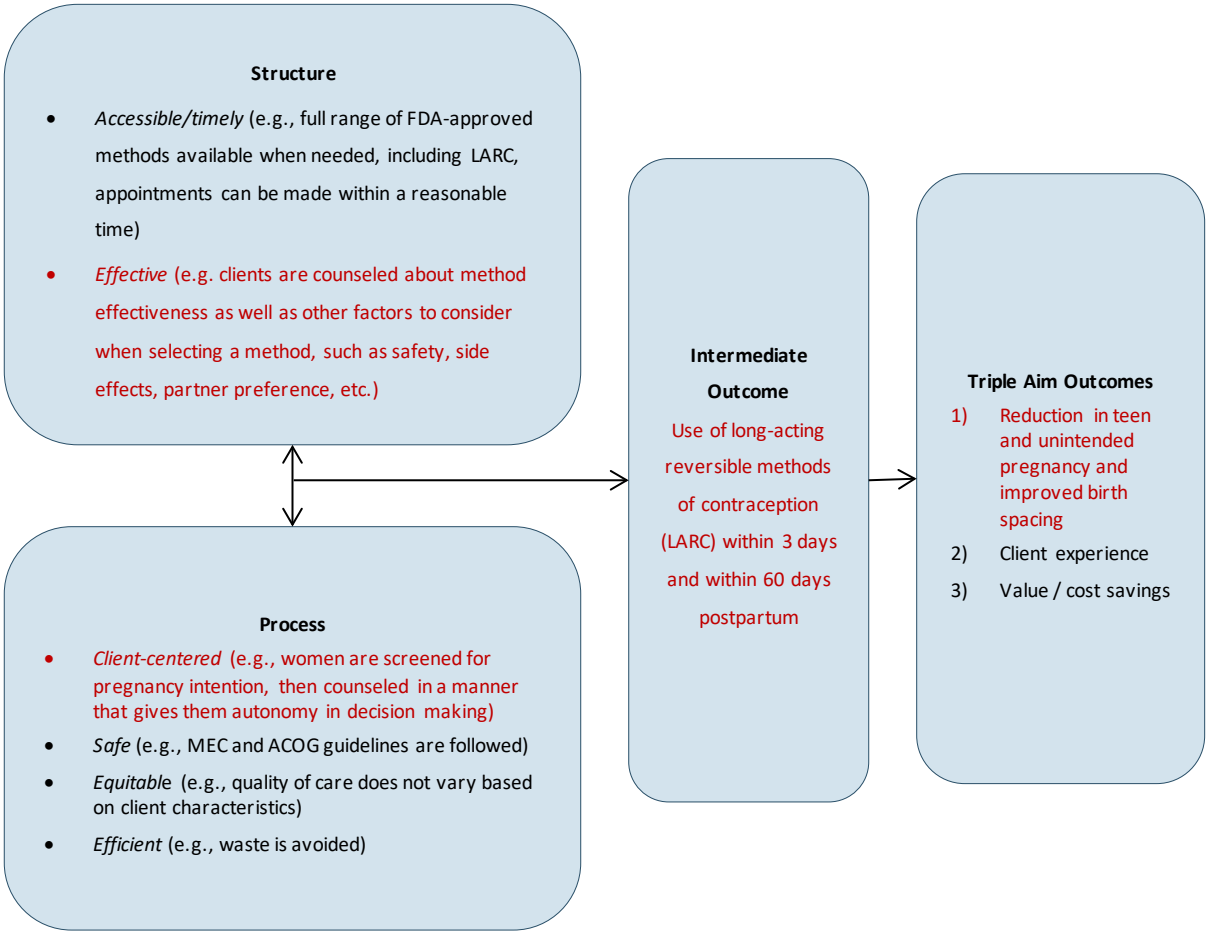
INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.

The diagram below illustrates the steps between the structure and process that influence the intermediate health outcome, and how the intermediate health outcome in turns influences the longer-term outcomes. The text highlighted in red shows the primary relationships that will be affected by use of the proposed measure: (a) increased use of the most and moderately effective methods of contraception will influence rates of unintended pregnancy; and (b) appropriate counseling of a client can lead to increased use of the most and moderately effective methods of contraception.

The type of contraceptive method used by a woman is strongly associated with her risk of unintended pregnancy. The **most** effective methods (sterilization and the long-acting reversible contraceptive [LARC] methods of intrauterine devices and implants) have a failure rate that is less than 1% per year under typical use; the **moderately** effective methods (shot, oral pills, patch, ring, and diaphragm) have a typical failure rate of 6-12% per year; the least effective methods have a typical failure rate of 18-28%; and if no method is used then 85 of every 100 women will become pregnant in a year (Trussell 2011).

The measure is secondarily supported by evidence that the way in which contraceptive counseling is offered (e.g., increased screening of clients for reproductive intention; the provision of client-centered counseling, which includes providing information about and ready access to the most and moderately effective methods of contraception; and ready access to all methods of contraception, ideally on a same-day basis) will lead to increased use of the most and moderately effective methods of contraception (i.e., the intermediate outcome).



1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?

☒ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***

☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***

☒ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***

☐ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

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

Clinical recommendations (from both government sources and professional organizations) are the best source of evidence about the relationship between contraceptive counseling and increased use of the most and moderately effective methods of contraception (see diagram above).

CDC/OPA (2014). Providing Quality Family Planning Services (QFP): Recommendations of CDC and the US Office of Population Affairs, MMWR Recommendations and Reports, April 24, 2014. http://www.cdc.gov/mmwr/preview/mmwrhtml/rr6304a1.htm

American College of Obstetricians and Gynecologists (ACOG), Committee on Gynecologic Practice. Increasing access to contraceptive implants and intrauterine devices to reduce unintended pregnancy. Committee Opinion Number 642; October 2015.

ACOG *Long-acting reversible contraception: Implants and intrauterine devices*, in *Practice Bulletin*. 2015 (reaffirmed), American College of Obstetricians and Gynecologists: Washington, DC. p. 1-13.

The American Academy of Pediatrics (AAP) (2014). Contraception for Adolescents. Pediatrics, 134:e1244–e1256.

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

"Providers are encouraged to present information on potential reversible methods of contraception by using a tiered approach (i.e., presenting information on the most effective methods first, before presenting information on less effective methods). This information should include an explanation that long-acting reversible contraceptive methods are safe and effective for most women, including those who have never given birth and adolescents. Information should be tailored and presented to ensure a client-centered approach. It is not appropriate to omit presenting information on a method solely because the method is not available at the service site. If not all methods are available at the service site, it is important to have strong referral links in place to other providers to maximize opportunities for clients to obtain their preferred method that is medically appropriate."

Source: CDC/OPA (2014). *Providing Quality Family Planning Services*, page 8 and Appendix B

"For all women at risk of unintended pregnancy, obstetrician-gynecologists should provide counseling on all contraceptive options, including implants and IUDs. Long-acting reversible contraception methods require a single action of motivation for long-term use, eliminating adherence and user dependence from the effectiveness equation. These top-tier methods share the highest continuation rates of all contraceptives, which is one of the most important factors in contraceptive success."

Source: ACOG (2015), page 1.

"The immediate postpartum period is a particularly favorable time for IUD or implant insertion. Women who have recently given birth are often highly motivated to use contraception, they are known not to be

pregnant and the hospital setting offers convenience for both the patient and the health care provider." ACOG (2015 Practice Bulletin), page 4.

"Contraceptive methods most commonly used by adolescents are listed below, ordered from most to least effective, starting with long-acting reversible contraception (LARC); implants and IUDs. *Pediatricians are encouraged to counsel adolescents in that order, discussing the most effective contraceptive methods first*." ACOG (2014), page e1246.

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

Not applicable

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

Not applicable

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

Not applicable

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

☒ Yes → ***complete section 1a.7***

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

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

Not applicable

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

Not applicable

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

Not applicable

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

Not applicable

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

Not applicable

Complete section 1a.7

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

Two systematic literature reviews are the best source of evidence about the relationship between use of long-acting reversible methods of contraception (LARC) and unintended pregnancy (see diagram in 1a.3, above). A third systematic review focused on the provision of LARC methods in the immediate postpartum period.

1. The first review was led by Professor James Trussell from Princeton University, which is repeated on an ongoing basis and published in a handbook entitled "Contraceptive Technology". The Trussell analyses serve as the primary source of information about contraceptive failure rates, and are cited by the World Health Organization, CDC, and leading professional associations in the U.S. and in other countries. Trussell used two sources of data when estimating contraceptive failure. The first was published research, which comprised results from clinical trials and surveys. The second source was the CDC's National Survey of Family Growth (NSFG), which was used to estimate *typical* use rates using data from a nationally representative sample of users.

- Trussell J (2011). Contraceptive efficacy. In: Hatcher RA, Trussell J, Nelson AL, Cates W, Kowal D, editors. Contraceptive technology: twentieth revised edition. New York: Ardent Media; 2011, pp. 777–861. This was subsequently summarized in: Trussell J (2011). Contraceptive failure in the United States. Contraception; 83(5):397-404.

- WHO/Department of Reproductive Health and Research & Johns Hopkins Bloomberg School of Public Health/Center for Communication Programs (2011). Family Planning: A Global Handbook for Providers. Baltimore and Geneva: CCP and WHO.

2. The second review was conducted by Mansour et al in 2010. They search Medline and Embase from January 1990 to February 2008 for publications reporting contraceptive failure rates.

    - Mansour D, Inki P, Gemzell-Danielsson K (2010). Efficacy of contraceptive methods: A review of the literature. The European Journal of Contraception and Reproductive Health

Care, 15:4-16.

3. A recent Cochrane systematic review examined the outcomes of IUD insertion immediately after placement delivery (within 10 minutes). Randomized clinical trials published through April 1, 2015 were identified in the following databases: PubMed, CENTRAL, POPLINE, Web of Science, EMBASE, LILACS, ClinicalTrials.gov, and ICTRP.

   - Lopez, L.M., et al., *Immediate postpartum insertion of intrauterine device for contraception.* Cochrane Database Syst Rev, 2015. **6**: p. CD003036.

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Evidence of contraceptive effectiveness & its impact on unintended pregnancy

- See 1a.6.1 above

Evidence of effectiveness of counseling or other interventions to affect patients' choice of method

- Zapata LB, Tregear SJ, Curtis KM, Tiller M, Pazol K, Mautone-Smith N, Gavin LE (2015). Impact of Contraceptive Counseling in Clinical Settings: A Systematic Review. Am J Prev Med. 2015 Aug;49(2 Suppl 1):S31-45.

Complete section 1a.7

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

Evidence of contraceptive effectiveness & its impact on unintended pregnancy

The studies examining contraceptive efficacy and effectiveness considered the impact of use of specific contraceptive methods on risk of pregnancy (i.e., contraceptive failure). Pregnancy risk can be assessed either through life table analyses (usually through 12 months) that show the percentage of women who become pregnant, or the score on the Pearl Index. The Pearl Index is a commonly used technique for reporting the effectiveness of a **birth control** method in clinical trials, and estimates the number of **unintended pregnancies** over a period of exposure (e.g. 100 women over one year of use, or 10 women over 10 years). Contraceptive failure rates are reported for *perfect use* and *typical use*. Perfect use reflects how effective methods can be in preventing pregnancy when used consistently and correctly

according to instructions. Typical use reflects how effective methods are for the average person who does not always use methods correctly or consistently. Pregnancy rates during typical use of adherence-dependent methods (such as the oral pill) generally vary widely for different groups using the same method, primarily due to differences in the propensity to use the method perfectly. The review by Lopez et al (2015) focused on immediate postpartum insertion of IUDs (within 10 minutes) compared immediate insertion to insertion at other postpartum times. Key outcomes were expulsion and method use.

Evidence of effectiveness of counseling or other interventions to affect patients' choice of method

The systematic review underpinning the CDC-OPA recommendation on contraceptive counseling used an analytic framework that considered the impact of providing contraceptive counseling and/or education on short (e.g., client knowledge, attitudes), medium (e.g., selection of more effective methods, correct and consistent use) and long-term (unintended pregnancy) outcomes (Zapata 2015).

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

Evidence of contraceptive effectiveness & its impact on unintended pregnancy

While the quality of the studies was not graded in either the Trussell (2011) or Mansour (2010) review, they were primarily comprised of randomized controlled trials. The Lopez (2015) review applied principles from GRADE (Grades of Recommendation, Assessment, Development and Evaluation) to assess the quality of evidence as shown below, and found the body of evidence to be of moderate quality:

- **High quality:** Further research is very unlikely to change our confidence in the estimate of effect.
- **Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
- **Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- **Very low quality:** We are very uncertain about the estimate.

Evidence of effectiveness of counseling or other interventions to affect patients' choice of method

The review did not grade the *overall* body of evidence. However, the quality of **individual studies** was graded in accordance with USPSTF methodologies for doing so, i.e., Level I, Level II-1, Level II-2, Level II-3, Level III.

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

Not applicable

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range

Trussell (2011): 1958-2010

Mansour (2010): January 1990 to February 2008

Lopez (2015): through April 1, 2015

Zapata (2015): 1985-February 2011 with supplemental searches through 2014

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

Evidence of contraceptive effectiveness & its impact on unintended pregnancy

- **Trussell et al 2011**: The review comprised results from clinical trials and surveys; the most recent review listed more than 350 studies, of which the majority was randomized controlled trials (Trussell 2011a).

- **Mansour et al 2010**: The authors identified and extracted information from 139 publications. Of the included studies, 47 assessed combined oral contraceptives (COCs), one assessed progestogen-only pills (POPs), three assessed the patch, three assessed the vaginal ring, 15 assessed implants, 16 assessed injectables, 31 assessed copper intrauterine devices (Cu-IUDs), nine assessed the levonorgestrel-releasing intrauterine system (LNGIUS), three assessed the male condom, four assessed other barrier methods, 11 assessed natural methods, and four assessed female sterilization. Overall, there were 64 publications of randomized controlled studies included in this review. A detailed description of each publication can be accessed from www.informahealthcare.com/doi/pdf/10.3109/13625180903427675.

- **Lopez (2015).** Fifteen RCTs were identified, with seven studies reported from 2010-2014.

Evidence of effectiveness of counseling or other interventions to affect patients' choice of method

- **Zapata et al (2015)**: 22 studies (from 23 articles) met the inclusion criteria; 8 studies included use of more effective methods as an outcome. Seven of the 8 studies were randomized controlled trials, while the eighth utilized a pre-posttest study design.

1a.7.6. What is the overall quality of evidence across studies in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

The quality of evidence is not described in either the Trussell (2011) or the Mansour (2010) publications. However, both reviews are substantially comprised of randomized controlled trials. The Lopez (2015) review determined that the overall body of evidence (comprised of 15 RCTs) was of moderate quality.

In Zapata et al (2011), 7 of the 8 studies were graded Level I (properly designed randomized controlled trial), and the 8th study was graded Level II-3 (evidence obtained from time series, uncontrolled trial).

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) across studies in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

Evidence of contraceptive effectiveness & its impact on unintended pregnancy

- **Trussell et al 2011**: The key findings of this review are estimated failure rates for a wide range of contraceptive methods under "perfect" and "typical" use. The most recent findings – published in 2011 -- are that the **most** effective methods (LARC and sterilization) have a failure rate that is less than 1% per year under typical use; the **moderately** effective methods (shot, PPR, diaphragm) have a typical failure rate of 6-12% per year; the least effective methods have a typical failure rate of 18-28%; and not using any method at all has a failure rate of 85%.

- **Mansour et al 2010**: "Information was identified and extracted from 139 studies. One-year Pearl Indices reported for short-acting user-dependent hormonal methods were generally less than 2.5. Gross life-table rates for long-acting hormonal methods (implants and the levonorgestrel releasing-intrauterine system [LNG-IUS]) generally ranged between 0–0.6 per 100 at one year, but wider ranges (0.1–1.5 per 100) were observed for the copper intrauterine devices (0.1–1.4 per 100 for Cu-IUDs with surface area _300 mm2 and 0.6–1.5 per 100 for those with surface area5300 mm2). Barrier and natural methods were the least effective." The authors conclude that "the review broadly confirmed the hierarchy of contraceptive effectiveness in descending order as: (1) female sterilisation, long-acting hormonal contraceptives (LNG-IUS and implants); (2) Cu-IUDs with_300 mm2 surface area; (3) Cu-IUDs with5300 mm2 surface area and short-acting hormonal contraceptives (injectables, oral contraceptives, the patch and vaginal ring), and (4) barrier methods and natural methods."

- **Lopez (2015)**: A meta-analysis showed that IUC use at six months was more likely with immediate insertion than with standard insertion (OR 2.04; 95% CI 1.10 to 4.09; participants=243; studies=4). Expulsion was more likely for the immediate group, but the confidence interval was wide (OR 4.89; 95% CI 1.47 to 16.32; participants =210; studies=4). The review concludes that the "benefit of effective contraception immediately after delivery may outweigh the disadvantage of increased risk for expulsion. Frequent prenatal visits during the third trimester provide the opportunity to discuss

effective contraceptive methods and desired timing for initiation. Clinical follow-up can help detect early expulsion, as can educating women about expulsion signs and symptoms."

Evidence of effectiveness of counseling or other interventions to affect patients' choice of method

- **Zapata (2015)**: Five of the 8 studies that examined use of more effective methods found an increased rate of use in the intervention vs control/comparison conditions. Three studies found no significant impact. No studies found a decreased rate of use of more effective contraceptive methods.

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

The harms were not noted in the cited reviews. However, CDC clinical recommendations on contraceptive safety explicitly address this question. CDC's "US Medical Eligibility Criteria for Contraceptive Use" (USMEC) describe what contraceptive methods are safe for women with a range of characteristics (e.g., age, postpartum) and medical conditions (e.g., infectious or chronic diseases). The citation for the USMEC recommendations is:

> CDC (2010). US Medical Eligibility Criteria for Contraceptive Use, MMWR Recommendations and Reports, 59 (RR04):1–85. Available online at: http://www.cdc.gov/reproductivehealth/UnintendedPregnancy/USMEC.htm."

The evidence on which the USMEC recommendations are based has been summarized in the following journal supplement:

> Contraception, Volume 82, Issue 1, Pages 1-118 (July 2010). Available online at: http://www.sciencedirect.com/science/journal/00107824/82/1

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for each new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

Results from two large studies have been recently published, which provide additional evidence that: (a) long-acting reversible methods of contraception (LARC) are associated with reduced risk of unintended pregnancy, and (b) that the type of counseling provided is associated with selection of LARC methods by the client. The first study is a cluster-randomized trial led by researchers at the University of California – San Francisco (Harper 2015) and the second is a prospective cohort study that is known as "Project CHOICE" (Winner 2012).

UCSF trial (Harper et al 2015)

A cluster randomized trial was conducted in 2011-2013 to assess the effects of an intervention to increase patients' access to long-acting reversible contraceptives (LARCs) on pregnancy rates. A total of 40 clinics participated: 20 clinics were randomly assigned to receive evidence-based training on providing counselling and insertion of intrauterine devices (IUDs) or progestin implants, and 20 to provide standard care. Usual costs for contraception were maintained at all sites. Women aged 18-25 years attending family planning or abortion care visits and not desiring pregnancy in the next 12 months were recruited. The primary outcome was selection of an IUD or implant at the clinic visit and secondary outcome was pregnancy within 12 months. Generalised estimating equations for clustered data were used to measure the intervention effect on contraceptive selection, and survival analysis was used to assess pregnancy rates. Of 1500 women enrolled, more at intervention than control sites reported receiving counselling on IUDs or implants (565 [71%] of 797 vs 271 [39%] of 693, odds ratio 3·8, 95% CI 2·8-5·2) and more selected LARCs during the clinic visit (224 [28%] vs 117 [17%], 1·9, 1·3-2·8). The pregnancy rate was lower in intervention group than in the control group after family planning visits (7·9 vs 15·4 per 100 person-years), but not after abortion visits (26·5 vs 22·3 per 100 person-years). We found a significant intervention effect on pregnancy rates in women attending family planning visits (hazard ratio 0·54, 95% CI 0·34-0·85).

- Harper C, Rocca CH, Thompson KM, Morfesis J, Goodman S, Darney PD, Westhoff CL, Speidel JJ (2015). Reductions in pregnancy rates in the USA with long-acting reversible contraception: a cluster randomised trial. Lancet. Volume 386, No. 9993, p562–568, 8 August 2015

Project CHOICE (Secura et al 2014, Winner et al 2015)

The Contraceptive CHOICE Project was a prospective cohort study involving 9256 St. Louis area adolescent and adult women 14 to 45 years of age, in which women were counseled about the use of LARC methods to prevent unintended pregnancy. Participants were educated about reversible contraception, with an emphasis on the benefits of LARC methods, were provided with their choice of reversible contraception at no cost, and were followed for 2 to 3 years. Almost three-quarters of enrolled participants chose a LARC method when they were counseled about effectiveness and offered their choice of method at no charge, and continuation rates were high 2 years (77% for LARC users vs 41% for non-LARC users) and 3 years (67% for LARC users vs 31% for non-LARC users) after insertion. The contraceptive failure rate among participants using pills, patch, or ring was 4.55 per 100 participant-years, as compared with 0.27 among participants using long-acting reversible contraception (hazard ratio after adjustment for age, educational level, and history with respect to unintended pregnancy, 21.8; 95% confidence interval, 13.7 to 34.9).

- Winner B, Peipert J, Qiuhong Z, Buckel C, Madden T et al (2012). Effectiveness of Long-Acting Reversible Contraception, The New England Journal of Medicine, 366 (21): 1998-2007
- Diedrich, J.T., et al., *Three-year continuation of reversible contraception.* Am J Obstet Gynecol, 2015. **213**(5): p. 662 e1-8.
- O'Neil-Callahan, M., et al., *Twenty-four-month continuation of reversible contraception.* Obstet Gynecol, 2013. **122**(5): p. 1083-91.

1a.8 OTHER SOURCE OF EVIDENCE – not applicable

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.8.1 What process was used to identify the evidence?

1a.8.2. Provide the citation and summary for each piece of 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.*

Unintended pregnancies and interpregnancy intervals of less than 18 months have been associated with poor perinatal outcomes such as preterm birth, low birth weight, small size for gestational age, as well as adverse maternal outcomes [1, 2]. Studies among U.S. women report that women at younger maternal age are at higher risk for unintended pregnancy [14] and older maternal age is associated with closely spaced pregnancies [15]. Contraception is a highly effective clinical preventive service that can assist women in reaching their reproductive health goals, like reducing unintended pregnancies and the percentage of births occurring within 18 months of a previous birth [3, 4]. The type of contraceptive method used by a woman is strongly associated with her risk of unintended pregnancy. The most effective methods (LARC and sterilization) have a failure rate that is less than 1% per year under typical use [4]. The moderately effective methods (injectable, pill, patch, ring) have a typical failure rate of 4-7% per year, while the less effective methods have a typical failure rate of 13-27% [4]. One recent study also indicates that the most used contraceptive methods in the United States have experienced reductions in their typical use failure rates [16]. Not using any method at all has a typical failure rate of 85% [4].

After NQF endorsed #2903 in 2016, OPA published multiple articles in peer-reviewed journals to inform health care providers in public and private settings (e.g., commercial health plans, Medicaid, community health centers, free-standing reproductive health clinics) about the new measure. These publications outline our conceptual framework for developing #2903 alongside its two complementary measures (NQF #2902 and #2904) and describe appropriate measure implementation and use. Furthermore, OPA highlighted systematic reviews which indicate that effective contraceptive method use increases the interbirth interval and reduces adolescent and unintended pregnancies. This association between use of most and moderately effective methods and positive reproductive health outcomes demonstrates the importance of contraceptive care measures to health care quality [17-19].

While NQF #2903 and the contraceptive care measures reflect that some contraceptive methods are more effective than others at preventing pregnancy, these measures and their guidelines for use are

designed to encourage providers to offer those clients seeking contraception the full range of methods. The goal of providing contraception should never be to recommend any one method or class of methods over women's individual choices. Women who want to delay or prevent pregnancy should have access to a broad range of contraceptive methods, preferably on a same-day, on-site basis. Furthermore, it is important that these contraceptive services are provided in a client-centered manner that treats each person as a unique individual with respect, empathy, and understanding, providing accurate, easy-to-understand information based on the client's self-identified needs, goals, preferences, and values [11]. Patients receiving client-centered care may feel motivated to continue seeking reproductive health care for contraception and if they become pregnant, prenatal care and birth [13]. Thus, efforts to provide client-centered contraceptive services aligned with the American Academy of Pediatrics (AAP), the American College of Obstetricians and Gynecologists (ACOG), the Centers for Disease Control and Prevention (CDC), and Office of Population Affairs (OPA) recommendations [7-12] may be strengthened by quality improvement processes based on standardized metrics of contraceptive care provision.

References

[1] Conde-Agudelo, A., Rosas-Bermúdez, A., & Kafury-Goeta, A. C. (2006). Birth spacing and risk of adverse perinatal outcomes: a meta-analysis. JAMA, 295(15), 1809–1823. https://doi.org/10.1001/jama.295.15.1809

[2] Conde-Agudelo A, Rosas-Bermúdez A, Kafury-Goeta AC. Effects of birth spacing on maternal health: a systematic review. Am J Obstet Gynecol. 2007 Apr;196(4):297-308. doi: 10.1016/j.ajog.2006.05.055. PMID: 17403398.

[3] Mansour, D., Inki, P., & Gemzell-Danielsson, K. (2010). Efficacy of contraceptive methods: A review of the literature. The European journal of contraception & reproductive health care: the official journal of the European Society of Contraception, 15(1), 4–16. https://doi.org/10.3109/13625180903427675

[4] Trussell, J., Aiken, A.R.A., Micks, E., Guthrie, K.A. (2018). Efficacy, safety, and personal considerations. In R.A. Hatcher, A.L. Nelson, J. Trussell, C. Cwiak, P. Cason, M.S. Policar, A. Edelman, A.R.A. Aiken, J. Marrazzo, D. Kowal (Eds.). Contraceptive technology (21st ed., pp. 95–128). Ayer Company Publishers, Inc.

[5] Winner, B., Peipert, J. F., Zhao, Q., Buckel, C., Madden, T., Allsworth, J. E., & Secura, G. M. (2012). Effectiveness of long-acting reversible contraception. The New England journal of medicine, 366(21), 1998–2007. https://doi.org/10.1056/NEJMoa1110855

[6] Harper, C. C., Rocca, C. H., Thompson, K. M., Morfesis, J., Goodman, S., Darney, P. D., Westhoff, C. L., & Speidel, J. J. (2015). Reductions in pregnancy rates in the USA with long-acting reversible contraception: a cluster randomised trial. Lancet (London, England), 386(9993), 562–568. https://doi.org/10.1016/S0140-6736(14)62460-0

[7] Committee on Gynecologic Practice Long-Acting Reversible Contraception Working Group (2015). Committee Opinion No. 642: Increasing Access to Contraceptive Implants and Intrauterine Devices to Reduce Unintended Pregnancy. Obstetrics and gynecology, 126(4), e44–e48. https://doi.org/10.1097/AOG.0000000000001106

[8] ACOG Committee Opinion No. 735: Adolescents and Long-Acting Reversible Contraception: Implants and Intrauterine Devices. (2018). Obstetrics and gynecology, 131(5), e130–e139. https://doi.org/10.1097/AOG.0000000000002632

[9] Menon, S., & COMMITTEE ON ADOLESCENCE (2020). Long-Acting Reversible Contraception: Specific Issues for Adolescents. Pediatrics, 146(2), e2020007252. https://doi.org/10.1542/peds.2020-007252

[10] Committee on Adolescence (2014). Contraception for adolescents. Pediatrics, 134(4), e1244–e1256. https://doi.org/10.1542/peds.2014-2299

[11] Gavin, L., Moskosky, S., Carter, M., Curtis, K., Glass, E., Godfrey, E., Marcell, A., Mautone-Smith, N., Pazol, K., Tepper, N., Zapata, L., & Centers for Disease Control and Prevention (CDC) (2014). Providing quality family planning services: Recommendations of CDC and the U.S. Office of Population Affairs. MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 63(RR-04), 1–54.

[12] Curtis, K. M., Tepper, N. K., Jatlaoui, T. C., Berry-Bibee, E., Horton, L. G., Zapata, L. B., Simmons, K. B., Pagano, H. P., Jamieson, D. J., & Whiteman, M. K. (2016). U.S. Medical Eligibility Criteria for Contraceptive Use, 2016. MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 65(3), 1–103. https://doi.org/10.15585/mmwr.rr6503a1

[13] Gomez, A. M., & Wapman, M. (2017). Under (implicit) pressure: young Black and Latina women´s perceptions of contraceptive care. Contraception, 96(4), 221–226. https://doi.org/10.1016/j.contraception.2017.07.007

[14] Finer, L.B. and Zolna, M.R. (2016). Declines in unintended pregnancy in the United States, 2008-2011. New Engl J Med, 374(9), 843-52

[15] Thoma, M. E., Copen, C. E., & Kirmeyer, S. E. (2016). Short Interpregnancy Intervals in 2014: Differences by Maternal Demographic Characteristics. NCHS data brief, (240), 1–8.

[16] Sundaram, A., Vaughan, B., Kost, K., Bankole, A., Finer, L., Singh, S., & Trussell, J. (2017). Contraceptive Failure in the United States: Estimates from the 2006-2010 National Survey of Family Growth. Perspectives on sexual and reproductive health, 49(1), 7–16. https://doi.org/10.1363/psrh.12017

[17] Gavin, L., Frederiksen, B., Robbins, C., Pazol, K., & Moskosky, S. (2017). New clinical performance measures for contraceptive care: their importance to healthcare quality. Contraception, 96(3), 149–157. https://doi.org/10.1016/j.contraception.2017.05.013

[18] Gavin, L. E., Ahrens, K. A., Dehlendorf, C., Frederiksen, B. N., Decker, E., & Moskosky, S. (2017). Future directions in performance measures for contraceptive care: a proposed framework. Contraception, 96(3), 138–144. https://doi.org/10.1016/j.contraception.2017.06.001

[19] Moniz, M. H., Gavin, L. E., & Dalton, V. K. (2017). Performance Measures for Contraceptive Care: A New Tool to Enhance Access to Contraception. Obstetrics and gynecology, 130(5), 1121–1125. https://doi.org/10.1097/AOG.0000000000002314

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.*

Performance scores for this contraceptive care measure (NQF #2903) are presented for nine programs: federal Medicaid efforts to support state use of the measures; five state Medicaid programs (i.e., the Iowa Medicaid Enterprise, Louisiana Medicaid, the Washington State Health Care Authority, MassHealth and Oregon Medicaid); and one outpatient clinic network within an academic health system (NewYork-Presbyterian Hospital/Columbia University). We also include data from two national organizations that focus on the delivery of reproductive health services (i.e., the Planned Parenthood Federation of America and the Title X program); however, the measure is calculated and interpreted somewhat

differently than the NQF specifications (e.g., the denominator is comprised of women seeking care from the reproductive health clinics). We analyzed NQF #2903 at the following levels: Clinician group/practice, Facility, Health Plan, Public Health Region, and State. When data were available, we also examined trends over time, starting in 2016, the year that NQF #2903 was initially endorsed. We include descriptive statistics for each program and level of analysis below. For more details, see the attached Testing Attachment.

1. Centers for Medicaid & Medicare Services (CMS): Maternal and Infant Health Initiative, Core Measure Set

The three contraceptive care measures were included as part of CMS' Maternal and Infant Health Initiative from 2015 to 2018 and the median measure scores were reported for Federal Fiscal Year (FFY) 2016 and 2017.

FFY 2016 Median Measure Scores:

Ages 15-20: 30.5

Ages 21-44: 26.3

FFY 2017 Median Measure Scores:

Ages 15-20: 30.8

Ages 21-44: 25.6

Although #2903 has been adopted into CMS' Adult and Child Core Set, the measure performance for adult women ages 21-44 have not yet been reported because fewer than 25 states have reported the measure. In FFY 2018, #2903 were reported for the first time in the Child Core Set for women ages 15-20 and then again in FFY 2019. The measure score went up slightly for #2903.

FFY 2018 Measure Scores Ages 15-20

Median: 28.1

Range: 7.6 – 39.0

FFY 2019 Measure Scores Ages 15-20

Median: 29.5

Range: 1.4 – 98.0

2. Iowa Medicaid Enterprise (IME)

The IME analysis included 116,892 women who received services from January 1 through December 31, 2018. The results showed that 30.7% of clients ages 15-44 were provided a most or moderately effective method of contraception. There was variation by public health region (n = 6) and clinician group/facility (n=3,081). For more details, see the Testing Attachment.

Dates included: January 1 through December 31, 2018

Number of measured entities: 3,081 Clinician Groups/Practices

Mean performance score: 26.77

Standard deviation: 26.94

Range: 0.00 – 100.0

Percentiles:

25th: 0.00

50th: 25.00

75th: 40.00

Scores by decile

0 – 10: 1086

11 – 20: 294

21 – 30: 458

31 – 40: 510

41 – 50: 370

51 – 60: 88

61 – 70: 57

71 – 80: 27

81 – 90: 5

91 – 100: 186

Number of measured entities: 6 Public Health Regions (Population Equivalents)

Mean performance score: 31.11

Standard deviation: 2.14

Range: 28.81 – 34.78

Overall Measure Scores for IME (State)

2015

Ages 15-44: 31.5

Ages 15-20: 37.1

Ages 21-44: 29.6

2016

Ages 15-44: 35.6

Ages 15-20: 41.2

Ages 21-44: 33.7

2017

Ages 15-44: 33.5

Ages 15-20: 39.8

Ages 21-44: 31.4

2018

Ages 15-44: 30.7

Ages 15-20: 36.8

Ages 21-44: 28.7

2019

Ages 15-44: 34.0

Ages 15-20: 41.5

Ages 21-44: 31.5

3. Louisiana Medicaid (LA Medicaid)

The LA Medicaid analysis included 279,100 female Medicaid clients who resided in 64 parishes and participated in 5 health plans. About 23.1% of clients aged 15-44 years were provided a most or moderately effective method of contraception; the measure scores varied by health plan. For more details, see the Testing Attachment.

Number of measured entities: 5 Health Plans

Number of female clients ages 15-44: 279,100

Dates included: January 1 through December 31, 2019

Mean performance score: 30.7

Range: 29.0 – 32.2

4. Washington State Health Care Authority (WA HCA)

The WA HCA analysis included 196,568 female Medicaid clients who resided in 39 counties and participated in 5 health plans. About 29.6% of clients aged 15-44 years were provided a most or moderately effective method of contraception; the measure scores varied by health plan. For more details, see the Testing Attachment.

Number of measured entities: 5 Health Plans

Number of female clients ages 15-44:

Dates included: January 1 through December 31, 2019

Mean performance score: 29.0

Range: 27.7 – 30.7

WA HCA published a report in October 2020 the presents trends over time in the measure scores. However, the age categories do not align with the measure specifications. For more information about these trends and steps that Washington State has taken to expand access to contraception, see: https://www.dshs.wa.gov/sites/default/files/rda/reports/research-7-119.pdf.

5. Massachusetts Medicaid (MassHealth)

MassHealth analysis included 197,529 female Medicaid clients who resided in 14 counties and participated in 21 health plans. Sixteen of these health plans were accountable managed care organizations. About 23.1% of clients aged 15-44 years were provided a most or moderately effective method of contraception; the measure scores varied by health plan. For more details, see the Testing Attachment.

Number of measured entities: 21 Health Plans

Number of female clients ages 15-44: 197,529

Dates included: January 1 through December 31, 2019

Mean performance score: 23.2

Range: 18.4 – 26.0

6. Oregon Medicaid

A recent state-level, claims-based cohort study in Oregon evaluated Oregon's use of the contraceptive measures and assessed whether an association exists between implementing an incentive metric and effective contraceptive use within the Oregon Medicaid program, which was introduced on January 1, 2015 [1].

The study period covered 2012-2017, and participants included adult women at risk of pregnancy (18-50 years of age) living in Oregon and enrolled in Medicaid or in commercial health insurance. Compared to the commercially insured comparison group, effective contraceptive use among Medicaid enrollees for all ages combined increased 3.6% (95%CI, 3.1%-4.1%) 1 year after the start of the incentive metric, 7.5% (95%CI, 6.8%-8.2%) at the end of 2 years, and 11.5% (95%CI, 10.5%-12.4%) at the end of 3 years. Before the incentive, contraceptive use rates among Medicaid enrollees 18-24 years of age were decreasing. When results were stratified by age, increased use rates were found in all groups.

7. NewYork-Presbyterian Hospital (NYP)/Columbia University Irving Medical Center Ambulatory Care Network (ACN)

NewYork-Presbyterian Hospital (NYP)/Columbia University Irving Medical Center Ambulatory Care Network (ACN) analysis included 31,084 female clients ages 15-44 who in calendar year 2018 received services from 31 NYP ACN facilities. Approximately 42.7% of clients ages 15-44 received a most or moderately effective method of contraception, and the measure scores varied across 31 facilities. For more details, see the Testing Attachment.

Number of measured entities: 31 facilities

Number of female patients ages 15-44: 31,084

Dates included: January 1 through December 31, 2018

Mean performance score: 32.80

Standard deviation: 13.82

Range: 3.7 – 59.2

Percentiles:

25th: 21.3

50th: 32.5

75th: 45.6

8. Planned Parenthood Federation of America (PPFA)

The PPFA final dataset analyzed included 123,978 female patients aged 15-44 years, who received services from 2 PPFA affiliates between January 1 and December 31, 2019. The measures were evaluated using all claims data among the eligible population, which included de-identified patient encounters, and identifiers for providers and health centers within affiliates. The results showed that 61.2% of clients ages 15-44 were provided a most or moderately effective method of contraception; variation existed across 56 facilities. For more details, see attached Testing Attachment.

Number of measured entities: 56 facilities

Dates included: January 1 through December 31, 2019

Mean performance score: 59.22

Standard deviation: 16.02

Range (minimum – maximum): 0.00 – 81.02

Percentiles:

25th: 59.0

50th: 63.0

75th: 66.0

Scores by decile

0 – 10: 3

11 – 20: 0

21 – 30: 2

31 – 40: 0

41 – 50: 1

51 – 60: 13

61 – 70: 32

71 – 80: 4

81 – 90: 1

91 – 100: 0

9. The Title X Family Planning Program

Enacted in 1970, the Title X Family Planning program is the only federal grant program dedicated solely to providing low-income individuals with comprehensive family planning and related preventive health services. The U.S. Department of Health and Human Services (HHS) Office of Population Affairs (OPA) oversees the Title X program. Calculated from the Title X Family Planning Annual Report (FPAR), the application includes Title X measure scores to demonstrate that even in a program committed to the provision of family planning services, considerable room for improvement exists in its delivery of contraceptive services. The FPAR data has several advantages over claims data, in that it documents sterilization or LARC insertion in a year preceding the measurement year, and whether the client was seeking pregnancy. The 2019 results showed that overall, 65.7% of clients ages 15-19 and 59.5% of clients ages 20-44 were provided a most or moderately effective method of contraception; variation by grantee existed (e.g., from 0 to 89.4% for adolescent clients, and from 0 to 82.9 % among adult clients). See 2018 and 2019 FPAR results below. For more details, see the attached appendix.

Number of measured entities: 99 grantees

FPAR 2018

Ages 15-19

Mean performance score: 67.9

Standard deviation: 0.17

Range (minimum – maximum): 0.00 – 92.3

Percentiles:

25th: 62.1

50th: 74.0

75th: 79.0

Scores by decile

0 – 10: 2

11 – 20: 0

21 – 30: 1

31 – 40: 6

41 – 50: 2

51 – 60: 10

61 – 70: 20

71 – 80: 32

81 – 90: 20

91 – 100: 1

Ages 20-44

Mean performance score: 61.3

Standard deviation: 0.15

Range (minimum – maximum): 0.00 – 88.2

Percentiles:

25th: 53.8

50th: 63.8

75th: 70.1

Scores by decile

0 – 10: 1

11 – 20: 1

21 – 30: 1

31 – 40: 6

41 – 50: 8

51 – 60: 20

61 – 70: 32

71 – 80: 20

81 – 90: 6

91 – 100: 0

Number of measured entities: 100 grantees

FPAR 2019

Ages 15-19

Mean performance score: 65.7

Standard deviation: 0.19

Range (minimum – maximum): 0.00 – 89.4

Percentiles:

25th: 58.7

50th: 72.2

75th: 78.7

Scores by decile

0 – 10: 2

11 – 20: 2

21 – 30: 1

31 – 40: 9

41 – 50: 2

51 – 60: 13

61 – 70: 16

71 – 80: 33

81 – 90: 21

91 – 100: 0

Ages 20-44

Mean performance score: 59.5

Standard deviation: 0.16

Range (minimum – maximum): 0.00 – 82.9

Percentiles:

25th: 55.4

50th: 63.2

75th: 70.3

Scores by decile

0 – 10: 2

11 – 20: 1

21 – 30: 5

31 – 40: 4

41 – 50: 6

51 – 60: 24

61 – 70: 31

71 – 80: 23

81 – 90: 3

91 – 100: 0

From 2016-2019, the percentage of all Title X family planning users provided a most or moderately effective method of contraception by year remained quite stable, with a very slight decrease in the percentage of women using most or moderately effective methods in 2018 and 2019 [2-5].

2016: 62%

2017: 62%

2018: 60%

2019: 59%

References

[1] Rodriguez MI, Meath T, Huang J, Darney BG, McConnell KJ. Association of Implementing an Incentive Metric in the Oregon Medicaid Program With Effective Contraceptive Use. JAMA Netw Open. Aug 3 2020;3(8):e2012540.

[2] Fowler, C. I., Gable, J., Lasater, B., & Asman, K. (2020, September). Family Planning Annual Report: 2019 National Summary. Washington, DC: Office of Population Affairs, Office of the Assistant Secretary for Health, Department of Health and Human Services. https://opa.hhs.gov/sites/default/files/2020-09/title-x-fpar-2019-national-summary.pdf

[3] Fowler, C. I., Gable, J., Wang, J., Lasater, B., & Wilson, E. (2019, August). Family Planning Annual Report: 2018 national summary. Research Triangle Park, NC: RTI International. https://opa.hhs.gov/sites/default/files/2020-07/title-x-fpar-2018-national-summary.pdf

[4] Fowler, C. I., Gable, J., Wang, J., & Lasater, B. (2018, August). Family Planning Annual Report: 2017 national summary. Research Triangle Park, NC: RTI International. https://opa.hhs.gov/sites/default/files/2020-07/title-x-fpar-2017-national-summary.pdf

[5] Fowler, C. I., Gable, J., Wang, J., & Lasater, B. (2017, August). Family Planning Annual Report: 2016 national summary. Research Triangle Park, NC: RTI International. https://opa.hhs.gov/sites/default/files/2020-07/title-x-fpar-2016-national.pdf

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.

A special analysis of data from the National Survey of Family Growth (NSFG), 2015-2017, was conducted to examine contraceptive use patterns among women who were at risk of unintended pregnancy because they had ever had sex, were fecund, and were neither pregnant nor seeking pregnancy. The analysis showed that 51.7% of adolescents and 60.8% of adult women used a most or moderately effective method (CDC/NCHS, unpublished data), which indicates there might be room for improvement (e.g., 15-20 percentage points). We have noted and published on our website (https://opa.hhs.gov/evaluation-research/title-x-services-research/contraceptive-care-measures/most-or-moderately) that: "No specific benchmark has been set for this measure, but the Office of Population Affairs (OPA) does not expect it to reach 100%, as some women will make informed decisions to choose methods in the lower tier of efficacy even when offered the full range of methods and all logistical or financial barriers to access are removed."

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.

In addition to calculating NQF #2903 by age group for reliability and validity testing, we examined two datasets for measure scores by race and ethnicity: Planned Parenthood Federation of America (PPFA) and Washington State Health Care Authority (WA HCA).

The PPFA final dataset analyzed included 123,978 female patients aged 15-44 years, who received services from two PPFA affiliates between January 1 and December 31, 2019. The results showed that the percentage of women ages 15-44 that were provided most and moderately effective methods differed by race/ethnicity reported:

African American: 53.50

Alaskan Native: 64.87

Asian: 68.23

Hispanic: 66.27

Multi-racial: 64.64

Native American: 59.83

Pacific Islander: 65.18

White: 66.53

Other race: 58.49

For 2014-2018, WA HCA reported NQF #2903 measure scores for female clients ages 15-44 by age group and race/ethnicity (https://www.hca.wa.gov/assets/program/ccw-contraceptive-care.pdf). The percentages of women that were provided most and moderately effective methods by race/ethnicity remained stable over these five years.

In 2018, the measure scores for ages 15-20 differed by race/ethnicity reported (note that race/ethnicity categories other than "Hispanic" report ethnicity as "Not Hispanic" or "Unknown"):

Hispanic: 24.4

White: 37.2

Asian: 19.4

Black: 24.5

American Indian/Alaska Native: 33.7

Hawaiian/Pacific Islander: 18.9

More than One Race: 34.9

Other/Unknown: 23.7

The 2018 measure scores for ages 21-44 also varied by race/ethnicity reported:

Hispanic: 33.1

White: 27.0

Asian: 26.0

Black: 26.1

American Indian/Alaska Native: 24.6

Hawaiian/Pacific Islander: 23.6

More than One Race: 29.9

Other/Unknown: 26.9

Based on these measure scores, opportunities for improvement exist to ensure that all race/ethnicity groups have equal access to the full range of contraceptive methods and receive patient-centered contraceptive care. These differences by socio-demographic characteristics could be explained in part by modifiable clinical and programmatic considerations rather than varying biological responses to contraception. Although providers may see some local variations by socio-demographic characteristics, we believe that these differences will be reduced if contraceptive services are offered in a client-centered manner, as defined by CDC and OPA's recommendations, Providing Quality Family Planning Services (https://www.cdc.gov/mmwr/preview/mmwrhtml/rr6304a1.htm).

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

To further investigate differences in most and moderately effective contraceptive use, a special analysis of data from the National Survey of Family Growth (NSFG) 2015-2017 was conducted (CDC/NCHS unpublished data). This analysis suggests that there are statistically significant differences by age group (for ages 20-29 compared to ages 30-44) and among women who have never been married (compared to women of other marital status). However, no significant differences occur between race/ethnicity, most categories of marital status, and poverty level. For more details, please see the Testing Attachment.

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):*

Perinatal Health, Perinatal Health : Newborn Care

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

Primary Prevention

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

Children, Women

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.)*

https://opa.hhs.gov/evaluation-research/title-x-services-research/contraceptive-care-measures/most-or-moderately

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 **:** NQF_2903_Codes_2021-637453719019907247.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.

Yes

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.

The measure specification has been changed to no longer include diaphragm as a moderately effective contraceptive method. A woman will no longer be included in the numerator if she only has codes indicating use of a diaphragm in the measurement year. This revision brings the measure up-to-date with the current edition of the clinical reference Contraceptive Technology (http://www.contraceptivetechnology.org/the-book/take-a-peek/contraceptive-efficacy), which classifies the diaphragm as a less effective method of contraception due to higher typical use failure rates. Many public and reproductive health organizations cite and use the typical use failure estimates from Contraceptive Technology in their educational materials for clients and providers. These updated typical use failure rates have also been reported by the Centers for Disease Control and Prevention (CDC) and World Health Organization (WHO). Furthermore, removal of diaphragm from the measure numerator should not greatly impact measure scores because only a small proportion of women utilize a diaphragm as contraception. For more information, see the Release Notes at the end of the Intent to Submit Form.

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).*

Women ages 15-44 at risk of unintended pregnancy who are provided a most (sterilization, intrauterine device, implant) or moderately (injectable, pill, patch, ring) effective method of contraception.

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 target population is eligible women ages 15-44 who are provided a most or moderately effective method of contraception. To identify the numerator, follow these steps:

Step 1 Define the numerator by identifying women who were provided a most (sterilization, IUD, implant) or moderately (injectable, pill, patch, or ring) effective method of contraception in the measurement year. To do this, use the codes in Table CCW-E.

Step 2 Calculate the rates by dividing the number of women who were provided a most or moderately effective method of contraception by the number of women in the denominator. Calculate the rates separately for adolescents and adults.

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

Women ages 15-44 who are at risk of unintended pregnancy.

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

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

The target population is women of reproductive age (i.e., ages 15-44 years). In a Medicaid population, this includes:

- Women in the general Medicaid program who were continuously enrolled during the measurement year, i.e., had no more than one gap in enrollment of up to 45 days. To determine continuous enrollment for a Medicaid enrollee for whom enrollment is verified monthly, the enrollee may not have more than a 1-month gap in coverage (i.e., an enrollee whose coverage lapses for 2 months is not considered continuously enrolled)
- All women participating in a state-sponsored family planning-specific Section 1115 waiver or in a family–planning specific state plan amendment (SPA) program, even if they were not continuously enrolled. This is because the primary intent of these waiver and/or SPA programs is to provide family planning services, including contraception.

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

The following categories of women are excluded from the denominator: (1) those who are infecund for non-contraceptive reasons; (2) those who had a live birth in the last 2 months of the measurement year; or (3) those who were still pregnant or their pregnancy outcome was unknown at the end of the measurement year.

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.)*

Follow the steps below to identify the denominator. The tables that are referenced are found in the attached Excel file (NQF_2903_Codes_2021.xlsx).

Step 1 Identify and exclude women who were infecund due to non-contraceptive reasons such as natural menopause or oophorectomy. To do this, use the codes listed in Table CCW-A.

Step 2 Identify women who were pregnant at any point in the measurement year by using the codes listed in Table CCW-B. We selected this list of codes by reviewing the following documents:

- CMS & NCHS (2020). ICD-10-CM Official Guidelines for Coding and Reporting FY 2021. Available online at: http://www.cdc.gov/nchs/icd/icd10cm.htm
- CMS & NCHS (2020). ICD-10-PCS Official Guidelines for Coding and Reporting FY2020. Available online at: https://www.cms.gov/Medicare/Coding/ICD10/2020-ICD-10-PCS

Step 3 Among women who were pregnant at any point in the measurement year, exclude those who:

- Had a live birth in the last 2 months of the measurement year because there may not have been an opportunity to provide them with contraception. A two-month period was selected because the American College of Obstetricians and Gynecologists (ACOG) recommends having a postpartum visit by 6 weeks, and an additional 2 weeks was added to allow for reasonable delays in attending the postpartum visit. To identify live births, use the codes listed in Table CCW-D. This table includes codes from the HEDIS measure of Prenatal and Postpartum Care, and ICD-10-CM codes for live births were added.
- Were still pregnant at the end of the measurement year because they did not have a pregnancy outcome code indicating a non-live birth (Table CCW-C) or a live birth (Table CCW-D). Codes for non-live births were also drawn from the HEDIS measure of Prenatal and Postpartum Care, and procedure codes (CPT, ICD-10-PCS codes) were added.

Once the exclusions are applied, the denominator includes women who:

- Were not pregnant at any point in the measurement year,
- Were pregnant during the measurement year but whose pregnancy ended in the first 10 months of the measurement year since there was adequate time to provide contraception in the postpartum period.
- Were pregnant during the measurement year but whose pregnancy ended in an ectopic pregnancy, stillbirth, miscarriage, or induced abortion.

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.)*

The primary stratification variable is age, so that adolescents can be examined separately from adult women for the purposes of quality improvement. Though their current clinical guidelines report that most and moderately effective contraceptive methods are safe and recommended for teen and nulliparous populations who wish to use them, the American Academy of Pediatrics (AAP), ACOG, the Centers for Disease Control and Prevention (CDC), and Office of Population Affairs (OPA) note that it can still be difficult for these populations to access these highly effective contraceptive methods. We utilize age groups that are consistent with Center for Medicaid and CHIP Services (CMCS) reporting requirements; adolescents are defined as 15-20 years and adults are 21-44 years of age.

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 Identify all women ages 15-44 who were enrolled in the health plan or program. In the case of general Medicaid, include women who were continuously enrolled (i.e., had no more than one gap in enrollment of up to 45 days). In the case of women enrolled in a family planning-specific expansion program (1115 waiver or state plan amendment), include all women even if they do not meet the continuous enrollment criteria because the reason for their visit is related to pregnancy prevention.

Step 2 Define the denominator by excluding women who: (a) are infecund for non-contraceptive reasons; (b) had a live birth in the last 2 months of the measurement year; or (c) were still pregnant or their pregnancy outcome was unknown at the end of the measurement year. Once exclusions are applied, the following groups of women will be included in the denominator: (a) those who were not pregnant at any point in the measurement year; (b) those who had a live birth in the first 10 months of the measurement year; and (c) those who had a known miscarriage, stillbirth, ectopic pregnancy, or induced abortion during the measurement year.

Step 3 Define the numerator by using claims codes to identify women in the denominator who were provided or continued use of one of the following methods of contraception in the measurement year: sterilization, IUD, implant, contraceptive injection, pills, patch, or ring.

Step 4 Calculate the rates by dividing the number who were provided or continued use of a most or moderately effective method of contraception by the number of women in the denominator. Calculate the rates for all women ages 15-44 and separately for adolescents and adults.

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.

The measure is based on data about all clients seen, not a sample.

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

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

Not applicable.

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

If other, please describe in S.18.

Claims

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

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

Administrative claims data are used to calculate the measure. The data request should include an eligibility file, all paid, suspended, pending, and denied claims with diagnosis codes (ICD-10-CM), procedure codes (HCPCS, CPT, ICD-10-PCS), and medication codes (NDC).

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)*

Available in attached appendix at A.1

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

Clinician : Group/Practice, Facility, Health Plan, Population : Regional and State

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

Other

If other: Primary care and reproductive health settings.

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.)*

Not applicable.

Validity – See attached Measure Testing Submission Form

MostMod_2903_nqf_testing_attachment_2021-4-27.docx

2.1 For maintenance of endorsement

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

Yes

2.2 For maintenance of endorsement

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

Yes

2.3 For maintenance of endorsement

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

No - This measure is not risk-adjusted

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

Measure Number (*if previously endorsed*)**:** 2903

Measure Title: Contraceptive Care – Most & Moderately Effective Methods

Date of Submission: **1/5/2021**

Type of Measure:

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

*cell intentionally left blank

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

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

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

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

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*).

2021 Submission

Claims data from seven organizations were used for testing:

(1) **The Planned Parenthood Federation of America (PPFA).** In 2019, PPFA comprised 49 independently incorporated affiliates, operating approximately 600 facilities in the United States, and providing reproductive health care to nearly 2.4 million patients. De-identified, encounter-level data are captured in a quality information warehouse for a subset of affiliates. The final dataset analyzed included female patients aged 15-44 years, who received services from 2 PPFA affiliates between January 1 and December 31, 2019. The measures were evaluated using all claims data among the eligible population, which included de-identified patient encounters, and identifiers for providers and health centers within affiliates. Affiliates vary in size and can cover geographic service areas that range from several counties within a single state, to an entire state population, up to multiple states. Among the 2 affiliates included in our dataset, there were 64 facilities and 188 unique providers nested among the facilities. One affiliate represents multiple less densely populated states, while the other includes several counties in one state. For the purposes of this application, OPA suggests that each affiliate be considered a proxy for a U.S. state. We utilized the PPFA data for reliability and validity testing.

(2) **The Iowa Medicaid Enterprise (IME)**. The IME dataset comprised all female Medicaid clients aged 15-44 years who resided in 99 counties and participated in either the general Medicaid program or the state-funded Family Planning Program (FPP). IME provides contraceptive services to women through these two programs. To be eligible for FPP services, the following guidelines apply: the individual is a man or woman between the ages of 12 and 54 years; family income is at or below 300 percent of the

federal poverty level; and women whose pregnancy and delivery was covered by Medicaid will have family planning services covered for an additional 12 months without having their eligibility re-determined. During fiscal year 2019, Medicaid services in Iowa were provided primarily through two managed care organizations (MCOs), although a small percentage of clients (approximately 7%) were provided care on a fee-for-service basis. We utilized the IME data for reliability and validity testing.

(3) **Iowa Department of Public Health (IDPH) Title X Grantee**. The IDPH dataset included a random sample of female clients ages 15-44 who visited six Title X sites in calendar year 2019. As a Title X grant recipient, IDPH provides funding to seven local agencies which cover 45 counties and offer a broad range of family planning methods and related preventive health services. Persons from low-income families are given priority. Costs for service at IDPH's family planning clinics are based on ability to pay and are often less than at other health centers. IDPH Title X services are free for people enrolled in Medicaid and those whose income is below 100% of FPL. We utilized the IDPH data for data element level validity testing.

(4) **NewYork-Presbyterian Hospital (NYP)/Columbia University Irving Medical Center.** In 2018, NYP Ambulatory Care Network (ACN) consisted of 14 primary care sites, 7 school-based facilities, 13 mental health school-based programs, and over 60 specialty practices. NYP ACN totaled 3,428,630 outpatient visits, 155,399 ambulatory surgeries, and 693,454 emergency department visits (including admissions) during 2018. The NYP dataset is comprised of female clients aged 15-44 years who in 2018 received services from 8 NYP outpatient locations. Within these 8 ACN locations are 31 facilities. We utilized the NYP data for reliability testing.

(5) **Washington State Health Care Authority (WA HCA)**. In 2019, the WA HCA dataset contained all female Medicaid clients aged 15-44 years who resided in 39 counties and participated in 5 health plans. WA HCA provided contraceptive services to these women via the general Medicaid program or the state's family planning waiver programs, Family Planning Only and Family Planning Only – Pregnancy Related. Formerly known as Take Charge, Family Planning Only is a 1115 demonstration waiver program that serves low-income (up to 260% of FPL) uninsured clients seeking to prevent unintended pregnancy, and teens and domestic violence victims who need confidential family planning services. The Family Planning Only – Pregnancy Related program (previously known as the Family Planning Only extension) provides services to recently pregnant women who lose Medicaid coverage 60 days post-pregnancy. During fiscal year 2019, Medicaid services in Washington were provided primarily through 5 MCOs; approximately 85% of Washington's Medicaid clients were enrolled in managed care. We utilized the WA HCA data for reliability testing.

(6) Massachusetts **MassHealth (MA)**. In 2019, the MA dataset contained all female Medicaid clients aged 15-44 years who resided in 14 counties and participated in 21 health plans. Sixteen of these health plans were accountable care organizations (ACO). An ACO is a group of doctors, hospitals, and other health care providers that work together with the goals of delivering better care to members, improving the population's health, and controlling costs. ACOs are accountable both for the health of their members and for the cost of the care their members receive. MA provided contraceptive services to these women via the general Medicaid program. Approximately 70% of Massachusetts Medicaid clients were enrolled in managed care. We utilized the MA data for reliability testing.

(7) **Louisiana Medicaid (LA Medicaid)**. In 2019, the LA Medicaid dataset contained all female Medicaid clients aged 15-44 years who resided in 64 parishes and participated in 5 health plans. LA Medicaid provided contraceptive services to these women via the general Medicaid program or its family planning state plan amendment, Take Charge Plus (which is a different program than WA HCA's family planning waiver program). Take Charge Plus provides family planning and/or family planning-related services to low-income women or men (138% of FPL). In 2019, Medicaid services in Louisiana (excluding Medicaid-Medicare dual-eligibles) were provided primarily by 5 managed care plans, which are administered by the state's Healthy Louisiana program. Approximately 15% of the Medicaid population not dually eligible was continuously enrolled in traditional fee-for-service Medicaid. We utilized the LA data for reliability testing.

2016 Submission

Claims data from three programs were used for testing:

(1) **The Planned Parenthood Federation of America** (PPFA). In 2014, PPFA comprised 66 independently incorporated affiliates, operating approximately 700 health centers in the United States, and providing reproductive health care to nearly 2.7 million patients. De-identified, encounter-level data are captured in a quality information warehouse for a subset of affiliates. The final dataset analyzed included 838,872 female patients aged 15-44 years, who received services from 25 PPFA affiliates between January 1 and December 31, 2014. The measures were evaluated using all claims data among the eligible population, which included de-identified patient encounters, and identifiers for billing providers and health centers within affiliates. Affiliates cover geographic service areas that range from several counties within a state, a state population, and multiple states. Among the 25 affiliates included in our dataset, there were 363 health centers, and 4,467 unique billing providers nested among the health centers. These data cover diverse geographic regions and extremely large member populations, and thus may be considered reasonably representative of the U.S. population of women of reproductive age. Hence, OPA suggests that the affiliate be considered a reasonable proxy for a U.S. state, for purposes of this application.

(2) **The Iowa Medicaid Program** (IME). The IME dataset comprised all female Medicaid clients aged 15-44 years who resided in 6 public health regions, participated in either fee-for-service care or in two health plans, and participated in either the general Medicaid program or the state's family planning waiver program. Iowa's Medicaid Enterprise (IME) provides contraceptive services to women through its general Medicaid program and its family planning waiver program (IFPN). Services are available to Iowa residents who are US citizens or qualified immigrants. To be eligible for IFPN services, the following guidelines apply: an individual does not have insurance or your insurance does not cover family planning services; the individual is a man or woman between the ages of 12 and 54; family income is at or below 300 percent of the federal poverty level; and women whose pregnancy and delivery was covered by Medicaid will have family planning services covered. In 2013, Medicaid services in Iowa were provided primarily on a fee-for-service basis, although a small percentage of clients (approximately 2%) were provided care through one of two managed care organizations (MCO).

(3) **The Wisconsin Medicaid Program** (WMP). The WMP dataset is comprised of all female Medicaid clients aged 15-44 years who in 2014 resided in Wisconsin, had a paid Medicaid claim, and participated in either the general Medicaid program or the state's Family Planning Only Services (FPOS) program. The

WMP provides contraceptive services to women through its general Medicaid program (BadgerCare Plus) and FPOS. FPOS members receive services on a fee for service basis only. Services are available to Wisconsin residents who are US citizens or qualified immigrants meeting income eligibility criteria (e.g., a child <18 years with household income at or below 300% FPL; an adult with income at or below 100% FPL). To be eligible for FPOS, individuals must not be covered by Medicaid for the Elderly, Blind, or Disabled or BadgerCare Plus and must be at or below 300% FPL. In December 2014, 65% of Wisconsin Medicaid members were enrolled in one of 18 health maintenance organizations (HMO).

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

2021 Submission

Data from PPFA, WA HCA, MA, LA Medicaid, and IDPH covered the period January 1, 2019 – December 31, 2019.

Data from IME and NYP covered the period January 1, 2018 – December 31, 2018.

2016 Submission

January 1 2013 – December 31 2014

Data from PPFA covered the period January 1 2014 – December 31 2014.

Data from IME covered the period January 1, 2013 – December 31, 2013.

Data from Wisconsin Medicaid covered the period January 1, 2014 – December 31, 2014.

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: Population/state equivalent, public health region, benefit type	☒ other: Public health region

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)

2021 Submission

Reliability

The measure was tested at several levels, as shown in the table below.

Level	Number of measured entities	Data Source
Facility	56	PPFA
Facility	31	NYP
Public health region	6	IME
Group billing provider	3,081	IME
Health plan	5	WA HCA
Health plan	21	MA
Health plan	5	LA Medicaid

Validity

Score Level Validity

The measure was tested at the facility and group billing provider levels as the reliability table shown above.

Data Element Validity

Six Iowa Department of Public Health Title X Family Planning Clinics provided data and the analysis was conducted using aggregated numbers across all 6 clinics.

2016 Submission

Reliability

The measure was tested at several levels, as shown in the table below.

Level	Number of measured entities	Data Source
Affiliate	25	PPFA
Health center	363	PPFA
Benefit type (general Medicaid vs FP waiver)	2	IME
Public health region	6	IME

Level	Number of measured entities	Data Source
Health plan (Medicaid health maintenance organization)	17	WMP

Validity

A panel of experts assessed the measure's face validity.

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*)

2021 Submission

Level of analysis	Number of patients: 15 – 20 years	Number of patients: 21-44 years	Number of patients: 15-44 years
Facility, n = 56 (PPFA)	*	*	*
TOTAL	28,454	95,524	123,978
Range	0 – 1,267	1 – 4,240	1 – 5,030
Facility, n =31 (NYP)	*	*	*
TOTAL	5,705	25,379	31,084
Range	2 – 1,568	0 – 8,894	87 – 10,462
Public Health Region, n = 6 (IME)	*	*	*
PHR 1	8,365	25,070	33,435
PHR 2	2,247	6,392	8,639
PHR 3	3,183	8,615	12,098
PHR 4	2,824	2,252	3,455
PHR 5	3,609	11,346	14,955
PHR 6	8,409	28,341	36,750
TOTAL	**28,637**	**88,255**	**116,892**
Group Billing Provider, n = 3,081 (IME)	*	*	*
TOTAL	24,162	75,627	99,789
Range	0 – 1,433	0 – 4,804	1 – 6,237

Level of analysis	Number of patients: 15 – 20 years	Number of patients: 21-44 years	Number of patients: 15-44 years
Title X grantee, n = 6 (IDPH)	*	*	*
Clinic 1	20	49	69
Clinic 2	24	52	76
Clinic 3	17	55	72
Clinic 4	21	47	68
Clinic 5	12	59	71
Clinic 6	20	47	67
TOTAL	**114**	**309**	**423**
Health Plan, n = 5 (WA HCA)	*	*	*
MCO 1	4,031	15,357	19,388
MCO 2	9,684	20,378	30,062
MCO 3	7,731	15,127	22,858
MCO 4	31,628	73,240	104,868
MCO 5	4,281	15,111	19,392
TOTAL	**57,355**	**139,213**	**196,568**
Health Plan, n = 21 (MA)	*	*	*
TOTAL	50,934	146,595	197,529
Range	0 – 8,036	351 – 17,779	351 – 22,499
Health Plan, n = 5 (LA Medicaid)	*	*	*
MCO 1	3,004	15,174	18,178
MCO 2	10,115	27,867	37,982
MCO 3	12,636	37,620	50,256
MCO 4	29,880	61,423	91,303
MCO 5	22,699	58,682	81,381
TOTAL	**78,334**	**200,766**	**279,100**

*cell intentionally left blank

2016 Submission

Level of analysis	Number of patients: 15 - 20 years	Number of patients: 21 -44 years	Number of patients: 15 - 44 years
Affiliate (PPFA), n=25	*	*	*
TOTAL	203,970	634,902	838,872
Range	294 - 42,698	1265 – 131,187	1701 – 173,885
Health centers within affiliate (PPFA), n=363	*	*	*
TOTAL	203,970	634,902	838,872
Range	Aug-84	31 – 11,391	48 – 13,335
Type of benefit (IME)	*	*	*
General Medicaid	5,254	9,483	14,737
Family planning waiver	6,445	23,568	30,013
TOTAL	**11,699**	**33,051**	**44,750**
Public health region (IME)	*	*	*
Region 1	3,460	9,588	13,048
Region 2	1,154	2,906	4,060
Region 3	1,176	3,175	4,351
Region 4	1,087	2,887	3,974
Region 5	1,701	4,359	6,060
Region 6	3,121	10,136	13,257
TOTAL	**11,699**	**33,051**	**44,750**
Health plan (WMP)	*	*	*
HMO 1	4,832	14,043	18,875
HMO 2	1,838	5,688	7,526
HMO 3	920	2,862	3,782
HMO 4	1,795	5,681	7,476
HMO 5	1,231	3,936	5,167
HMO 6	219	725	944
HMO 7	558	1,608	2,166
HMO 8	352	1,096	1,448
HMO 9	1,623	6,164	7,787

Level of analysis	Number of patients: 15 - 20 years	Number of patients: 21 -44 years	Number of patients: 15 - 44 years
HMO 10	618	1,683	2,301
HMO 11	4,898	15,166	20,064
HMO 12	1,239	4,290	5,529
HMO 13	2,69	853	1,122
HMO 14	2,149	5,596	7,745
HMO 15	56	240	296
HMO 16	5,114	18,875	23,989
HMO 17	559	1,533	2,092
TOTAL	**28,270**	**90,039**	**118,309**

*cell intentionally left blank

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.

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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.

2021 Submission

Reliability and validity of the measure were assessed after stratifying by age group (e.g., adolescent compared to adult women of reproductive age). Given different care delivery models by age group, calculating the measure in this way is important to develop successful quality improvement initiatives and public health interventions. We utilized the age group categories developed by the Center for Medicaid and CHIP Services (CMCS). CMCS defines adolescents as individuals aged 15 through 20 years (15-20), while adults of reproductive age are individuals aged 21 through 44 years (21-44).

2016 Submission

We assessed reliability of the measures after stratifying by age, i.e., adolescent versus adult. Teen pregnancy is worthy of a separate focus because of the large potential negative impact on the life of the

teen and her child(ren), and the existence of unique programs and contraceptive counseling approaches tailored to this population. To define age groups, we used the categories developed by the Center for Medicaid and CHIP Services (CMCS), i.e., individuals aged 15 through 20 years (15-20) were defined as adolescents, and individuals aged 21 through 44 years (21-44) were defined as adults.

2a2. RELIABILITY TESTING

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

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

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

2021 Submission

Several methods have been suggested to assess the reliability of provider-level performance measures (Adams, 2010; Scholle et al, 2008; Fung et al, 2010). These methods may focus on different facets of reliability such as consistency across time, consistency across raters or units, or variability at different levels of aggregation. The NQF has suggested a *signal-to-noise* approach as one way to evaluate measure reliability (Adams, 2009). For this application, reliability was estimated from a Beta-binomial model using parametric empirical Bayes methods. Two distributional shape parameters (alpha and beta) were estimated from the observed quality scores, and reliability was then calculated as a function of alpha, beta, and total patient count for each unit of analysis. Overall reliability in this context represents the ability of the proposed measure to confidently distinguish the performance of one entity (e.g., facility) from another. A detailed description of this method is demonstrated in the Appendix, where we lay out the formulation of the method and describe how it improves upon the Beta-binomial approach applied in previous studies (Adams, 2009; Adams and Paddock, 2017; Blair et. al., 2015; Kazis et. al., 2017; Staggs and Cramer, 2016).

Measure developers frequently recommended setting a minimum patient size for performance measurement when estimating at the facility or provider level because patient size has a large impact on reliability (HEDIS, 2007; Safran, 2007). In this analysis, we tested reliability using 75 as a cutoff of total patients served at each unit of analysis to show how such threshold impacts reliability.

Structure of the Data

PPFA dataset. Two PPFA affiliates included in our dataset contain a total of 64 facilities. Eight of the 64 facilities were follow-up call centers or labs that did not serve any eligible women, resulting in 56 client

facilities included in the analysis. Reliability testing could only be performed at the facility level due to the limited number of affiliate data partners.

Iowa Medicaid Enterprise dataset (IME). For Iowa Medicaid, we performed reliability testing at 2 different levels: public health region (n=6) and group billing provider (n=3081). Iowa Medicaid data does not contain facility information. We used billing providers who registered as "organizations" to represent group practices. The additional use of a cutoff to exclude group billing providers who served less than 75 patients during the measurement year further ensures that we are only examining reliability among large group practices, rather than small rural practices that may only have one doctor, even if it is registered as a group practice.

NewYork-Presbyterian (NYP) Hospital dataset. The NYP network included 31 facilities. Reliability testing was performed at the facility level.

Washington State Health Care Authority dataset (WA HCA). The Washington Medicaid program included 5 health plans. Reliability testing was performed at the health plan level.

Massachusetts MassHealth dataset (MA). The Massachusetts Medicaid program included 21 health plans. Reliability testing was performed at the health plan level.

Louisiana Medicaid dataset (LA Medicaid). The Louisiana Medicaid program included 5 health plans. Reliability testing was performed at the health plan level.

2016 Submission

Several methods have been suggested to assess the reliability of provider-level performance measures (Adams, 2010; Scholle et al, 2008; Fung et al, 2010). These methods may focus on different facets of reliability such as consistency across time, consistency across raters or units, or variability at different levels of aggregation. The NQF has suggested a ***signal-to-noise*** approach as one way to evaluate measure reliability. According to Adams (2009), reliability can be assessed by the proportion of variance in a performance measure due to systemic differences across measured units (signal) in relation to random error (noise) within units.

When analytic units fall into a natural hierarchy (e.g. clients nested within health centers nested within health plan organizations), one can estimate multilevel variance components using hierarchical generalized linear modeling (HGLM) (Raudenbush and Bryk, 2002; Woltman et al, 2012). In this approach the within-provider regression coefficients are allowed to vary across providers as random effects. The covariance parameter for the random effect estimates the true between-provider variance after accounting for within-provider variance. HGLM methods are robust and well-developed for continuous outcomes, and have more recently been applied to binary outcomes (Ridout, 1999; Molenberghs et al, 2007).

In the present analyses, multi-level mixed models were fit to each dataset using a hierarchical SAS 9.3 GLIMMIX procedure with a log link function. Parameters were estimated by pseudo-maximum-likelihood using the Laplace method (Ene et al, 2012). Modeling proceeded in a top-down manner

starting from the largest unit of aggregation; the variance component (random coefficient) was always estimated for the top level.

Reliability was then calculated as a function of the intraclass correlation (ICC) and the median number cases per unit, using the Spearman-Brown prophecy (Eijkenaar et al, 2013). ICCs are derived using the estimated variance component for the level of interest divided by the total variance (Wu et al, 2012; He et al, 2014). ICCs conceptually represent the proportion of total variation accounted for by the between-provider level, and thus follows the signal-to-noise framework suggested by NQF.

The HGLM method of estimation assumes a normally distributed error component; some authors have noted that ICCs on the logit scale can be inflated under certain circumstances when population rates are near the extremes (Wu et al, 2012). To provide more conservative estimation, medians were used in the Spearman-Brown reliability formula; the use of means would tend to bias estimates upward due to one or two atypically large provider units.

Structure of the Data

PPFA dataset. PPFA affiliates offer services within health centers. Inside each health center a group of billing providers offer care to clients. Modeling began at the topmost affiliate level (n=25), where all clients were aggregated within affiliate for the calculation of rate of most/moderately effective contraceptive use. The next level of analysis was performed within each of the 25 affiliates to examine health center rates (n=363 across all affiliates). This provided a basic 2-level structure of clients aggregated within each hierarchical unit. The top-down modeling approach enabled us to ignore small sample size problems and attribution error among individual billing providers; it also allowed us to explore the lowest level of 'granularity' for distinguishing performance among health centers of smaller size.

Iowa Medicaid Enterprise dataset. For IME data, modeling similarly proceeded from the top down starting with public health region (n=6). Unlike the PPFA data, IME data could not be examined by health facility. Instead the analysts were interested in reporting on public health region and benefit type (family planning waiver or general Medicaid benefit). Since the benefit type categories exist across regions, there is no nesting of units. Therefore, in Iowa the six regions were simply crossed with the type of benefit (n=12). Both of these crossed analyses were thought to provide useful and potentially actionable information about the interplay of regional and administrative influences on service delivery.

Wisconsin Medicaid dataset. For WMP data, modeling similarly proceeded from the top down starting with health maintenance organization (data from 17 of 18 HMOs was available).

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

2021 Submission

The table below shows summary results of the reliability analyses at four levels (facility, public health region, group billing provider, and health plan), stratified by three age categories (i.e., 15-20, 21-44, and

15-44 years). More detailed information including reliability estimates for each unit at each level (except group billing providers) can be found in Tables 1-6 (appended at the end of the form).

Beta-binomial reliability estimates by age group

Level	Age group	Results: Median N (all units)	Results: Reliability (all units)	Results: Median N (unit size ≥ 75)	Results: Reliability (unit size ≥ 75)
Facility (PPFA) 1	15-44	2,915	**.972**	2,929	**.989**
Facility (PPFA) 2	21-44	2,180	.966	2,201	.983
Facility (PPFA) 3	15-20	604	.943	644	.951
Facility (NYP) 1	15-44	597	**.970**	597	**.970**
Facility (NYP) 2	21-44	560.5	.869	629	.985
Facility (NYP) 3	15-20	87.5	.770	145	.916
Public health region (IME) 1	15-44	14,955	**.960**	14,955	**.960**
Public health region (IME) 2	21-44	11,346	.924	11,346	.924
Public health region (IME) 3	15-20	3609	.960	3,609	.960
Group billing provider (IME) 1	15-44	5	**.323**	148.5	**.909**
Group billing provider (IME) 2	21-44	3	.298	148	.914
Group billing provider (IME) 3	15-20	1	.313	129	.935
Health plan (WA HCA) 1	15-44	26,460	**.951**	26,460	**.951**
Health plan (WA HCA) 2	21-44	17,867.5	.930	17,867.5	.930
Health plan (WA HCA) 3	15-20	8,707.5	.908	8,707.5	.908
Health plan (MA) 1	15-44	7,362.5	**.916**	7,362.5	**.916**
Health plan (MA) 2	21-44	5,320	.876	5,320	.876
Health plan (MA) 3	15-20	1,683	.940	1,683	.940
Health plan (LA Medicaid) 1	15-44	65,818.5	**.966**	65,818.5	**.966**
Health plan (LA Medicaid) 2	21-44	48,151	.927	48,151	.927
Health plan (LA Medicaid) 3	15-20	17,667.5	.848	17,667.5	.848

Beta-binomial reliability estimates by age group

2016 Submission

The table below shows summary results of the reliability analyses at five levels (i.e., affiliate, health center, health plan, public health region and region by benefit type), stratified by three age categories (i.e., 15-20, 21-44, and 15-44). More detailed information about the analyses at each level can be found in Tables 1-4 (appended at the end of the form).

Level	Age group	Results: Median N	Results: ICC	Results: Reliability
Affiliate (PPFA) 1	15-20	4,839	.1164	.9984
Affiliate (PPFA) 2	21-44	11,648	.1232	.9994
Affiliate (PPFA) 3	15-44	16,590	.1191	.9996
Health centers (PPFA) (estimated within each affiliate) 1	15-20	366	.0612 (median)	.6096 - .9985
Health centers (PPFA) (estimated within each affiliate) 2	21-44	1,016	.0484 (median)	.6709 - .9990
Health centers (PPFA) (estimated within each affiliate) 3	15-44	1,379	.0581 (median)	.7056 - .9997
Public health region (IME) 1	15-20	1,438	.0121	.9461
Public health region (IME) 2	21-44	3,767	.0041	.9399
Public health region (IME) 3	15-44	5,205	.0034	.9461
Benefit type (IME) 1	15-20	11,699	.1268	.9988
Benefit type (IME) 2	21-44	33,051	.0057	.9895
Benefit type (IME) 3	15-44	44,750	.0463	.9991
Region by benefit type (IME) 1	15-20	716	.1929	.9942
Region by benefit type (IME) 2	21-44	2,325	.2148	.9984
Region by benefit type (IME) 3	15-44	2,954	.1920	.9986
Health plan/HMO (WMP) 1	15-20	1,231	.0017	.6767
Health plan/HMO (WMP) 2	21-44	3,936	.0029	.9206
Health plan/HMO (WMP) 3	15-44	5,167	.0018	.9048

For each level, the overall reliability was estimated using the medians as previously mentioned. ICCs, an

indicator of the proportion of variance explained by the groupings, are also shown. Similar studies of hierarchical binary outcomes estimate ICCs in a typical range of .02 - .18 (Fung et al, 2010). The moderate ICCs found in our analyses, combined with the large volume of patients at most levels, tend to generate high reliability estimates. Using the 'floor' of reliability, we also calculate the minimum number of cases required to achieve acceptable reliability thresholds for each level.

The estimated reliabilities remain above .90 for affiliates, for most affiliate groupings of health centers, for region, for benefit type, for region by benefit type, and 2 of 3 age groups at the health plan level. The ICCs at these levels were variable, ranging from low (e.g., <1%) to high (up to 21%). Of note, reliability did decline slightly in the analysis of the health centers within each affiliate and for the 15-20 age group at the health plan level. This would be expected since the volume of patients decreased, the cases per unit were less stable, and the rates were slightly more consistent among the health centers. The estimated reliabilities remain above .90 for most affiliates, and most levels below affiliate, due to sufficient patient volume. An exception occurred with two of the affiliates that contained only a single health center. Since there can be no variance in rates for a single unit, the health center level ICCs (and reliabilities) for those two affiliates are not included above.

It is commonly advised that reliability should be $\geq$ .90 for making decisions, and $\geq$ .70 for general reporting/monitoring (Eijkenaar, 2013; Adams, 2010). The Spearman-Brown prophecy allows one to test different values for ICC and patient volume per unit in order to predict expected reliability. Using an ICC value near the 20th percentile as a conservative expected correlation within units, we can compute the minimum recommended case load per level for each threshold of reliability. For example, for within-affiliate reporting of health centers, we used a conservative expected floor of .02 ICC to recommend that health centers have at least 115 patient cases for reporting rates to maintain >.70 reliability, and 450 cases to maintain >.90 reliability. The median ICC from actual data was nearly 3 times our conservative floor value (and most health centers exceeded this minimum number of cases) thus our reported reliabilities were considerably higher.

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?*)

2021 Submission

It is commonly advised that reliability should be $\geq$ .90 for making high-stakes decisions, and $\geq$ .70 for general reporting/monitoring (Eijkenaar, 2013; Adams, 2010). Our tested reliability is consistently greater than .70 at the facility, public health region, and health plan levels, showing adequate to high reliability at these levels. This was mostly driven by the large number of patients per unit at these levels.

Iowa Medicaid data does not contain data on clinical service sites; large group billing providers (with eligible female patient volume of >75 per year) were used to represent group practices for these data. This minimum threshold was selected since the entire distribution of group billing providers is positively skewed, with a high number of small office practices (many seeing fewer than 10 eligible patients annually). As one would expect (and as shown in many prior studies), including very small practices makes estimates unstable and less reliable (falling below .70). However, with the minimum threshold of 75 eligible patients annually, reliability improves greatly, exceeding .90 in all cases. Measure developers frequently recommend the minimum patient size approach for performance measurement when

estimating at the facility or provider level (HEDIS, 2007; Safran, 2007) and our analysis suggests that a minimum of 75 patients yields sufficient reliability.

2016 Submission

Despite the challenges of recoding claims data to obtain contraceptive rates, having large and diverse datasets available made a positive impact on reliability. For the PPFA data both at the affiliate level and at the next level down (groups of health centers within affiliate), we found reliabilities well above the commonly accepted .90 reliability threshold for reporting and decision-making. This was largely driven by two factors. First, the data exhibited adequate variation in the rates of contraceptive use at both the affiliate and lower levels. Second, the number of patients per unit at the affiliate level was mostly in the thousands, and at the lower levels, usually exceeded several hundred. For the IME data, the rates were much more uniform by region resulting in lower ICCs, but the volume of clients still enabled adequate reliability for distinguishing performance. When region was crossed by type of health plan or benefit the contraceptive rates were more variable among the units, so even given the smaller size of these analytic units the estimated reliabilities were higher.

In performing this analysis, we attempted to provide a conservative estimate of reliability wherever possible. Using medians rather than means, and presenting the 'floor' of reliability that may be observed for the smallest units, we bracket the results with worst-case scenarios. We further utilized a conservative value of ICC to recommend minimum patients per unit to maintain the .70 and .90 levels of reliability. In future years, analyses could examine the actual ICCs in order to make appropriate determinations about cases per unit. Yet even with these conservative methods, the 2014 data at the affiliate (state) and lower levels appears to provide sufficient reliability for reporting contraceptive rates.

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

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

☒ **Performance measure score**

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

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

2021 Submission

Empirical validity testing

We tested for convergent validity of the most or moderately effective contraceptive measure by exploring whether it was correlated with other similar quality measures listed below:

- **Cervical cancer screening:** Percentage of continuously eligible women ages 21 to 44 who were screened for cervical cancer using either of the following criteria:
    - Women ages 21 to 44 who had cervical cytology performed every 3 years;
    - Women ages 30 to 44 who had cervical cytology/human papillomavirus (HPV) co-testing performed every 5 years.

  The original cervical cancer screening specification includes women ages 21 to 64. We restricted the calculation of this measure to be among women ages 21 to 44 in order to match with the age range of the contraceptive care measure. For IME, we only had one year of data and thus were not able to include the look back period of 3-5 years as originally specified. The measure numerator only included women who received service during the measurement year. For PPFA, we had one year of data with variables indicating the dates of last cervical cytology and HPV testing. We used these variables to identify women who received services in the 3-5 year period.
- **Chlamydia screening**: Percentage of women ages 16 to 24 who were identified as sexually active and who had at least one test for chlamydia during the measurement year.

  When testing the correlation with chlamydia screening, we restricted the contraceptive care measure calculation to women ages 16 to 24 in order to match with the age range of chlamydia screening measure.
- **Encounter for contraceptive counseling**: Percentage of women ages 15 to 44 who received any contraceptive counseling during the measurement year.
- **Encounter for gynecological exam**: Percentage of women ages 15 to 44 who received any gynecological exam during the measurement year.

We hypothesized that facilities/groups that perform well on contraceptive care should perform well on cervical cancer screening, chlamydia screening, contraceptive counseling, and gynecological exams. Therefore, these related measures should be positively correlated to the contraceptive care measure. This hypothesis is based on the assumption that facilities offering contraceptive services are also more likely to provide other women's health services. We also hypothesize that the magnitude of correlations may be weak for cervical cancer screening and chlamydia screening due to the difference in recommended screening frequency and target population for these two measures compared to the contraceptive care measure. To test these correlations, we used two different approaches.

In the first approach, we used a Pearson's correlation test. This test estimates the strength of the linear association between two continuous variables. The correlation coefficient ranges from -1 to +1. A value of 1 indicates a perfect positive linear correlation between two variables. A value of 0 indicates no linear association. A value of -1 indicates a perfect negative linear relationship between two variables. We used a threshold of $p < .05$ to evaluate the statistical significance of test results.

Even though Pearson's correlation test is widely used to evaluate the correlation between two measures, it is only optimal in cases where linearity can be assumed. Crucially, the bounded nature of the variation in the proportion of contraceptive care measure (i.e., 0 and 1) means that estimates of association that assume linearity on the contraceptive care measure rates will be biased. This is a particular concern when the count of service events is either very high or very low relative to the total number of patients in a cluster. In addition, the correlations captured by the Pearson correlation matrix are averaged over the "true" and error variances. As a result, Pearson's correlation could downwardly

bias the correlation substantially in cases when the clusters are small with few patients, and where the measurement error is high.

Given these limitations with Pearson's correlation test we present a novel alternative approach. We employ a multilevel correlation estimation method to test the relationship between the contraceptive care measure and the related measures. The model is based on a multivariate generalized linear mixed model framework (Coull and Agresti, 2000). By employing a logit transformation of the binomial proportions, the model relaxes the linearity assumption on the original measurement scale. In addition, it analytically separates "true" score variance from measurement error by presenting measurement error as a random, binomial deviate, conditional on each cluster's "true" quality measure. Thus, the multilevel correlation estimation approach captures the correlation more accurately when the cluster size is small.

In the present analyses, the parameters of the multilevel model were estimated using a hierarchical SAS 9.4 GLIMMIX procedure with a log link function and fully unstructured residual error. Parameters were estimated by pseudo-maximum-likelihood using the Laplace method. The error structure was reported as correlation coefficients and variances. We are also able to provide 95% confidence limits for the estimates using likelihood bounds, which is far more informative than the single p-value for statistical significance. Rather than estimating all possible pairwise associations simultaneously, we estimated each pairwise association in a separate model in order to speed up and improve model convergence. In the appendix of the application, we provide a detailed description of the model with example statistical programing code.

Since Iowa Medicaid data does not contain facility information, we used billing providers who registered as "organizations" to represent group practices. In addition, we used a set of cutoffs to exclude group billing providers who served only a small number of patients during the measurement year. We did this to avoid including small rural practices that only have one doctor, even if it is registered as a group practice. We used 25, 50, and 75 as the cutoffs to show how the choice of a cutoff impacts the analysis. Using both the "organization" type of billing provider and the patient count cutoff, we ensure that we are only analyzing score level validity among large group practices.

Critical data elements

For each of the 6 Iowa Department of Public Health Title X Family Planning Clinics, about 70 female patients aged 15-44 years in 2019 were randomly sampled, resulting in a total of 423 patients. For each of these patients, data elements used for contraceptive care measure calculations were compared between the claims records and the patient charts, and agreement numbers were summarized in a 2 by 2 table (yes/yes, yes/no, no/yes, and no/no) for each element. We compared 10 data elements in total, including 7 most or moderately effective methods (Female sterilization, Implantable, IUD, Injectables, Contraceptive pills, Contraceptive patch, and Vaginal ring) and 3 exclusion criteria elements (Infecund, Currently pregnant or unknown pregnancy outcome, and Live births in the last 2 months of the year). Using the patient chart as the authoritative source, we calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), Cohen's Kappa statistics (McHugh, 2012; Viera and Garrett, 2005; Watson and Petrie, 2010) with 95% confidence intervals (CIs), and percent agreement for each data element.

2016 Submission

We used a systematic process to assess the face validity of the performance measure, i.e., whether the corresponding measure scores correctly reflect the quality of care provided and adequately identify differences in quality. Nine experts with the following characteristics were identified: (1) expertise in the delivery of contraceptive services, as evidenced by employment in a clinical or managerial capacity for at least 3 years during which they delivered contraceptive services in a clinical setting (i.e., public and private family planning and primary care providers, or health administrators); and (2) expertise in the use of performance measures, as evidenced by participation in at least one effort to collect and use performance measurement data for the purpose of improving clinical services in the setting(s) in which they work. Below is the final list of experts who participated in the assessment:

1. Carol Brady, MA, Project Director, Florida Association of Healthy Start Coalitions, Inc.
2. Anne Burke, MD, Associate Professor, School of Medicine, Johns Hopkins Bayview Medical Center
3. Vanessa Dalton, MD, MPH, Associate Professor, Director, Program on Women's Health Care Effectiveness Research, University of Michigan
4. Anne Dunlop, MD, MPH, Program Director, Preventive Medicine Division, Emory University School of Medicine
5. Daryn Eikner, MS, Vice President of Health Care Delivery, National Family Planning & Reproductive Health Association
6. Jan Engstrom, PhD, RN, CNM, WHNP-BC, Professor & Acting Chairperson, Department of Women, Children and Family Nursing, College of Nursing, Armour Academic Center
7. Mark Hathaway, MD, MPH, Senior Technical Advisor, Jhpiego – Johns Hopkins University
8. Michael Policar, MD, MPH, Clinical Professor of Obstetrics, Gynecology, and Reproductive Sciences, UCSF School of Medicine
9. Linda Wheal, Maternal Health Program Manager, Bureau of Quality Management, Illinois Department of Healthcare and Family Services

We contacted the selected experts to confirm consent to participate via email. Each expert panelist was sent a disclosure form to report any relevant financial or other competing interests; disclosures were compiled with brief biographies and shared with all panelists. Upon receipt of the disclosure form we sent the participant information about the measure specifications and other background information about the measure. Participants then participated in a webinar designed to provide important background information about the measure, how it is computed, the NQF endorsement process, and how the face validity assessment will be used in the application package that will be submitted to NQF. After reviewing the measure specifications and participating in the webinar the participants completed a survey (anonymous) that asked the following question about the measure:

The scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good and poor quality in contraceptive services:

1=Strongly Disagree
3=Neither Agree nor Disagree
5=Strongly Agree

ICD-10 Conversion:

We tested the measure specifications based on 2014 codes, but have also included the codes needed to calculate the measure using ICD-10 and 2015 NDC codes. Both sets of codes are attached. Our goal was to convert the measure to a new code set, fully consistent with the intent of the original measure. A description of how we converted from ICD-9 to ICD-10 is provided below, for each table in the measure specifications.

- **Sterilization for non-contraceptive reasons (Table UCM-A)**

We identified the 2015 ICD-10 codes for this table by using ICD-10 online conversion tools and confirming codes in the ICD-10-CM Expert for Physicians complete official code set, as well as with a clinical expert. These were confirmed with a clinical expert, Denise Wheeler, MS, Family Planning Director at the Iowa Department of Public Health.

- **Pregnancy codes (Table UCM-B)**

We identified the 2015 ICD-10 codes for this table by searching the NCHS/CMS publication, "ICD-10-CM Official Guidelines for Coding and Reporting, FY 2015". Pregnancy-related codes were found in "Chapter 15: Pregnancy, Childbirth and the Puerperium (O00-O9A)", and also Z codes for "outcome of delivery".

- **Known miscarriage, ectopic pregnancy, stillbirth, or induced abortion (Table UCM-C)**

These codes were identified by copying the Non-live Births Value Set from NCQA's Prenatal & Postpartum Care (PPC) measure (NQF#1517), as well as non-live birth codes in "Chapter 15: Pregnancy, Childbirth and Purperium (O00-O9A)". In the PPC measure, these codes are used to identify live births.

- **Delivery resulting in a live birth (Table UCM -D)**

These codes were identified by copying the Deliveries Value Set from NCQA's Prenatal & Postpartum Care (PPC) measure (NQF#1517), excluding extraction of products of conception retained and ectopic. In the PPC measure, these codes are used to identify live births.

- **Contraceptive codes (Tables UCM E, F and G)**

We used ICD-10 online conversion tools and confirming codes in the ICD-10-CM Expert for Physicians complete official code set. They were cross-checked against a ICD-10 conversion chart for family planning services that was prepared by Dr Michael Policar, from the University of California-San Francisco, and confirmed with a clinical expert, Denise Wheeler, MS, Family Planning Director at the Iowa Department of Public Health. NDC codes for 2015 were updated by using the codes for contraception contained in the HEDIS specifications for Chlamydia screening.

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

2021 Submission

Empirical validity testing

Tables below show summary results of the score level validity analyses. We utilized two statistical methods in this validity analysis to assess correlations between the contraceptive care measure and related measures at two levels of analysis (PPFA: facility, IME: group billing provider) and stratified by three age categories (i.e., 15-20, 21-44, and 15-44). Results from two methods are shown side-by-side. At the IME group billing provider level, we ran the analyses using 3 different minimum thresholds to exclude billing providers with fewer than 25, 50, and 75 eligible patients. Estimates for the cutoff of 75 is shown below and results using the cutoffs of 25 and 50 are shown in Table 7 (appended at the end of the form).

Correlation with selected related measures, Facility, PPFA 2019 — Multilevel correlation estimation

Related measures	Age Group	Median unit size of related measures	Pearson r	Correlation coefficients	95% Confidence Limits (lower, upper)
Contraceptive Counseling (1)	15-44	3,075	**.26**	**.55***	**(.28, .71)**
Contraceptive Counseling (2)	21-44	2,266	.25	.48*	(.19, .66)
Contraceptive Counseling (3)	15-20	623	.27*	.57*	(.30, .73)
Gynecological Examination (1)	15-44	3,075	**.36***	**.37***	**(.06, .60)**
Gynecological Examination (2)	21-44	2,266	.35*	.30	(-.02, .54)
Gynecological Examination (3)	15-20	623	.19	.55*	(.05, .81)
Cervical Cancer Screening†	21-44	2,217	**.40***	**.16**	**(-.13, .42)**
Chlamydia Screening†(1)	16-24	1,233	**.13**	**.02**	**(-.27, .30)**
Chlamydia Screening†(2)	21-24	657	.09	-.13	(-.40, .17)
Chlamydia Screening†(3)	16-20	503	.17	-.04	(-.33, .25)

*statistically significant at $p < .05$

†Age range of the related measure differs from that of the contraceptive care measure and the analysis was conducted among the overlapping population only, which was hypothesized to potentially attenuate the magnitude of the associations.

Correlation with selected related measures, Group Billing Provider, IME 2018

Related measures	Age Group	Median unit size of related measures	Pearson r	Multilevel correlation estimation: Correlation coefficients	Multilevel correlation estimation: 95% Confidence Limits (lower, upper)
Contraceptive Counseling (1)	15-44	**.56***	**.56***	**.61***	**(.52, .68)**
Contraceptive Counseling (2)	21-44	.54*	.54*	.58*	(.47, .67)
Contraceptive Counseling (3)	15-20	.59*	.59*	.63*	(.42, .77)
Gynecological Examination (1)	15-44	**.32***	**.32***	**.29***	**(.15, .41)**
Gynecological Examination (2)	21-44	.32*	.32*	.29*	(.14, .42)
Gynecological Examination (3)	15-20	.42*	.42*	.42*	(.06, .66)
Cervical Cancer Screening†	21-44	**.30***	**.30***	**.31***	**(.15, .44)**
Chlamydia Screening†(1)	16-24	**.23**	**.23**	**.21**	**(-.09, .46)**
Chlamydia Screening†(2)	21-24	.29	.29	.22	(-.21, .55)
Chlamydia Screening†(3)	16-20	.29	.29	.27	(-.16, .59)

*statistically significant at $p < .05$

†Age range of the related measure differs from that of the contraceptive care measure and the analysis was conducted among the overlapping population only, which was hypothesized to potentially attenuate the magnitude of the associations.

Critical data elements

The table below shows results of the data element level validity analyses. We calculated sensitivity, specificity, PPV, NPV, Cohen's Kappa statistics with 95% CIs, and percent agreement for each data element.

Data element validity test results, Iowa Department of Public Health Title X Grantee, 2019

Data elements	Age group	Sensitivity	Specificity	PPV	NPV	% agreement	Kappa	95% CI
Sterilization	15-44	**0.63**	**1.00**	**1.00**	**0.99**	**99.3%**	**0.766**	**0.502, 1.030**
*	21-44	0.57	1.00	1.00	0.99	99.0%	0.721	0.409, 1.034
*	15-20	1.00	1.00	1.00	1.00	100.0%	0.997	0.997, 0.997
IUD	15-44	**0.73**	**1.00**	**0.97**	**0.97**	**97.2%**	**0.820**	**0.719, 0.922**
*	21-44	0.71	1.00	0.96	0.96	96.4%	0.785	0.666, 0.905
*	15-20	0.86	1.00	1.00	0.99	99.1%	0.899	0.740, 1.058
Implantable	15-44	**0.76**	**1.00**	**0.98**	**0.95**	**95.3%**	**0.834**	**0.761, 0.907**
*	21-44	0.71	1.00	0.98	0.94	94.5%	0.774	0.677, 0.871
*	15-20	0.88	1.00	1.00	0.97	97.4%	0.843	0.751, 0.936
Injectables	15-44	**0.80**	**1.00**	**0.99**	**0.95**	**95.5%**	**0.860**	**0.796, 0.924**
*	21-44	0.77	1.00	0.98	0.94	94.8%	0.810	0.729, 0.891
*	15-20	0.88	1.00	1.00	0.97	97.4%	0.843	0.751, 0.936
Contraceptive pills	15-44	**0.56**	**0.99**	**0.94**	**0.87**	**88.2%**	**0.635**	**0.539, 0.731**
*	21-44	0.59	1.00	0.98	0.89	90.3%	0.662	0.552, 0.772
*	15-20	0.50	0.96	0.85	0.82	82.5%	0.476	0.287, 0.666
Contraceptive patch	15-44	**0.25**	**1.00**	**1.00**	**0.99**	**99.3%**	**0.398**	**-0.281, 1.077**
*	21-44	0.25	1.00	1.00	0.99	99.0%	0.397	-0.283, 1.076
*	15-20	NA†	1.00	NA†	1.00	100.0%	NA†	NA†
Vaginal ring	15-44	**0.70**	**1.00**	**1.00**	**0.99**	**99.3%**	**0.820**	**0.618, 1.023**
*	21-44	0.71	1.00	1.00	0.99	99.4%	0.828	0.594, 1.063
*	15-20	0.67	1.00	1.00	0.99	99.1%	0.790	0.391, 1.188
Infecund	15-44	**1.00**	**1.00**	**1.00**	**1.00**	**100.0%**	**1.000**	**1.000, 1.000**
*	21-44	1.00	1.00	1.00	1.00	100.0%	1.000	1.000, 1.000
*	15-20	1.00	1.00	1.00	1.00	100.0%	0.997	0.997, 0.997
Currently pregnant or unknown pregnancy outcome	15-44	**0.67**	**1.00**	**1.00**	**0.99**	**98.6%**	**0.794**	**0.629, 0.958**
*	21-44	0.64	1.00	1.00	0.99	98.7%	0.769	0.546, 0.991
*	15-20	0.71	1.00	1.00	0.98	98.2%	0.809	0.567, 1.050

Live birth in the last 2 months of the measurement year	15-44	**0.50**	**1.00**	**1.00**	**0.99**	**99.3%**	**0.664**	**0.284, 1.043**
*	21-44	0.40	1.00	1.00	0.99	99.0%	0.567	0.080, 1.054
*	15-20	1.00	1.00	1.00	1.00	100.0%	0.997	0.997, 0.997

†NA indicates that the validity statistics could not be calculated because the denominator to calculate the statistics was 0. E.g. the patient chart (authoritative source) indicates no women in the 15-20 age group used the contraceptive patch.

*cell intentionally left blank

2016 Submission

The mean rating from the face validity assessment for this measure was 4.67 with a median of 5 (Strongly Agree), range 4-5. There were 66.7% (n = 6) of respondents who strongly agreed and 33.3% (n = 3) of respondents who agreed that the scores obtained from this measure, as specified, will provide an accurate reflection of quality and can be used to distinguish good and poor quality in contraceptive services. One respondent replied that he or she thinks that "the proposed measures are valid measure of quality contraceptive care for healthy women" and one responded he or she "feels STRONGLY that the adoption of these measures will promote providers' and practices' attention to reproductive planning and contraceptive care as part and parcel of women's primary health care." One respondent strongly agrees "that the measure has excellent face validity as currently specified." He or she also responded, "However, in the future, we would suggest considering the use of a look-back period using claims data to identify previous use of long-acting contraceptives." One respondent pointed out that "quality of the indicator will in part depend on how well 'unintended' is characterized."

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?*)

2021 Submission

Empirical validity testing

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. Using the multilevel correlation estimation method, we observed mostly moderate to strong positive correlations between the contraceptive care measure with contraceptive counseling and gynecological exam measures at both facility and group billing provider levels among the 15-44 age group. Pearson's correlation test showed similar positive correlations except for a non-significant correlation with contraceptive counseling. We also found positive associations among the sub-age groups with contraceptive counseling and gynecological exam, although some of the associations were not statistically significant, likely due to smaller number of units in the analysis. For cervical cancer screening, both methods showed positive correlations, although the correlation was not statistically significant at the facility level when using the multilevel correlation estimation. For chlamydia screening, we did not observe any statistically significant associations at either facility or group billing provider levels.

The magnitude of correlation was generally slightly weaker using Pearson's correlation, as expected, since the distributional assumptions of this method are a poor fit to binary outcomes, resulting in underestimation. Although the Pearson correlation can be a rough approximation of correlation in binary outcomes for large units, cluster sizes become much smaller at the billing provider level, resulting in further attenuation. When we increased the minimum threshold to exclude billing providers with fewer than 25, 50, and 75 eligible patients, as shown in Table 7, the magnitude of Pearson's correlation increased, supporting this theory. We demonstrate that our generalized linear multilevel estimation more closely captures the "true" correlation between two measures, and is better suited for binary outcomes and smaller units of analysis.

Overall, we observed positive correlations between the contraceptive care measure and those services that (in theory) should be related (contraceptive counseling, gynecological examination, and cervical cancer screening); these were highly consistent with our hypotheses and provide good evidence for validity of the contraceptive care measure at the score level. We observed no significant associations for chlamydia screening. We speculate that the absence of significant association may be due to the censoring of age to enable comparisons; or from application of standardized clinical guideline (e.g., from the Centers for Disease Control and Prevention https://www.cdc.gov/std/prevention/screeningreccs.htm) for this service, which could limit variation of the measure. It is also possible that many women visit a doctor for sexually transmitted disease screening when they are concerned or experiencing symptoms and may not want to obtain contraception at that time; whereas contraception is a more routine part of well woman visits such as gynecological visits.

Critical data elements

Sensitivity was above 0.5 for the majority of the data elements, except for contraceptive patch, whereas specificity, PPV, and NPV were above 0.8 for all data elements. Percent agreement was consistently over 80% for all data elements. We also observed statistically significant Kappa above 0.6 for all data elements except for contraceptive patch, indicating moderate to almost perfect agreement between the claims records and the patient charts (Watson and Petrie, 2010). Overall, our data provide fairly strong evidence for validity of the contraceptive care measure at the data element level.

2016 Submission

We think that the responses to the face validity assessment indicate that the measure will provide an accurate reflection of quality and can be used to distinguish good and poor quality in contraceptive services.

2b2. EXCLUSIONS ANALYSIS

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*)

2021 Submission

The rationale for exclusion is due to the fact that some women are not at risk of unintended pregnancy due to infecundity or pregnancy. Also, women with live births that occurred in the last 2 months of the measurement year might not have had a chance to receive postpartum contraceptive care in the 60-day time frame and were therefore excluded. After limiting our datasets to women 15-44 years of age, the following exclusions were analyzed for frequency and variability across various units included in our analysis. Codes utilized for the exclusions are in the tables referenced (see the Excel file named NQF 2903 Codes 2021.xlsx).

1. Women who were infecund due to non-contraceptive reasons such as natural menopause or oophorectomy. The codes (ICD-10-CM, ICD-10-PCS, and CPT) utilized to exclude these women are listed in Table CCW-A, Codes Indicating Sterilization for Non-Contraceptive Reasons.
2. Women who were still pregnant at the end of the year because they did not have a pregnancy outcome code indicating a non-live birth (Table CCW-C) or a live birth (Table CCW-D).
    - Codes for non-live births were drawn from the HEDIS measure of Prenatal and Postpartum Care, and procedure codes (CPT, ICD-10-PCS codes) were added.
    - Codes for live birth include CPT and ICD-10-PCS codes also from the HEDIS measure of Prenatal and Postpartum Care, and ICD-10-CM codes for live births were added.
3. Women who had a live birth in the last 2 months of the measurement year. A two-month period was selected because the American College of Obstetricians and Gynecologists (ACOG) recommends having a postpartum visit by 6 weeks, and an additional 2 weeks was added to allow for reasonable delays in attending the postpartum visit. Live births were identified for this exclusion by the codes listed in Table CCW-D.

To exclude women with a live birth in the last 2 months or those still pregnant at the end of the year, women who were pregnant at any point in the measurement year were first identified by using the codes listed in Table CCW-B, Codes Indicating a Pregnancy. We selected this list of codes by reviewing the following documents:

- CMS & NCHS (2020). ICD-10-CM Official Guidelines for Coding and Reporting FY 2021. Available online at: http://www.cdc.gov/nchs/icd/icd10cm.htm.
- CMS & NCHS (2020). ICD-10-CM Official Guidelines for Coding and Reporting FY2020. Available online at: https://www.cms.gov/Medicare/Coding/ICD10/2020-ICD-10-PCS

Exclusions were performed in a hierarchical manner in the order listed above.

2016 Submission

Exclusions were not formally tested. The rationale for exclusion was due to the fact that some women are not at risk of unintended pregnancy due to infecundity or pregnancy.

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*)

2021 Submissio

We examined the overall frequencies and proportions of women excluded for each exclusion criterion in 3 datasets. Categories are not mutually exclusive.

Frequency of denominator exclusions for the contraceptive care measure, 129,652 women 15-44 years of age in 56 PPFA health centers, 2019

Category	N (%)	Distribution across health centers (in percentiles): 25th	Distribution across health centers (in percentiles): 50th	Distribution across health centers (in percentiles): 75th
Exclusion: Infecund for non-contraceptive reasons	**18 (.01)**	**.00**	**.00**	**.01**
Exclusion: Had a live birth in the last 2 months of the measurement year	**0 (.00)**	**0 (.00)**	**0 (.00)**	**0 (.00)**
Exclusion: Pregnant or their pregnancy outcome was unknown at the end of the measurement year	**5,656 (4.4)**	**1.8**	**4.1**	**5.6**
Number of women 15-44 years of age, after exclusions	**123,978**	*	*	*

*cell intentionally left blank

Frequency of denominator exclusions for the contraceptive care measure, 208,709 women 15-44 years of age in 5 WA HCA health plans, 2019

Category	N (%)	Distribution across health centers (in percentiles): 25th	Distribution across health centers (in percentiles): 50th	Distribution across health centers (in percentiles): 75th
Exclusion: Infecund for non-contraceptive reasons	3,568 (1.7)	1.5	1.6	1.6
Exclusion: Had a live birth in the last 2 months of the measurement year	1,785 (.9)	.8	.9	.9
Exclusion: Pregnant or their pregnancy outcome was unknown at the end of the measurement year	6,936 (3.3)	3.2	3.2	3.4
Number of women 15-44 years of age, after exclusions	196,568	*	*	*

*cell intentionally left blank

Frequency of denominator exclusions for the contraceptive care measure, 126,069 women 15-44 years of age in 6 IME public health regions, 2018

*cell intentionally left blank

Category	N (%)	Distribution across health centers (in percentiles): 25th	Distribution across health centers (in percentiles): 50th	Distribution across health centers (in percentiles): 75th
Exclusion: Infecund for non-contraceptive reasons	1,889 (1.5)	1.4	1.6	1.7
Exclusion: Had a live birth in the last 2 months of the measurement year	5,733 (4.6)	4.5	4.5	4.6
Exclusion: Pregnant or their pregnancy outcome was unknown at the end of the measurement year	1,555 (1.2)	1.1	1.2	1.3
Number of women 15-44 years of age, after exclusions	116,892	*	*	*

2016 Submission

The table below shows the number of women excluded in each of the two datasets, presented by the reason for exclusion.

Category	Number of women: PPFA, 2014	Number of women: IME, 2013	Number of women: WMP, 2014
Women 15-44 years of age	950,647	49,232	132,940
Exclusion: Infecund for non-contraceptive reasons	83	169	2,025
Exclusion: Had a live birth in the last 2 months of the measurement year	7	520	2,995
Exclusion: Pregnant or their pregnancy outcome was unknown at the end of the measurement year	111,685	3793	9,611
Number of women 15-44 years of age, after exclusions	838,872	44,750	118,309

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*)

2021 Submission

The frequency of exclusions for the datasets analyzed is low. About 4.4% (PPFA), 7.3% (IME) and 5.1% (WA HCA) of women 15-44 years of age were excluded from the measure denominator. The distributions across units of analysis were as expected. The relative contribution of each type of exclusion varied by data set (e.g., live births in the last 2 months of the year were a larger population in IME dataset than the PPFA dataset). These differences likely exist because the programs emphasize different areas of health services. The PPFA program focuses primarily on delivery of outpatient reproductive health care while the state Medicaid programs (IME, WA HCA) offer a wider range of primary, acute, and curative care services.

The exclusions are utilized so that women who may not need nor have an opportunity to obtain contraception to prevent unintended pregnancy are removed from the denominator. Without these exclusions for the denominator, it may appear that fewer women have access to a wide range of most and moderately effective contraception, making it difficult to distinguish true differences in measure scores across health facilities, health plans, clinician groups/practices, regions, and states. Thus, we believe that the benefits of applying the exclusion criteria outweigh their burden.

2016 Submission

When combined, the total number of exclusions in each of the two data sets comprised 11.8% (PPFA), 9.1% (IME) and 11% (WMP) of all women 15-44 years of age, although the relative contribution of each type of exclusion varied by data set (e.g.., live births in the last 2 months of the year were a relatively larger population in IME dataset than the PPFA dataset). These differences are likely explained by the fact that the emphasis of each program is slightly different, with the PPFA program more heavily focused on delivery of reproductive health care while the IME and WMP programs offer a wider range of primary, acute and curative care services. The number of women excluded will have a noticeable impact on the rates, and will be important to reassure providers that the measure is as 'fair' in terms of identifying the population at risk as claims data will allow it to be. For these reasons, we believe that the burden of applying the exclusion criteria is outweighed by the benefits of doing so.

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

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

2b3.1. What method of controlling for differences in case mix is used?

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with risk factors**

☐ **Stratification by risk categories**

☐ **Other,**

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

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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.

2021 Submission

We do not believe that risk adjustment is justified. Variations in contraceptive use by socio-demographic characteristics exist in part due to modifiable clinical and programmatic considerations, and not different biological responses to contraception. Providers may also see variation by socio-demographic characteristics locally, but we believe that these differences will be reduced if contraceptive services are offered in a client-centered manner, as defined by CDC-OPA recommendations for providing quality family planning services (Gavin 2014, Gavin 2016, Gavin 2017).

To investigate differences in most and moderately effective contraceptive use, a special analysis of data from the National Survey of Family Growth (NSFG) 2015-2017 was conducted (see table below). This analysis suggests that there are statistically significant differences by age group (for ages 20-29 compared to ages 30-44) and among women who have never been married (compared to women of other marital status). However, no significant differences occur between race/ethnicity, most categories of marital status, and poverty level.

Percentage of women 15-44 years of age at risk of unintended pregnancy* that used a most or moderately effective method of contraception,

National Survey of Family Growth, 2015-2017

Measures	Frequency	Weighted Frequency	Percent	95% Confidence Limits
Age: 15-19	163	2,142,115	56.31	46.34 – 66.28
Age: 20-29	697	8,676,773	52.79	47.96 – 57.61
Age: 30-44	1,234	17,661,227	64.66	60.91 – 68.41
Race/ethnicity: Hispanic	409	5,599,163	55.91	50.06 – 61.76
Race/ethnicity: Non-Hispanic White	1,060	16,580,506	63.34	59.14 – 67.55
Race/ethnicity: Non-Hispanic Black	456	3,741,514	58.52	53.22 – 63.82
Marital status: Married	793	12,740,525	64.08	59.30 – 68.85
Marital status: Cohabitating	306	4,714,726	64.86	57.05 – 72.67
Marital status: Widowed/divorced/separated	239	2,952,812	64.12	57.42 – 70.83
Marital status: Never married	756	8,072,052	51.1	47.07 – 55.13
Percent Federal poverty level: <100%	595	6,694,429	61.75	56.58 – 66.93
Percent Federal poverty level: 100-199%	462	5,984,161	53.48	47.60 – 59.38
Percent Federal poverty level: 200-399%	549	8,242,219	63.59	59.17 – 68.01
Percent Federal poverty level: 400-499%	158	2,216,392	57.37	47.69 -67.04
Percent Federal poverty level: 500+%	330	5,342,914	61.41	55.97 -66.85

*Women are considered to be at risk of unintended pregnancy if they are not pregnant, not seeking pregnancy, are fecund, and have ever had sex.

2016 Submission

We do not believe that risk adjustment is justified. Although there are potential variations in contraceptive use by socio-demographic characteristics, the reason for those patterns is based on modifiable clinical and programmatic considerations rather than differing biological responses to contraception. Although providers may see some local variations by socio-demographic characteristics, we do not believe that these differences will be maintained if contraceptive services are offered in a

client-centered manner, as defined by CDC-OPA recommendations for providing quality family planning services (CDC-OPA, 2014).

A special analysis of data from the National Survey of Family Growth (NSFG), 2011-2013, was conducted to explore disparities in the use of most and moderately effective methods of contraception (see table below). This analysis suggests that there are statistically significant differences by age and for women who were never married. However, there were no significant differences by race/ethnicity, most categories of marital status, and poverty level.

Percentage of women 15-44 years of age at risk of unintended pregnancy* that used a most or moderately effective method of contraception, National Survey of Family Growth, 2011-2013

Measures	Frequency	Weighted Frequency	Percent	95% Confidence Limits
Age	- -	- -	- -	- -
15-19	183	1,740,000	43.5	35.98 - 51.10
20-29	919	9,341,000	56.6	52.90 - 60.36
30-44	1,356	17,342,000	67.2	64.06 - 70.31
Race/ethnicity	- -	- -	- -	- -
Hispanic	576	5,229,000	57.3	52.92 - 61.64
NH White	1,211	17,373,000	64.7	61.28 - 68.10
NH Black	494	3,657,000	56.6	51.09 - 62.16
Marital status	- -	- -	- -	- -
Married	941	13,629,000	70.9	67.86 - 73.98
Cohab	402	4,481,000	62.4	56.26 - 68.58
Wid/div/sep	335	3,173,000	62.1	56.25 - 67.90
Never married	780	7,139,000	48.3	43.69 - 52.81
Percent Federal poverty level	- -	- -	- -	- -
<100%	825	7,335,000	57.6	53.78 - 61.42
100-199%	555	6,015,000	60.6	56.31 - 64.88
200-399%	656	8,608,000	63.8	59.84 - 67.69

Measures	Frequency	Weighted Frequency	Percent	95% Confidence Limits
400-499%	152	2,462,000	66.6	57.65 - 75.45
500+%	270	4,001,000	62.1	54.83 - 69.41

- - cell intentionally left blank

* Women are considered to be at risk of unintended pregnancy if they are not pregnant, not seeking pregnancy, are fecund, and have ever had sex.

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?

2021 Submission

We recommend stratifying by age group so that percentages for adolescent and adult women can be calculated separately for quality improvement (QI) purposes. Given different care delivery models among adolescents, HHS's Centers for Disease Control and Prevention (CDC), American Academy of Pediatrics (AAP), and American College of Obstetricians and Gynecologists (ACOG) have published patient-centered counseling recommendations specifically for this population (Gavin 2014, ACOG 2017 Committee Opinion 710, Menon 2020). Though current guidance notes that most and moderately effective methods, including LARC methods, are safe and recommended for teen and nulliparous populations, it can still be difficult for these populations to access these highly effective contraceptive methods (Ott 2014, ACOG 2017 Committee Opinion 699, Menon 2020). Studies report that adolescents experience more unintended pregnancies (Coles 2011, Ahrens 2018) which may result in adverse outcomes for mothers and infants. For these reasons, it is particularly important to measure most and moderately effective contraceptive provision among the adolescent population.

2016 Submission

We recommended stratifying the client population by age so that rates for adolescents can be tracked separately from those for adult women. We propose this stratification for purposes of QI but not as a method of risk-adjustment. Teen pregnancy is worthy of a separate focus because of the large potential negative impact on the life of the teen and her child(ren), and the existence of unique programs and contraceptive counseling approaches tailored to this population.

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)**

2021 Submission

To affirm stratifying by age group to calculate measure scores for adolescent and adult women separately, we reviewed current clinical guidelines for contraception for women of reproductive age (i.e., women ages 15-44) as well as women ages 15-20. We also examined published studies and systematic reviews that focused on facilitators and barriers to contraception among women who wish to prevent pregnancy. The literature is summarized in section 2b3.3a above.

2016 Submission

Not asked in previous submission

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

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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.**

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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*)**:**

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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

2021 Submission

Not applicable.

2016 Submission

Not applicable.

2b3.9. Results of Risk Stratification Analysis:

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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*)

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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*)

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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

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

2021 Submission

Because our datasets are designed to represent the census of all claims available, rates are assumed to reflect "true" rates by unit for the data year. Non-sampling error (such as coding or measurement error) is not estimable given our limited access to the claims data and processes. Any differences in rates must therefore be evaluated based on practical or clinically meaningful impact. We present calculated measure rates at each level for all age groups for all data sources for such consideration.

One can set up a model in which the units of performance measurement (despite our census of all extant units) represent a sample from the 'infinite universe' of possible units. These units are modeled as if they were a random sampling of units from an infinitely large entity of units. We considered differences in performance using the PPFA data to illustrate this hypothetical example, with the caveat that the discussion is strictly speculative to support this section. To examine differences we simply calculated 95% confidence intervals for the unadjusted metric results for women 15-44 years of age in all facilities. If a facility's confidence interval did not include the grand mean rate across all facilities, then the facility was identified as better or worse than average. Note that a statistically significant difference is largely dependent on size of the measured units. A small facility with few patient cases might exhibit low rates, but not be "statistically different" from the average; or alternatively, a large entity with many patients being identified as "below average" when the difference might be negligible from a quality-of-care perspective. Other rubrics for identifying differences might be considered including nonparametric rank-order methods such as lowest percentiles.

Because the measure is most appropriately utilized to identify entities with very low rates of contraceptive provision relative to other units (perhaps suggesting structural barriers to access), we also developed a convenient empirical Bayes tool for setting a user-specified 'floor' value and identifying all units that fall below the floor value (with 95% confidence accounting for unit size). This tool is included as an appendix for consideration and might be generally applicable within the clinical quality improvement field.

Given the sensitive and context-dependent nature of quality improvement activities for contraceptive care, we strongly recommend that any methods for addressing performance gaps are developed carefully in conjunction with established guidelines for patient-centered contraceptive care. Because the interpretation of these measures is context dependent, clinically meaningful differences are best evaluated by subject matter experts who are familiar with the healthcare delivery organizations and their populations.

2016 Submission

Due to the fact that our dataset represents a census of all claims available, rates are assumed to reflect 'true' rates by unit for the data year. Non-sampling error (such as coding or measurement error) is not estimable given our limited access to the claims data and processes. Thus we do not present any

confidence intervals for inferential testing results. These assumed-true differences in rates must therefore be evaluated based on practical or clinically meaningful impact.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinical/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*)

2021 Submission

We excluded 2 PPFA facilities with denominators less than 75, resulting in 54 facilities in the analysis. The distribution for facility rate is shown in the table below.

Distribution for facility Most & Moderately Effective methods rate (%) in 54 PPFA facilities, 2019

Mean	SD	Minimum	10th percentile	25th percentile	Median	75th percentile	90th percentile	Maximum
61.2	12.3	2.1	54.9	58.7	62.9	65.8	69.9	81.0

Using the approach described in 2b4.1, 24 facilities (44.4%) of 54 PPFA facilities were rated as higher than the mean (i.e. the lower limit of facility's 95% confidence interval was > 61.2) and 15 facilities (27.8%) were identified as lower than the mean (i.e. the upper limit of facility's 95% confidence interval was < 61.2). Another 15 facilities were either higher or lower than the mean (61.2) but their results were not statistically significant.

The table below summarizes measure rates at each level. More detailed information about the variation of rates by unit within each level can be found in Tables 1-6, which are appended at the end of this document.

Provision of Most & Moderately Effective methods

Level	Age group	Rate Mean (range)
Facility (PPFA), n=56	15-44	.612 (.329 - 1.00)
*	21-44	.583 (.328 - 1.00)
*	15-20	.712 (.340 - 1.00)
Facility (NYP), n=31	15-44	.427 (.037 - 1.00)
*	21-44	.429 (.018 - 1.00)
*	15-20	.422 (.071 - 1.00)

Level	Age group	Rate Mean (range)
Public Health Region (IME), n=6	15-44	.307 (.288 - .349)
*	21-44	.287 (.272 - .319)
*	15-20	.368 (.331 - .429)
Group Billing Provider (IME), n=3081	15-44	.331 (.000 - 1.00)
*	21-44	.309 (.000 - 1.00)
*	15-20	.401 (.000 - 1.00)
Health Plan (WA HCA), n=5	15-44	.296 (.277 - .307)
*	21-44	.287 (.265 - .298)
*	15-20	.318 (.280 - .328)
Health Plan (MA), n=21	15-44	.231 (.184 - .260)
*	21-44	.242 (.205 - .312)
*	15-20	.197 (.102 - .285)
Health Plan (LA Medicaid), n=5	15-44	.313 (.290 - .322)
*	21-44	.297 (.281 - .305)
*	15-20	.356 (.334 - .366)

*cell intentionally left blank

2016 Submission

The table below summarizes rates at each level. As noted above, since our data contain the entirety of the defined population, estimation of sampling error and related inferential statistics such as confidence intervals are not applicable. More detailed information about the variation of rates by unit within each level can be found in Tables 1-3, which are appended at the end of this document.

Level	Age group	Rate (Provision of most and moderately effective methods)
Affiliate (PPFA), n=25 Mean (range)	15-20	.73 (.37-.90)
*	21-44	.66 (.28-.84)
*	15-44	.68 (.31-.86)
Health center (PPFA), n=363 Mean (range)	15-20	.73 (.00-1.0)
*	21-44	.66 (.00-1.0)
*	15-44	.68 (.00-1.0)
Public health region (IME) Mean (range)	15-20	.62 (.52-.67)
*	21-44	.60 (.56-.65)
*	15-44	.60 (.58-.66)
Benefit type (IME) Mean (range)	15-20	.62 (.40-.79)
*	21-44	.60 (.28-.73)
*	15-44	.60 (.32-.74)
PH Region by benefit type (IME) Mean (range)	15-20	.62 (.39-.84)
*	21-44	.60 (.27-.74)

Level	Age group	Rate (Provision of most and moderately effective methods)
*	15-44	.60 (.31-.78)
Health plan/HMO (WMP) Mean (range)	15-20	.46 (.42-.50)
*	21-44	.38 (.34-.43)
*	15-44	.40 (.36-.45)

*cell intentionally left blank

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?*)

2021 Submission

This measure can reliably distinguish facilities with better- and worse-than-average performance. Facilities that were identified as statistically significantly better or worse than the average had scores that were on average 12% (range: 2% - 59%) lower or 7% (range: 2% - 20%) higher than the mean. However, as noted, this is only one of many potential methods for examining performance differences. As noted, only subject matter experts with an understanding of the healthcare delivery context should determine meaningful differences in performance. We also provided a tool for identifying those units falling below a user-specified 'floor' value with 95% confidence (while accounting for unit size and empirical distribution), to aid in assessments by quality improvement professionals.

Measure rates vary considerably across almost all levels, which suggest that identifying meaningful differences in performance among measured entities is possible. These differences also demonstrate that sizeable room for improvement exists in measure scores. To discourage coercion into use of contraception or a certain contraceptive method, no specific target has been determined for this measure. We do not expect measure rates to reach 100% because some women will make informed decisions to choose less effective contraception, even when offered the full range of methods and with financial or logistical barriers to access removed.

Since 2017, OPA has met with an expert panel three times to discuss appropriate measure use and interpretation in different health systems (e.g. programs with a reproductive health services focus compared to general health care providers). To ensure that health systems employ a client-centered approach to implementation, the expert panel has recommended using this measure with a patient-reported outcome performance measure (PRO-PM) for contraceptive counseling. This PRO-PM, the Person-Centered Contraceptive Counseling (PCCC) measure, gathers information on the patient's contraceptive services experience. Together, these two measures may provide a more complete understanding of factors involved in clients' contraceptive care. Through a multi-organization partnership led by UCSF and the National Association of Community Health Centers (NACHC), several federally qualified health center (FQHC) networks are currently testing the contraceptive care and PCCC measures in tandem use.

Members of the expert panel have also developed guidance for implementing the measure in various programmatic contexts. For example, PPFA released a policy paper in collaboration with Manatt Health that helps state policymakers and payers implement contraceptive care quality measures to improve access to all forms of contraception. Serving as a tool for policymakers, this paper details how to incorporate contraceptive care quality measures in Value Based Payment (VBP) initiatives to both ensure agency in women's contraceptive choices and develop strategies to improve people's access to contraception (https://www.plannedparenthood.org/uploads/filer_public/7e/90/7e90b4cb-4b3d-499f-8c6c-f31ab865b621/ppfa-manatt_measuring_quality_contraceptive_care.pdf).

If the measure maintains its NQF endorsement, OPA will continue to meet with its expert panel to further develop and refine recommendations promoting client-centered measure interpretation and utilization, which includes tandem use of this contraceptive provision measure and the PCCC.

2016 Submission

There are very large and meaningful differences in rates across almost all levels. These differences suggest that it will be possible to identify meaningful differences in performance across measured entities. It also reinforces that there is substantial room for improvement in measure scores.

As more experience is gained from using the measures in different programmatic contexts (e.g., in programs focused on reproductive health services versus general health care providers), it will be possible to recommend benchmarks for the different programmatic contexts. These benchmarks may also take into account subject matter expertise, economic costs, risks of maternal health, pregnancy or birth outcomes, and other contextual criteria. If the measure is endorsed by NQF, OPA expects to convene an expert panel within 2 years to identify appropriate criteria and apply those criteria for the development of recommended benchmark.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

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

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

2021 Submission

Not applicable.

2016 Submission

Not applicable.

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*)
2021 Submission

Not applicable.

2016 Submission

Not applicable.

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*)
2021 Submission

Not applicable.

2016 Submission

Not applicable.

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*)
2021 Submission

Not applicable.

2016 Submission

Not applicable.

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

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

2021 Submission

The data source for this measure is claims data. Claims data usually has very little missing data because it is used for billing, which also makes determining when claims data is missing challenging.

2016 Submission

The data source for this measure is claims data. Due to the nature of claims data (i.e., for billing purposes), there is typically very little missing data; further, it is difficult to ascertain when claims data is or is not missing.

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*)

2021 Submission

Not addressed due to the nature of claims data.

2016 Submission

Not addressed due to the nature of claims data.

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 non-responders) 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*)

2021 Submission

Not applicable.

2016 Submission

Not applicable.

References

Adams, John L. (2009) *The Reliability of Provider Profiling: A Tutorial*. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/technical_reports/TR653.html.

Adams JL, Mehrotra A, Thomas JW, McGlynn EA (2010). Physician cost profiling-reliability and risk of misclassification. New England Journal of Medicine; 362:1014–1021.

Adams, J. L., & Paddock, S. M. (2017). Misclassification Risk of Tier-Based Physician Quality Performance Systems. *Health services research*, 52(4), 1277-1296.

Ahrens, K. A., Thoma, M. E., Copen, C. E., Frederiksen, B. N., Decker, E. J., & Moskosky, S. (2018). Unintended pregnancy and interpregnancy interval by maternal age, National Survey of Family Growth. *Contraception*, *98*(1), 52–55. https://doi.org/10.1016/j.contraception.2018.02.013

Blair, R., Liu, J., Rosenau, M., Brannan, M., Hazelwood, N., Gray, K. F. & Schmitt, A. (2015). Development of Quality Measures for Inpatient Psychiatric Facilities. *Development*, 2, 04.

Coles, M. S., Makino, K. K., & Stanwood, N. L. (2011). Contraceptive experiences among adolescents who experience unintended birth. *Contraception*, *84*(6), 578–584. https://doi.org/10.1016/j.contraception.2011.03.008

Committee on Adolescent Health Care (2017). Committee Opinion No 699: Adolescent Pregnancy, Contraception, and Sexual Activity. *Obstetrics and gynecology*, *129*(5), e142–e149. https://doi.org/10.1097/AOG.0000000000002045

Committee on Adolescent Health Care (2017). Committee Opinion No. 710: Counseling Adolescents About Contraception. *Obstetrics and gynecology*, *130*(2), e74–e80. https://doi.org/10.1097/AOG.0000000000002234

Coull, B. A., & Agresti, A. (2000). Random effects modeling of multiple binomial responses using the multivariate binomial logit-normal distribution. Biometrics, 56(1), 73-80.

Eijkenaar F, van Vliet RC. (2013). Profiling individual physicians using administrative data from a single insurer: variance components, reliability, and implications for performance improvement efforts. Med Care. 51(8):731-9

Gavin, L., Moskosky, S., Carter, M., Curtis, K., Glass, E., Godfrey, E., Marcell, A., Mautone-Smith, N., Pazol, K., Tepper, N., Zapata, L., & Centers for Disease Control and Prevention (CDC) (2014). Providing quality family planning services: Recommendations of CDC and the U.S. Office of Population Affairs. *MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 63*(RR-04), 1–54.

Gavin, L., & Pazol, K. (2016). Update: Providing Quality Family Planning Services - Recommendations from CDC and the U.S. Office of Population Affairs, 2015. *MMWR. Morbidity and mortality weekly report*, *65*(9), 231–234. https://doi.org/10.15585/mmwr.mm6509a3

Gavin, L., Pazol, K., & Ahrens, K. (2017). Update: Providing Quality Family Planning Services - Recommendations from CDC and the U.S. Office of Population Affairs, 2017. *MMWR. Morbidity and mortality weekly report*, *66*(50), 1383–1385. https://doi.org/10.15585/mmwr.mm6650a4

HEDIS, N. (2007). Technical Specifications for Physician Measurement. Washington DC: National Committee for Quality Assurance.

Kazis, L. E., Rogers, W. H., Rothendler, J., Qian, S., Selim, A., Edelen, M. O.& Butcher, E. (2017). *Outcome performance measure development for persons with multiple chronic conditions*. RAND.

Kumar, N., & Brown, J. D. (2016). Access Barriers to Long-Acting Reversible Contraceptives for Adolescents. *The Journal of adolescent health: official publication of the Society for Adolescent Medicine*, *59*(3), 248–253. https://doi.org/10.1016/j.jadohealth.2016.03.039

Menon, S., & COMMITTEE ON ADOLESCENCE (2020). Long-Acting Reversible Contraception: Specific Issues for Adolescents. *Pediatrics*, *146*(2), e2020007252. https://doi.org/10.1542/peds.2020-007252

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. Biochemia medica: Biochemia medica, 22(3), 276-282.

Pritt, N. M., Norris, A. H., & Berlan, E. D. (2017). Barriers and Facilitators to Adolescents' Use of Long-Acting Reversible Contraceptives. *Journal of pediatric and adolescent gynecology*, *30*(1), 18–22. https://doi.org/10.1016/j.jpag.2016.07.002

Safran, D. G., Karp, M., Coltin, K., Chang, H., Li, A., Ogren, J., & Rogers, W. H. (2006). Measuring patients' experiences with individual primary care physicians: results of a statewide demonstration project. Journal of General Internal Medicine, 21(1), 13-21.

Staggs, V. S., & Cramer, E. (2016). Reliability of pressure ulcer rates: how precisely can we differentiate among hospital units, and does the standard signal-noise reliability measure reflect this precision? *Research in nursing & health*, 39(4), 298-305.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. Fam med, 37(5), 360-363.

Watson, P. F., & Petrie, A. (2010). Method agreement analysis: a review of correct methodology. Theriogenology, 73(9), 1167-1179.

LIST OF (LANDSCAPE) TABLES

2021 Submission

2016 Submission

2021 Submission

Table 1. Rates and reliabilities for most or moderately effective contraceptive method provision by facility, NewYork-Presbyterian Hospital system, 2018.

Facility ID	MEM_Method Most: 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability (all units))	MEM_Method Most: 15 to <21 Years (Reliability (unit size≥75))	MEM_MethodMost: 21 to 44 years (Most/Mod Provision)	MEM_MethodMost: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability (all units))	MEM_Method Most: 21 to 44 years (Reliability (unit size≥75))	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (Reliability (all units))	MEM_MethodMost: all age groups (Reliability (unit size≥75))
101	2	2	1.000	0.110	*	20	90	0.222	0.922	0.926	22	92	0.239	0.901	0.901
102	3	12	0.250	0.426	*	59	436	0.135	0.983	0.984	62	448	0.138	0.978	0.978
103	12	38	0.316	0.701	*	182	1083	0.168	0.993	0.993	194	1121	0.173	0.991	0.991
201	194	614	0.316	0.974	0.978	5	17	0.294	0.691	*	199	631	0.315	0.984	0.984
202	41	88	0.466	0.845	0.867	323	1032	0.313	0.993	0.993	364	1120	0.325	0.991	0.991
203	53	114	0.465	0.876	0.894	577	1184	0.487	0.994	0.994	630	1298	0.485	0.992	0.992
204	16	35	0.457	0.684	*	208	522	0.398	0.986	0.986	224	557	0.402	0.982	0.982
205	944	1568	0.602	0.990	0.991	4878	8894	0.548	0.999	0.999	5822	10462	0.556	0.999	0.999
301	2	7	0.286	0.302	*	79	489	0.162	0.985	0.986	81	496	0.163	0.980	0.980
302	48	271	0.177	0.944	0.952	1	2	0.500	0.208	*	49	273	0.179	0.964	0.964
303	111	228	0.487	0.934	0.944	1068	2161	0.494	0.996	0.997	1179	2389	0.494	0.996	0.996
304	74	165	0.448	0.911	0.924	640	1709	0.374	0.996	0.996	714	1874	0.381	0.995	0.995
401	1	14	0.071	0.464	*	21	127	0.165	0.944	0.947	22	141	0.156	0.933	0.933
402	28	129	0.217	0.889	0.905	358	1289	0.278	0.994	0.994	386	1418	0.272	0.993	0.993
403	6	19	0.316	0.540	*	5	281	0.018	0.974	0.975	11	300	0.037	0.967	0.967
404	12	42	0.286	0.722	*	195	407	0.479	0.982	0.983	207	449	0.461	0.978	0.978
405	13	53	0.245	0.766	*	203	460	0.441	0.984	0.985	216	513	0.421	0.981	0.981
501	6	23	0.261	0.587	*	103	540	0.191	0.986	0.987	109	563	0.194	0.982	0.982
502	43	166	0.259	0.911	0.925	1	4	0.250	0.345	*	44	170	0.259	0.944	0.944
503	37	83	0.446	0.837	0.860	324	684	0.474	0.989	0.99	361	767	0.471	0.987	0.987
504	25	48	0.521	0.748	*	283	677	0.418	0.989	0.99	308	725	0.425	0.986	0.986
601	27	67	0.403	0.805	*	216	902	0.239	0.992	0.992	243	969	0.251	0.990	0.990
602	48	109	0.440	0.871	0.890	517	965	0.536	0.992	0.993	565	1074	0.526	0.991	0.991
603	20	54	0.370	0.769	*	267	836	0.319	0.991	0.992	287	890	0.322	0.989	0.989
701	53	81	0.654	0.834	0.857	339	581	0.583	0.987	0.988	392	662	0.592	0.985	0.985
801	230	710	0.324	0.978	0.981	3	6	0.500	0.441	*	233	716	0.325	0.986	0.986
802	236	518	0.456	0.970	0.975	0	0	*	*	*	236	518	0.456	0.981	0.981
803	23	93	0.247	0.852	0.873	0	0	*	*	*	23	93	0.247	0.902	0.902
804	60	161	0.373	0.909	0.922	1	1	1.000	0.116	*	61	162	0.377	0.941	0.941
805	19	87	0.218	0.843	0.865	0	0	*	*	*	19	87	0.218	0.896	0.896
806	22	106	0.208	0.868	0.887	0	0	*	*	*	22	106	0.208	0.913	0.913

Facility ID	MEM_Method Most: 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability (all units))	MEM_Method Most: 15 to <21 Years (Reliability (unit size≥75))	MEM_MethodMost: 21 to 44 years (Most/Mod Provision)	MEM_MethodMost: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability (all units))	MEM_Method Most: 21 to 44 years (Reliability (unit size≥75))	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (Reliability (all units))	MEM_MethodMost: all age groups (Reliability (unit size≥75))
Total or Mean	**2409**	**5705**	**0.422**	*	*	**10876**	**25379**	**0.429**	*	*	**13285**	**31084**	**0.427**	*	*
*	*	*	*	**Overall Reliability**	**Overall Reliability**	*	*	*	**Overall Reliability**	**Overall Reliability**	*	*	*	**Overall Reliability**	**Overall Reliability**
*	Median n	87.5	*	**0.770**	**0.916**	Median n	560.5	*	**0.869**	**0.985**	Median n	597	*	**0.970**	**0.970**
*	Min n	2	*	*	*	Min n	0	*	*	*	Min n	87	*	*	*

*cell intentionally left blank

Table 2. Rates and reliabilities for most or moderately effective contraceptive method provision by facility, 56 PPFA facilities, 2019.

Facility ID	MEM_MethodMost: 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability (all units))	MEM_Method Most: 15 to <21 Years (Reliability (unit size≥75))	MEM_Method Most: 21 to 44 years (Most/Mod Provision)	MEM_Method Most: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability (all units))	MEM_Method Most: 21 to 44 years (Reliability (unit size≥75))	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (Reliability (all units))	MEM_MethodMost: all age groups (Reliability (unit size≥75))
1	312	422	0.739	0.963	0.947	626	1083	0.578	0.987	0.984	938	1505	0.623	0.992	0.990
2	514	737	0.697	0.979	0.969	2895	4999	0.579	0.997	0.996	3409	5736	0.594	0.998	0.997
3	523	747	0.700	0.979	0.970	1482	2605	0.569	0.995	0.993	2005	3352	0.598	0.996	0.995
4	653	884	0.739	0.982	0.974	2157	3529	0.611	0.996	0.995	2810	4413	0.637	0.997	0.997
5	875	1213	0.721	0.987	0.981	2368	4002	0.592	0.996	0.996	3243	5215	0.622	0.998	0.997
6	611	868	0.704	0.982	0.974	1706	2923	0.584	0.995	0.994	2317	3791	0.611	0.997	0.996
7	535	759	0.705	0.979	0.970	1743	2951	0.591	0.995	0.994	2278	3710	0.614	0.997	0.996
8	451	587	0.768	0.973	0.962	945	1612	0.586	0.991	0.989	1396	2199	0.635	0.995	0.993
9	1049	1400	0.749	0.989	0.984	2668	4253	0.627	0.997	0.996	3717	5653	0.658	0.998	0.997
10	638	812	0.786	0.981	0.972	1603	2541	0.631	0.994	0.993	2241	3353	0.668	0.996	0.995
11	180	225	0.800	0.933	0.906	476	668	0.713	0.979	0.974	656	893	0.735	0.987	0.983
12	559	721	0.775	0.978	0.969	1445	2222	0.650	0.994	0.992	2004	2943	0.681	0.996	0.995
13	457	803	0.569	0.980	0.972	2240	3489	0.642	0.996	0.995	2697	4292	0.628	0.997	0.996
14	667	886	0.753	0.982	0.974	1862	3136	0.594	0.995	0.994	2529	4022	0.629	0.997	0.996
15	605	781	0.775	0.980	0.971	1941	3016	0.644	0.995	0.994	2546	3797	0.671	0.997	0.996
16	132	167	0.79	0.912	0.877	291	438	0.664	0.968	0.961	423	605	0.699	0.981	0.975
17	161	220	0.732	0.932	0.904	442	719	0.615	0.980	0.976	603	939	0.642	0.987	0.984

Facility ID	MEM_MethodMost 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability (all units))	MEM_Method Most: 15 to <21 Years (Reliability (unit size≥75))	MEM_Method Most: 21 to 44 years (Most/Mod Provision)	MEM_Method Most: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability (all units))	MEM_Method Most: 21 to 44 years (Reliability (unit size≥75))	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (Reliability (all units))	MEM_MethodMost: all age groups (Reliability (unit size≥75))
18	521	769	0.678	0.980	0.970	1407	2514	0.560	0.994	0.993	1928	3283	0.587	0.996	0.995
19	691	979	0.706	0.984	0.977	1416	2481	0.571	0.994	0.993	2107	3460	0.609	0.997	0.996
20	219	306	0.716	0.950	0.929	433	724	0.598	0.981	0.976	652	1030	0.633	0.988	0.985
21	79	99	0.798	0.860	0.809	284	490	0.580	0.972	0.965	363	589	0.616	0.980	0.974
22	392	604	0.649	0.974	0.963	2268	4050	0.560	0.996	0.996	2660	4654	0.572	0.997	0.997
23	304	373	0.815	0.959	0.941	543	926	0.586	0.985	0.981	847	1299	0.652	0.991	0.988
24	155	236	0.657	0.936	0.910	802	1401	0.572	0.990	0.987	957	1637	0.585	0.993	0.991
25	584	833	0.701	0.981	0.973	1234	2136	0.578	0.993	0.992	1818	2969	0.612	0.996	0.995
26	463	612	0.757	0.974	0.963	1717	2512	0.684	0.994	0.993	2180	3124	0.698	0.996	0.995
27	315	510	0.618	0.969	0.956	1475	2594	0.569	0.994	0.993	1790	3104	0.577	0.996	0.995
28	528	772	0.684	0.980	0.971	1757	3128	0.562	0.995	0.994	2285	3900	0.586	0.997	0.996
29	192	268	0.716	0.943	0.920	481	989	0.486	0.986	0.982	673	1257	0.535	0.991	0.988
30	233	285	0.818	0.947	0.924	851	1053	0.808	0.987	0.983	1084	1338	0.810	0.991	0.989
31	167	202	0.827	0.926	0.896	450	573	0.785	0.976	0.970	617	775	0.796	0.985	0.980
32	239	340	0.703	0.955	0.936	369	615	0.600	0.977	0.972	608	955	0.637	0.988	0.984
33	533	730	0.730	0.978	0.969	961	1581	0.608	0.991	0.989	1494	2311	0.646	0.995	0.993
34	262	358	0.732	0.957	0.939	418	705	0.593	0.980	0.975	680	1063	0.640	0.989	0.986
35	355	478	0.743	0.967	0.953	622	1133	0.549	0.987	0.984	977	1611	0.606	0.993	0.991
36	1	11	0.091	0.406	*	5	48	0.104	0.770	*	6	59	0.102	0.831	*
37	616	774	0.796	0.980	0.971	875	1493	0.586	0.990	0.988	1491	2267	0.658	0.995	0.993

*cell intentionally left blank

Table 2. Rates and reliabilities for most or moderately effective contraceptive method provision by facility, 56 PPFA facilities, 2019. (cont.)

Facility ID	MEM_Method Most: 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability (all units))	MEM_Method Most: 15 to <21 Years (Reliability (unit size≥75))	MEM_Method Most: 21 to 44 years (Most/Mod Provision)	MEM_Method Most: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability (all units))	MEM_Method Most: 21 to 44 years (Reliability (unit size≥75))	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_MethodMost: all age groups (Reliability (all units))	MEM_MethodMost: all age groups (Reliability (unit size≥75))
38	255	340	0.750	0.955	0.936	606	918	0.660	0.985	0.981	861	1258	0.684	0.991	0.988
39	476	676	0.704	0.977	0.967	1495	2739	0.546	0.995	0.993	1971	3415	0.577	0.996	0.996

Facility ID	MEM_Method Most: 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability (all units))	MEM_Method Most: 15 to <21 Years (Reliability (unit size≥75))	MEM_Method Most: 21 to 44 years (Most/Mod Provision)	MEM_Method Most: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability (all units))	MEM_Method Most: 21 to 44 years (Reliability (unit size≥75))	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_MethodMost: all age groups (Reliability (all units))	MEM_MethodMost: all age groups (Reliability (unit size≥75))
40	381	518	0.736	0.970	0.957	1468	2397	0.612	0.994	0.993	1849	2915	0.634	0.996	0.995
41	274	327	0.838	0.953	0.933	502	713	0.704	0.980	0.975	776	1040	0.746	0.989	0.985
42	612	823	0.744	0.981	0.972	1264	2180	0.580	0.993	0.992	1876	3003	0.625	0.996	0.995
43	832	1258	0.661	0.987	0.982	1759	3368	0.522	0.996	0.995	2591	4626	0.560	0.997	0.997
44	0	0	*	*	*	0	1	0	0.065	*	0	1	0	0.077	*
45	524	800	0.655	0.980	0.972	1475	2843	0.519	0.995	0.994	1999	3643	0.549	0.997	0.996
46	927	1231	0.753	0.987	0.981	1705	2840	0.600	0.995	0.994	2632	4071	0.647	0.997	0.996
47	135	452	0.299	0.966	0.951	716	2441	0.293	0.994	0.993	851	2893	0.294	0.996	0.995
48	281	437	0.643	0.964	0.949	1097	1967	0.558	0.993	0.991	1378	2404	0.573	0.995	0.994
49	724	939	0.771	0.983	0.976	2097	3332	0.629	0.996	0.995	2821	4271	0.661	0.997	0.996
50	633	942	0.672	0.983	0.976	1290	2257	0.572	0.994	0.992	1923	3199	0.601	0.996	0.995
51	813	1040	0.782	0.985	0.978	2295	3432	0.669	0.996	0.995	3108	4472	0.695	0.997	0.997
52	84	288	0.292	0.947	0.925	499	1899	0.263	0.992	0.991	583	2187	0.267	0.995	0.993
53	196	342	0.573	0.955	0.936	778	1646	0.473	0.991	0.989	974	1988	0.490	0.994	0.992
54	0	16	0	0.499	*	2	79	0.025	0.846	0.815	2	95	0.021	0.888	0.860
55	55	59	0.932	0.786	*	159	213	0.746	0.937	0.922	214	272	0.787	0.958	0.946
56	150	199	0.754	0.925	0.895	282	463	0.609	0.970	0.963	432	662	0.653	0.982	0.977
Total or Mean	**23123**	**32458**	**0.712**	*	*	**64747**	**111060**	**0.583**	*	*	**87870**	**143518**	**0.612**	*	*
*	*	*	*	**Overall Reliability**	**Overall Reliability**	*	*	*	**Overall Reliability**	**Overall Reliability**	*	*	*	**Overall Reliability**	**Overall Reliability**
*	Median n	604	*	**0.943**	**0.951**	Median n	2180	*	**0.966**	**0.983**	Median n	2915	*	**0.972**	**0.989**
*	Min n	0	*	*	*	Min n	1	*	*	*	Min n	1	*	*	*

*cell intentionally left blank

Table 3. Rates and reliabilities for most or moderately effective contraceptive method provision by public health region, Iowa Medicaid Enterprise, 2018.

Public Health Region	MEM_Method Most: 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability)	MEM_Method Most: 21 to 44 years (Most/Mod Provision)	MEM_Method Most: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability)	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (Reliability)
1	2772	8365	0.331	0.982	6962	25070	0.278	0.965	9734	33435	0.291	0.982
2	963	2247	0.429	0.935	2042	6392	0.319	0.875	3005	8639	0.348	0.934
3	1057	3183	0.332	0.953	2429	8915	0.272	0.907	3486	12098	0.288	0.952
4	1203	2824	0.426	0.948	2252	8191	0.275	0.900	3455	11015	0.314	0.948
5	1432	3609	0.397	0.959	3239	11346	0.285	0.926	4671	14955	0.312	0.961
6	3117	8409	0.371	0.982	8407	28341	0.297	0.969	11524	36750	0.314	0.984
Total or Mean	**10544**	**28637**	**0.368**	*	**25331**	**88255**	**0.287**	*	**35875**	**116892**	**0.307**	*
*	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**
*	Median n	3609	*	**0.960**	Median n	11346	*	**0.924**	Median n	14955	*	**0.960**
*	Min n	2247	*	*	Min n	6392	*	*	Min n	8639	*	*

Note: Reliability estimates are the same regardless of using the unit size cutoff of 75 because all unit sizes are above 75.

*cell intentionally left blank

Table 4. Rates and reliabilities for most or moderately effective contraceptive method provision by health plan, Washington State Health Care Authority, 2019.

Health Plan	MEM_Method Most: 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability)	MEM_Method Most: 21 to 44 years (Most/Mod Provision)	MEM_Method Most: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability)	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (Reliability)
MCO 1	1288	4031	0.320	0.852	4077	15357	0.265	0.913	5365	19388	0.277	0.934
MCO 2	2708	9684	0.280	0.932	5796	20378	0.284	0.933	8504	30062	0.283	0.957
MCO 3	2520	7731	0.326	0.917	4505	15127	0.298	0.912	7025	22858	0.307	0.944
MCO 4	10373	31628	0.328	0.978	21410	73240	0.292	0.980	31783	104868	0.303	0.987
MCO 5	1330	4281	0.311	0.859	4097	15111	0.271	0.911	5427	19392	0.280	0.934
Total or Mean	**18219**	**57355**	**0.318**	*	**39885**	**139213**	**0.287**	*	**58104**	**196568**	**0.296**	*
*	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**
*	Median n	8707.5	*	**0.908**	Median n	17867.5	*	**0.930**	Median n	26460	*	**0.951**
*	Min n	4031	*	*	Min n	15111	*	*	Min n	19388	*	*

Note: Reliability estimates are the same regardless of using the unit size cutoff of 75 because all unit sizes are above 75.

*cell intentionally left blank

Table 5. Rates and reliabilities for most or moderately effective contraceptive method provision by health plan, Massachusetts MassHealth, 2019.

Health Plan	MEM_MethodMost: 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability)	MEM_Method Most: 21 to 44 years (Most/Mod Provision)	MEM_Method Most: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability)	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (Reliability)
ACOA 1	108	1057	0.102	0.919	1096	5237	0.209	0.911	1204	6294	0.191	0.935
ACOA 2	127	445	0.285	0.827	746	2934	0.254	0.851	873	3379	0.258	0.886
ACOA 3	102	634	0.161	0.872	832	3201	0.260	0.862	934	3835	0.244	0.898
ACOA 4	1644	8036	0.205	0.989	268	858	0.312	0.626	1912	8894	0.215	0.953
ACOA 5	343	1444	0.238	0.939	1035	3944	0.262	0.885	1378	5388	0.256	0.925
ACOA 6	310	1895	0.164	0.953	1517	6354	0.239	0.925	1827	8249	0.221	0.950
ACOA 7	407	2270	0.179	0.961	1556	6408	0.243	0.926	1963	8678	0.226	0.952
ACOA 8	141	534	0.264	0.852	892	3446	0.259	0.870	1033	3980	0.260	0.902
ACOA 9	340	1557	0.218	0.944	1122	4541	0.247	0.898	1462	6098	0.240	0.933
ACOA 10	838	4720	0.178	0.981	4130	17779	0.232	0.972	4968	22499	0.221	0.981
ACOA 11	385	1809	0.213	0.951	1068	4305	0.248	0.894	1453	6114	0.238	0.934
ACOA 12	239	1073	0.223	0.920	1431	5403	0.265	0.913	1670	6476	0.258	0.937
ACOA 13	162	1235	0.131	0.930	642	3136	0.205	0.859	804	4371	0.184	0.910
ACOB 1	920	4690	0.196	0.981	3677	14398	0.255	0.966	4597	19088	0.241	0.978
ACOB 2	1109	5243	0.212	0.983	3667	15009	0.244	0.967	4776	20252	0.236	0.979
ACOB 3	947	5450	0.174	0.983	3913	16436	0.238	0.970	4860	21886	0.222	0.981
Non-ACO 1	0	0	*	*	718	2884	0.249	0.849	718	2884	0.249	0.869
Non-ACO 2	426	1909	0.223	0.954	2456	10238	0.240	0.952	2882	12147	0.237	0.965
Non-ACO 3	1288	5668	0.227	0.984	3034	11833	0.256	0.958	4322	17501	0.247	0.976
Non-ACO 4	213	1265	0.168	0.931	1661	7900	0.210	0.939	1874	9165	0.204	0.955
Non-ACO 5	0	0	*	*	82	351	0.234	0.406	82	351	0.234	0.447
Total or Mean	**10049**	**50934**	**0.197**	*	**35543**	**146595**	**0.242**	*	**45592**	**197529**	**0.231**	*
*	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**
*	Median n	1683	*	**0.940**	Median n	5320	*	**0.876**	Median n	7362.5	*	**0.916**
*	Min n	0	*	*	Min n	351	*	*	Min n	351	*	*

Note: Reliability estimates are the same regardless of using the unit size cutoff of 75 because all unit sizes are above 75.

*cell intentionally left blank

Table 6. Rates and reliabilities for most or moderately effective contraceptive method provision by health plan, Louisiana Medicaid, 2019.

Health Plan	MEM_Method Most: 15 to <21 Years (Most/Mod Provision)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (Reliability)	MEM_Method Most: 21 to 44 years (Most/Mod Provision)	MEM_Method Most: 21 to 44 years (Total N)	MEM_Method Most: 21 to 44 years (Rate)	MEM_Method Most: 21 to 44 years (Reliability)	MEM_Method Most: all age groups (Most/Mod Provision)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (Reliability)
MCO 1	1004	3004	0.334	0.633	4266	15174	0.281	0.859	5270	18178	0.290	0.927
MCO 2	3387	10115	0.335	0.853	7909	27867	0.284	0.918	11296	37982	0.297	0.964
MCO 3	4493	12636	0.356	0.879	10977	37620	0.292	0.938	15470	50256	0.308	0.972
MCO 4	10653	29880	0.357	0.945	18553	61423	0.302	0.961	29206	91303	0.320	0.985
MCO 5	8317	22699	0.366	0.929	17906	58682	0.305	0.959	26223	81381	0.322	0.983
Total or Mean	**27854**	**78334**	**0.356**	*	**59611**	**200766**	**0.297**	*	**87465**	**279100**	**0.313**	*
*	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**
*	Median n	17667.5	*	**0.848**	Median n	48151	*	**0.927**	Median n	65818.5	*	**0.966**
*	Min n	3004	*	*	Min n	15174	*	*	Min n	18178	*	*

Note: Reliability estimates are the same regardless of using the unit size cutoff of 75 because all unit sizes are above 75.

*cell intentionally left blank

Table 7. Correlations between most or moderately effective contraceptive method provision and selected related measures by group billing provider, Iowa Medicaid Enterprise, 2018.

Related measures	Age Group	Results (unit size≥25): Number of units in analysis	Results (unit size≥25): Pearson r	Results (unit size≥25): Multilevel correlation estimation Correlation coefficients	Results (unit size≥25): Multilevel correlation estimation (95% CL (lower, upper))	Results (unit size≥50): Number of units in analysis	Results (unit size≥50): Pearson r	Results (unit size≥50): Multilevel correlation estimation (Correlation coefficients)	Results (unit size≥50): Multilevel correlation estimation (95% CL (lower, upper))	Results (unit size≥75): Number of units in analysis	Results (unit size≥75): Pearson r	Results (unit size≥75): Multilevel correlation estimation (Correlation coefficients)	Results (unit size≥75): Multilevel correlation estimation (95% CL (lower, upper))
Contraceptive Counseling	15-44	633	**.47***	**.62***	**(.55, .68)**	393	**.53***	**.62***	**(.54, .69)**	270	**.56***	**.61***	**(.52, .68)**
- -	21-44	525	.46*	.61*	(.53, .68)	297	.53*	.61*	(.52, .69)	201	.54*	.58*	(.47, .67)
- -	15-20	202	.52*	.69*	(.57, .77)	96	.55*	.65*	(.49, .76)	56	.59*	.63*	(.42, .77)
Gynecological Examination	15-44	633	**.34***	**.39***	**(.30, .47)**	393	**.29***	**.32***	**(.21, .42)**	270	**.32***	**.29***	**(.15, .41)**
- -	21-44	525	.32*	.37*	(.27, .47)	297	.29*	.31*	(.18, .42)	201	.32*	.29*	(.14, .42)
- -	15-20	202	.36*	.49*	(.31, .64)	96	.39*	.43*	(.19, .62)	56	.42*	.42*	(.06, .66)
Cervical Cancer Screening†	21-44	523	**.33***	**.42***	**(.33, .51)**	296	**.29***	**.33***	**(.20, .45)**	198	**.30***	**.31***	**(.15, .44)**
Chlamydia Screening†	16-24	186	**.00**	**.00**	**(-.17, .18)**	87	**.03**	**.00**	**(-.22, .23)**	53	**.23**	**.21**	**(-.09, .46)**
- -	21-24	82	.06	.11	(-.16, .36)	40	.26	.24	(-.11, .52)	24	.29	.22	(-.21, .55)
- -	16-20	99	.02	.03	(-.19, .25)	40	.23	.20	(-.14, .49)	27	.29	.27	(-.16, .59)

*statistically significant at p < .05
†Age range of the related measure differs from that of the contraceptive care measure and the analysis was conducted among the overlapping population only
- - cell intentionally left blank

2016 Submission (All following tables are from 2016 submission)

Table 1. Rates and reliabilities for moderate or most effective contraceptive method, 25 PPFA affiliates, 2014.

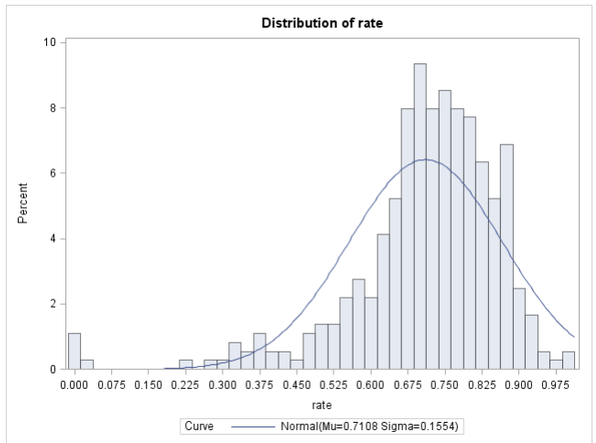
Affiliate ID	*	MEM_Method Most: 15 to <21 Years (Used Most/Mod)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_MethodMost: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (HC Within Affiliate Reliability)	*	MEM_Method Most: 21 to 45 years (Used Most/Mod)	MEM_Method Most: 21 to 45 years (Total N)	MEM_Method Most: 21 to 45 years (Rate)	MEM_Method Most: 21 to 45 years (HC Within Affiliate Reliability)	*	MEM_Method Most: all age groups (Used Most/Mod)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (HC Within Affiliate Reliability)
1	*	6765	7869	0.860	**0.9822**	*	23438	29638	0.791	**0.9990**	*	30203	37507	0.805	**0.9993**
3	*	19611	26591	0.738	**0.9850**	*	63945	88881	0.719	**0.9984**	*	83556	115472	0.724	**0.9981**
4	*	2713	4147	0.654	**0.6096**	*	12900	21430	0.602	**0.9597**	*	15613	25577	0.610	**0.9630**
5	*	33452	42698	0.783	**0.9910**	*	92727	131187	0.707	**0.9955**	*	126179	173885	0.726	**0.9971**
6	*	1710	2651	0.645	**0.8348**	*	4390	7362	0.596	**0.9469**	*	6100	10013	0.609	**0.9563**
9	*	17261	25268	0.683	**0.9852**	*	58446	88455	0.661	**0.9936**	*	75707	113723	0.666	**0.9952**
10	*	11239	15188	0.740	**0.9090**	*	30853	47698	0.647	**0.9281**	*	42092	62886	0.669	**0.9637**
12	*	4355	4839	0.900	**0.8011**	*	8599	10209	0.842	**0.9419**	*	12954	15048	0.861	**0.9604**
37	*	1318	1965	0.671	**0.6152**	*	2187	4194	0.521	**0.6709**	*	3505	6159	0.569	**0.7056**
38	*	4071	6093	0.668	**0.9502**	*	5683	10645	0.534	**0.9529**	*	9754	16738	0.583	**0.9743**
40	*	3628	5030	0.721	**0.9150**	*	5991	10843	0.553	**0.9264**	*	9619	15873	0.606	**0.9544**
41	*	4130	5466	0.756	**0.9512**	*	11021	17562	0.628	**0.8917**	*	15151	23028	0.658	**0.9620**
44	*	9903	11489	0.862	**0.9647**	*	27664	33620	0.823	**0.9848**	*	37567	45109	0.833	**0.9879**
47	*	4383	5644	0.777	**0.9793**	*	10699	16648	0.643	**0.9899**	*	15082	22292	0.677	**0.9946**
53	*	5136	8741	0.588	**0.9696**	*	12678	28791	0.440	**0.9691**	*	17814	37532	0.475	**0.9850**
54	*	2277	3122	0.729	**0.8237**	*	3636	6614	0.550	**0.9522**	*	5913	9736	0.607	**0.9765**
59	*	2851	3682	0.774	**0.9903**	*	6470	9778	0.662	**0.9991**	*	9321	13460	0.692	**0.9994**
60	*	258	436	0.592	**0.0000**	*	599	1265	0.474	**0.0000**	*	857	1701	0.504	**0.0000**
70	*	1944	4154	0.468	**0.9985**	*	3507	12436	0.282	**0.9996**	*	5451	16590	0.329	**0.9997**
73	*	406	996	0.408	**0.9863**	*	789	2825	0.279	**0.9874**	*	1195	3821	0.313	**0.9933**
75	*	431	1171	0.368	**0.9828**	*	1605	5070	0.317	**0.9946**	*	2036	6241	0.326	**0.9958**
76	*	1737	3817	0.455	**0.9368**	*	4819	11648	0.414	**0.8733**	*	6556	15465	0.424	**0.9425**
77	*	8635	11359	0.760	**0.9693**	*	21324	31393	0.679	**0.9655**	*	29959	42752	0.701	**0.9833**
79	*	721	1260	0.572	**0.8878**	*	2162	5149	0.420	**0.9440**	*	2883	6409	0.450	**0.9625**
81	*	182	294	0.619	**0.0000**	*	806	1561	0.516	**0.0000**	*	988	1855	0.533	**0.0000**
Total or Mean	*	**149117**	**203970**	**0.731**	*	*	**416938**	**634902**	**0.657**	*	*	**566055**	**838872**	**0.675**	*
*	*	*	**σ Level 2**	**ICC**	**Overall Affiliate Reliability**	*	*	**σ Level 2**	**ICC**	**Overall Affiliate Reliability**	*	*	**σ Level 2**	**ICC**	**Overall Affiliate Reliability**
Reliability using Median Affiliate Patient Volume	**Median n**	**4839**	**0.4334**	**0.1164**	**0.9984**	**Median n**	**11648**	**0.4624**	**0.1232**	**0.9994**	**Median n**	**16590**	**0.4448**	**0.1191**	**0.9996**

Affiliate ID	*	MEM_Method Most: 15 to <21 Years (Used Most/Mod)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_MethodMost: 15 to <21 Years (Rate)	MEM_Method Most: 15 to <21 Years (HC Within Affiliate Reliability)	*	MEM_Method Most: 21 to 45 years (Used Most/Mod)	MEM_Method Most: 21 to 45 years (Total N)	MEM_Method Most: 21 to 45 years (Rate)	MEM_Method Most: 21 to 45 years (HC Within Affiliate Reliability)	*	MEM_Method Most: all age groups (Used Most/Mod)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	MEM_Method Most: all age groups (HC Within Affiliate Reliability)
Reliability using Minimum Patient Volume (Floor)	Min n	294	0.4334	0.1164	0.9748	Min n	1265	0.4624	0.1232	0.9944	Min n	1701	0.4448	0.1191	0.9957

*cell intentionally left blank

Table 2. Distributions of rates and ICCs among health centers (n=363) for moderate/most effective methods among 25 PPFA affiliates, 2014

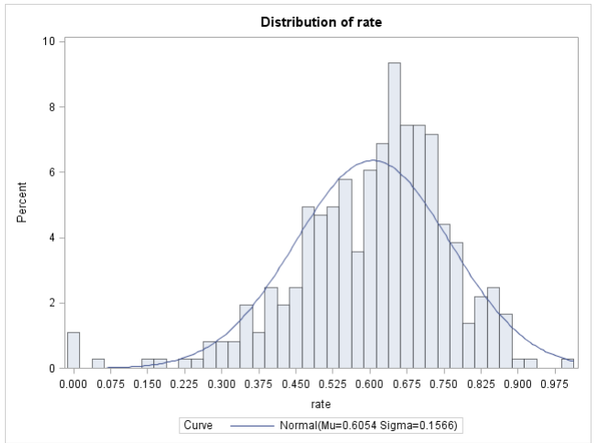
MEM_MethodMost: 15 to <21 Years



	cases (n)	rate
Mean	561.9	0.71
Median	366	0.73
SD	552.3	0.16
Variance	305043	0.02
Range	2976	1.00
Interquartile	629	0.15
Median ICC		0.06

Quantile	cases (n)	rate
100% Max	2984	1.00

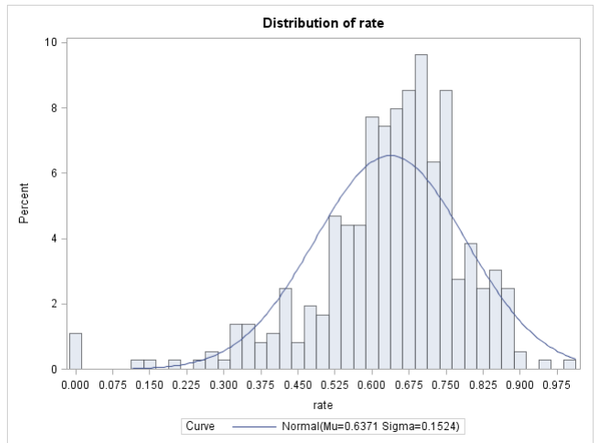
MEM_MethodMost: 21 to 45 years



	cases (n)	rate
Mean	1749.0	0.61
Median	1016	0.63
SD	1909	0.16
Variance	3645550	0.02
Range	11360	1.00
Interquartile	2145	0.19
Median ICC		0.05

Quantile	cases (n)	rate
100% Max	11391	1.00

MEM_MethodMost: all age groups



	cases (n)	rate
Mean	2310.9	0.64
Median	1379	0.66
SD	2424	0.15
Variance	5875321	0.02
Range	13287	1.00
Interquartile	2757	0.17
Median ICC		0.06

Quantile	cases (n)	rate
100% Max	13335	1.00

95%	1766	0.90	95%	5489	0.83	95%	7198	0.85
90%	1410	0.87	90%	4544	0.78	90%	5872	0.80
75% Q3	787	0.81	75% Q3	2516	0.71	75% Q3	3315	0.73
50% Med	366	0.73	50% Med	1016	0.63	50% Med	1379	0.66
25% Q1	158	0.66	25% Q1	371	0.52	25% Q1	558	0.57
10%	83	0.54	10%	149	0.41	10%	240	0.44
5%	53	0.41	5%	92	0.34	5%	141	0.35
0% Min	8	0.00	0% Min	31	0.00	0% Min	48	0.00

Table 3. Rates and reliabilities for moderate or most effective contraceptive method, Iowa Medicaid Enterprise, 2013, by region and type of benefit

Public Health Region	MEM_MethodMost: 15 to <21 Years (Not Used)	MEM_Method Most: 15 to <21 Years (Used Most/Mod)	MEM_Method Most: 15 to <21 Years (Total N)	MEM_Method Most: 15 to <21 Years (Rate)	*	*	MEM_Method Most: 21 to 45 years (Not Used)	MEM_Method Most: 21 to 45 years (Used Most/Mod)	MEM_Method Most: 21 to 45 years (Total N)	MEM_Method Most: 21 to 45 years (Rate)	*	*	MEM_Method Most: all age groups (Not Used)	MEM_Metho Most: all age groups (Use Most/Mod)	MEM_Method Most: all age groups (Total N)	MEM_Method Most: all age groups (Rate)	*	*
1	1384	2076	3460	0.600	*	*	3750	5838	9588	0.609	*	*	5134	7914	13048	0.607	*	*
2	380	774	1154	0.671	*	*	1013	1893	2906	0.651	*	*	1393	2667	4060	0.657	*	*
3	565	611	1176	0.520	*	*	1261	1914	3175	0.603	*	*	1826	2525	4351	0.580	*	*
4	415	672	1087	0.618	*	*	1262	1625	2887	0.563	*	*	1677	2297	3974	0.578	*	*
5	572	1129	1701	0.664	*	*	1890	2469	4359	0.566	*	*	2462	3598	6060	0.594	*	*
6	1148	1973	3121	0.632	*	*	4079	6057	10136	0.598	*	*	5227	8030	13257	0.606	*	*
Total or Mean	**4464**	**7235**	**11699**	**0.618**	*	*	**13255**	**19796**	**33051**	**0.599**	*	*	**17719**	**27031**	**44750**	**0.604**	*	*
*	*	*	*	**VarL1**	**ICC**	Region Reliability (Var L1)	*	*	*	**VarL1**	**ICC**	Region Reliability (Var L1)	*	*	*	**VarL1**	**ICC**	Region Reliability (Var L1)
Median Patient Volume Among Affiliates	*	Median n	1438.5	0.04018	0.0121	**0.9461**	*	Median n	3767	0.01365	0.0041	**0.9399**	*	Median n	5205.5	0.01109	0.0034	**0.9461**
Minimum Patient Volume (Floor)	*	Min n	1087	0.04018	0.0121	0.9299	*	Min n	2887	0.01365	0.0041	0.9229	*	Min n	3974	0.01109	0.0034	0.9305
Type of benefit (family planning waiver vs general Medicaid)	**MEM_Method Most: 15 to <21 Years** (Not Used)	**MEM_Method Most: 15 to <21 Years** (Used Most/Mod)	**MEM_Method Most: 15 to <21 Years** (Total N)	**MEM_Method Most: 15 to <21 Years** (Rate)	*	*	**MEM_Method Most: 21 to 45 years** (Not Used)	**MEM_Method Most: 21 to 45 years** (Used Most/Mod)	**MEM_Method Most: 21 to 45 years** (Total N)	**MEM_Method Most: 21 to 45 years** (Rate)	*	*	**MEM_Method Most: all age groups** (Not Used)	**MEM_Method Most: all age groups** (Used Most/Mod)	**MEM_Method Most: all age groups** (Total N)	**MEM_Method Most: all age groups** (Rate)	*	*
Family Planning Waiver	1333	5112	6445	0.793	*	*	6430	17138	23568	0.727	*	*	7763	22250	30013	0.741	*	*
General Medicaid	3131	2123	5254	0.404	*	*	6825	2658	9483	0.280	*	*	9956	4781	14737	0.324	*	*
Total or Mean	**4464**	**7235**	**11699**	**0.618**	*	*	**13255**	**19796**	**33051**	**0.599**	*	*	**17719**	**27031**	**44750**	**0.604**	*	*
*	*	*	*	**VarL2**	**ICC**	Type of Benefit Reliability (Var L2)	*	*	*	**VarL2**	**ICC**	Type of Benefit Reliability (Var L2)	*	*	*	**VarL2**	**ICC**	Type of Benefit Reliability (Var L2)
Reliability Based on Median Patient Volume Among Health Centers	*	Median n	5849.5	0.4778	0.1268	**0.9988**	*	Median n	16525.5	0.01878	0.0057	**0.9895**	*	Median n	22375	0.1598	0.0463	**0.9991**
Calculated Based on Minimum Patient Volume (Floor)	*	Min n	5254	0.4778	0.1268	0.9987	*	Min n	9483	0.01878	0.0057	0.9819	*	Min n	4444	0.1598	0.0463	0.9986
Region 1/Family Planning Waiver	941	595	1536	0.387	*	*	1801	743	2544	0.292	*	*	2742	1338	4080	0.328	*	*
Region 1/ General Medicaid	443	1481	1924	0.770	*	*	1949	5095	7044	0.723	*	*	2392	6576	8968	0.733	*	*
Region 2/Family Planning Waiver	271	192	463	0.415	*	*	477	178	655	0.272	*	*	748	370	1118	0.331	*	*
Region 2/ General Medicaid	109	582	691	0.842	*	*	536	1715	2251	0.762	*	*	645	2297	2942	0.781	*	*
Region 3/Family Planning Waiver	405	204	609	0.335	*	*	535	240	775	0.310	*	*	940	444	1384	0.321	*	*
Region 3/ General Medicaid	160	407	567	0.718	*	*	726	1674	2400	0.698	*	*	886	2081	2967	0.701	*	*
Region 4/Family Planning Waiver	297	239	536	0.446	*	*	640	261	901	0.290	*	*	937	500	1437	0.348	*	*
Region 4/ General Medicaid	118	433	551	0.786	*	*	622	1364	1986	0.687	*	*	740	1797	2537	0.708	*	*
Region 5/Family Planning Waiver	404	337	741	0.455	*	*	1089	408	1497	0.273	*	*	1493	745	2238	0.333	*	*
Region 5/ General Medicaid	168	792	960	0.825	*	*	801	2061	2862	0.720	*	*	969	2853	3822	0.746	*	*
Region 6/Family Planning Waiver	813	556	1369	0.406	*	*	2283	828	3111	0.266	*	*	3096	1384	4480	0.309	*	*
Region 6/ General Medicaid	335	1417	1752	0.809	*	*	1796	5229	7025	0.744	*	*	2131	6646	8777	0.757	*	*
Total or Mean	**4464**	**7235**	**11699**	**0.618**	*	*	**13255**	**19796**	**33051**	**0.599**	*	*	**17719**	**27031**	**44750**	**0.604**	*	*

Public Health Region	MEM_ Method Most: 15 to <21 Years (Not Used)	MEM_ Method Most: 15 to <21 Years (Used Most/Mod)	MEM_MethodMost: 15 to <21 Years (Total N)	MEM_MethodMost: 15 to <21 Years (Rate)	*	*	MEM_ Method Most: 21 to 45 years (Not Used)	MEM_ Method Most: 21 to 45 years (Used Most/Mod)	MEM_MethodMost: 21 to 45 years (Total N)	MEM_MethodMost: 21 to 45 years (Rate)	*	*	MEM_MethodMost: all age groups (Not Used)	MEM_MethodMost: all age groups (Used Most/ Mod)	MEM_MethodMost: all age groups (Total N)	MEM_MethodMost: all age groups (Rate)	*	*
*	*	*	*	**VarL2**	**ICC**	Type of Benefit by Public Health Region Reliability (Var L2)	*	*	*	**VarL2**	**ICC**	Type of Benefit by Public Health Region Reliability (Var L2)	*	*	*	**VarL2**	**ICC**	Type of Benefit by Public Health Region Reliability (Var L2)
Reliability Based on Median Patient Volume Among Health Centers	*	Median n	716	0.7862	0.1929	**0.9942**	*	Median n	2325.5	0.8998	0.2148	**0.9984**	*	Median n	2954.5	0.782	0.1920	**0.9986**
Calculated Based on Minimum Patient Volume (Floor)	*	Min n	463	0.7862	0.1929	0.9910	*	Min n	655	0.8998	0.2148	0.9944	*	Min n	1118	0.782	0.1920	0.9963

*cell intentionally left blank

Table 4. Rates and reliabilities for moderate or most effective contraceptive method, Wisconsin Medicaid, 2014, by health plan

HMO	MEM_MethodMost: 15 to <21 Years (Not Used)	MEM_MethodMost: 15 to <21 Years (Used Most/Mod)	MEM_MethodMost: 15 to <21 Years (Total N)	MEM_MethodMost: 15 to <21 Years (Rate)	*	*	MEM_MethodMost: 21 to 45 years (Not Used)	MEM_MethodMost: 21 to 45 years (Used Most/Mod)	MEM_MethodMost: 21 to 45 years (Total N)	MEM_MethodMost: 21 to 45 years (Rate)	*	*	MEM_MethodMost: all age groups (Not Used)	MEM_MethodMost: all age groups (Used Most/Mod)	MEM_MethodMost: all age groups (Total N)	MEM_MethodMost: all age groups (Rate)	*	*
1	2827	2005	4832	0.415	*	*	8462	5581	14043	0.397	*	*	11289	7586	18875	0.402	*	*
2	972	866	1838	0.471	*	*	3560	2128	5688	0.374	*	*	4532	2994	7526	0.398	*	*
3	490	430	920	0.467	*	*	1852	1010	2862	0.353	*	*	2342	1440	3782	0.381	*	*
4	924	871	1795	0.485	*	*	3471	2210	5681	0.389	*	*	4395	3081	7476	0.412	*	*
5	649	582	1231	0.473	*	*	2614	1322	3936	0.336	*	*	3263	1904	5167	0.368	*	*
6	113	106	219	0.484	*	*	451	274	725	0.378	*	*	564	380	944	0.403	*	*
7	281	277	558	0.496	*	*	1032	576	1608	0.358	*	*	1313	853	2166	0.394	*	*
8	178	174	352	0.494	*	*	663	433	1096	0.395	*	*	841	607	1448	0.419	*	*
9	930	693	1623	0.427	*	*	4019	2145	6164	0.348	*	*	4949	2838	7787	0.364	*	*
10	331	287	618	0.464	*	*	1023	660	1683	0.392	*	*	1354	947	2301	0.412	*	*
11	2656	2242	4898	0.458	*	*	9082	6084	15166	0.401	*	*	11738	8326	20064	0.415	*	*
12	660	579	1239	0.467	*	*	2776	1514	4290	0.353	*	*	3436	2093	5529	0.379	*	*
13	146	123	269	0.457	*	*	517	336	853	0.394	*	*	663	459	1122	0.409	*	*
14	1158	991	2149	0.461	*	*	3692	1904	5596	0.340	*	*	4850	2895	7745	0.374	*	*
15	29	27	56	0.482	*	*	158	82	240	0.342	*	*	187	109	296	0.368	*	*
16	2748	2366	5114	0.463	*	*	11763	7112	18875	0.377	*	*	14511	9478	23989	0.395	*	*
17	283	276	559	0.494	*	*	879	654	1533	0.427	*	*	1162	930	2092	0.445	*	*
Total or Mean	**15375**	**12895**	**28270**	**0.456**	*	*	**56014**	**34025**	**90039**	**0.378**	*	*	**71389**	**46920**	**118309**	**0.397**	*	*
*	*	*	*	**VarL1**	**ICC**	Overall HMO Reliability (Var L1)	*	*	*	**VarL1**	**ICC**	Overall HMO Reliability (Var L1)	*	*	*	**VarL1**	**ICC**	Overall HMO Reliability (Var L1)
Median Patient Volume Among Affiliates	*	Median n	1231	0.005593	0.0017	**0.6767**	*	Median n	3936	0.009698	0.0029	**0.9206**	*	Median n	5167	0.006053	0.0018	**0.9048**
Minimum Patient Volume (Floor)	*	Min n	56	0.005593	0.0017	0.0869	*	Min n	240	0.009698	0.0029	0.4143	*	Min n	296	0.006053	0.0018	0.3526

*cell intentionally left blank

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.

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

If other:

3b. Electronic Sources

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

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

ALL data elements are in defined fields in electronic claims

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

In 2019, OPA funded the University of California San Francisco (UCSF) to develop and submit to NQF for endorsement an eMeasure (aka eCQM) for the provision of most and moderately effective contraceptive methods. The goal of this collaboration is to enhance the quality of contraceptive services, particularly in underserved populations through widespread use of validated performance measures for contraceptive care. These contraceptive eCQMs would be disseminated and utilized in diverse health care settings, including Community Health Centers (CHCs), and calculated alongside the NQF-endorsed Person-Centered Contraceptive Counseling (PCCC) measure (NQF #3543). Building upon previous work completed by OPA, UCSF's project team is refining the specifications of an eCQM version of this measure to utilize a new data element that enables patients to self-report their need for pregnancy prevention. Data collection for reliability and validity analyses required for submitting the eCQM for NQF endorsement is also underway.

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.

NQF #2903 was one of three contraceptive care measures included as part of the Centers for Medicaid & Medicare Services' (CMS) Maternal and Infant Health Initiative (MIHI). From 2015 to 2018, thirteen MIHI grantees tested and developed these first metrics for contraceptive care. NQF #2903 became publicly reported as part of CMS' Adult and Child Core Sets of Health Care Quality Measures in 2018. This allows states and territories access to the measure specifications, code sets, and technical assistance for calculation so that they can voluntarily submit their annual their measure scores to CMS. Overall, these experiences have confirmed that the measures can be feasibly calculated using existing claims data. As documented in an analytic brief (https://www.medicaid.gov/medicaid/quality-of-care/downloads/mihi-contraceptive-measures.pdf), several lessons learned from the CMS MIHI are summarized below:

OPA and MIHI grantees participated in a "co-design process" to develop and refine the measure specifications together, which furthered the collaborative learning process for the measure steward and users. The collaborative learning helped to expand the code sets used to define the numerator for NQF #2903, as several grantees shared the codes that they used for contraceptive care that were missing from the early specifications. OPA continues to ask states to share any additional administrative codes or state-specific policies they utilize for measure calculation. OPA then considers these codes for future measure updates. This is consistent with the approach used by NCQA for its Chlamydia Screening in Women measure for HEDIS (NQF #0033).

U.S. territories require technical assistance for NQF #2903 calculation specific to the unique features of their available data and health care delivery system. One MIHI grantee was a U.S. territory, and its analysis data included only LARC methods provided in the hospital plus a subset of most or moderately effective methods received in the public health clinics. As a result of missing contraceptive services data from private and public clinics, the grantee's reported rates were noticeably lower than the other MIHI grantees.

Since its NQF endorsement in 2016, NQF #2903 has implemented in other programmatic contexts besides Medicaid, including Title X Family Planning Program and the Planned Parenthood Federation of America. Regardless of setting, users have noted that the measure calculation is time-consuming and complex, even after the measure specification was simplified to no longer account for LARC removals. Furthermore, while OPA has provided a set of SAS programs to compute NQF #2903, this syntax can be difficult to troubleshoot and adapt across data systems. OPA provides technical assistance to users requesting clarification and help with the SAS programs. Some ask for assistance in revising programs customized to their computing environment and creating a dataset of women eligible to be included in the measure denominator, which can require customized coaching sessions. OPA plans to explore ways to improve the efficiency of the SAS syntax and other platforms for syntax.

Other measured entities indicated that barriers exist to access and understanding claims data for computing NQF #2903 measure scores. One state that already reports the measure to CMS had to complete a lengthy data user agreement process to gain access to Title X Family Planning Program data to monitor changes in NQF #2903 for a quality improvement initiative, only to find that some providers did not see many clients for contraceptive services. The initiative may have also been affected by concurrent statewide and provider-based initiatives to improve access to most and moderately effective methods and application for continued Title X funding.

Finally, existing administrative claims data has several known limitations in the measurement of unintended pregnancy. Claims data does not capture the client's history of sexual experience, their desire to become pregnant, or sterilization or LARC insertion prior to the measurement year; information about these patient

attributes can affect a client's decision to use contraception. Building upon a 2018 pilot conducted in partnership with CDC, OPA has funded the University of California San Francisco (UCSF) to develop an electronic clinical quality measure (eCQM) for the provision of most and moderately effective contraceptives. This new eCQM will utilize a new data element that enables patients to self-report their need for pregnancy prevention. Contraceptive eCQMs would be calculated alongside the NQF-endorsed Person-Centered Contraceptive Counseling (PCCC) measure (NQF #3543). Data collection for reliability and validity analyses required for submitting the eCQM for NQF endorsement is currently underway.

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)*.

Not applicable. The measure specifications, code lists, programming code and NSFG tables needed to interpret scores will all be available at no charge on the OPA website.

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 CMCS Core Set of Adult and Child Health Care Quality Measures https://www.medicaid.gov/medicaid/quality-of-care/performance-measurement/adult-and-child-health-care-quality-measures/index.html Quality Improvement (Internal to the specific organization) https://dhs.iowa.gov/ime/members/medicaid-a-to-z Iowa Medicaid Enterprise Louisiana Medicaid https://qualitydashboard.ldh.la.gov/ MassHealth https://www.mass.gov/orgs/masshealth NewYork-Presbyterian Hospital/Columbia University Irving Medical Center Ambulatory Care Network https://www.nyp.org/acn Washington State Health Care Authority https://www.hca.wa.gov/about-hca/reproductive-health Title X Family Planning Program https://rhntc.org/resources/contraceptive-access-change-package, Title X Family Planning Program https://opa.hhs.gov/evaluation-research/title-x-services-research/family-planning-annual-report

*cell intentionally left blank

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

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

NQF #2903 current use is presented for eight programs: federal Medicaid efforts to publicly report and support state use of the measures; four state Medicaid programs (i.e., the Iowa Medicaid Enterprise, the Washington State Health Care Authority, Louisiana Medicaid, and MassHealth); and one outpatient clinic network within an academic health system (NewYork-Presbyterian Hospital/Columbia University). We also include data from two national organizations that focus on the delivery of reproductive health services (i.e., the Planned Parenthood Federation of America and the Title X program).

1. Centers for Medicaid & Medicare Services (CMS): Maternal and Infant Health Initiative, Core Measure Set

CMS' Center for Medicaid and CHIP Services (CMCS) incorporated the contraceptive care measures into the publicly reported Core Set for Adult and Child Health Care Quality Measures, which evaluates quality of care accessed by over 73 million Medicaid and CHIP beneficiaries in the United States. NQF #2903 was added in 2018, which allows all 50 states to report the measure scores on a voluntary basis. While CMCS has collected NQF #2903 rates since 2015 from 13 Maternal and Infant Health Initiative (MIHI) grantees, it only releases yearly Adult and Child Core Set data for measures that were reported by at least 25 states and met its internal standards for data quality. For federal fiscal year (FFY) 2018, NQF #2903 met CMCS's threshold for public reporting of state-specific results, and thus CMS published these rates among ages 15-20 for 26 states for the first time (24 states reported the rates among ages 21-44). For FFY 2019, 28 states reported measure scores for ages 15-20 (23 states reported the rates among ages 21-44). Measure scores are calculated from inpatient, outpatient, and pharmacy administrative claims from facilities delivering primary care and reproductive health services. These scores are reported to CMCS at the state population level by age group, and some states compute and publish NQF #2903 by health plan. For more details on the CMCS's Core set, see: https://www.medicaid.gov/medicaid/quality-of-care/performance-measurement/adult-and-child-health-care-quality-measures/index.html.
The state agencies that administer Medicaid in Iowa, Louisiana, Massachusetts, and Washington report measure scores to CMCS and utilize NQF #2903 for internal quality improvement.

2. Iowa Medicaid Enterprise (IME)

Approximately 25% of Iowa's population in fiscal year (FY) 2020 is estimated to be served by IME, which provides contraceptive services to female Medicaid beneficiaries ages 15-44 residing in 99 counties and participating in either the general Medicaid program or the state-funded Family Planning Program (FPP). During FY 2019, Medicaid services in Iowa were provided primarily through two managed care organizations (MCOs), although a small percentage of clients (approximately 7%) were provided care on a fee-for-service basis. In partnership with CMCS MIHI grantee Iowa Department of Public Health, IME has annually calculated and publicly reported NQF #2903 for the past six years at the levels of state and public health region populations. Approximately 116,892 eligible women ages 15-44 were included in the measure denominator in 2018; in 2019, the number of women included was 110,218.

3. Louisiana Medicaid (LA Medicaid)

The 2019 LA Medicaid dataset included all female Medicaid enrollees aged 15-44 years who resided in 64 parishes. Almost 40% of Louisiana's population is enrolled in its Medicaid program, which provides contraceptive services to women through its general Medicaid program and its family planning state-plan amendment, Take Charge Plus (which is a different program than WA HCA's family planning waiver program). Services are available to uninsured Louisiana residents not eligible for Medicaid, Louisiana's CHIP program, or Medicare and who do not have private insurance. The guidelines for Take Charge Plus include women or men of any age with income at or below 138% of the federal poverty level. In 2019, Medicaid services in Louisiana (excluding Medicaid-Medicare dual-eligibles) were provided primarily by five managed care plans, which are

administered by the state's Healthy Louisiana program. Approximately 15% of the Medicaid population that is not dual-eligible was continuously enrolled in traditional fee-for-service Medicaid. Since 2017, LA Medicaid has calculated and publicly reported NQF #2903 by health plan via its Medicaid Quality Dashboard [1]. In 2019, about 279,100 eligible women ages 15-44 were included in the NQF #2903 denominator.

4. Massachusetts Medicaid (MassHealth)

In 2019, MassHealth delivered contraceptive services to female Medicaid clients aged 15-44 who resided in 14 counties and participated in 21 health plans. Sixteen of these health plans were managed care organizations. During fiscal year 2019, almost half of MassHealth's 1.8 million members are now enrolled in an accountable care organization (ACO); about 32% of clients receive care on a fee-for-service basis. Through the CMCS MIHI funding awarded to the Commonwealth of Massachusetts, MassHealth has annually calculated and reported NQF #2903 for the past six years for the state. In 2019, approximately 197,529 eligible women ages 15-44 were included in the measure denominator.

5. Washington State Health Care Authority (WA HCA)

In 2019, the WA HCA provided contraceptive services to female Medicaid clients aged 15-44 years who resided in 39 counties. WA HCA delivered contraceptive services to these women via the general Medicaid program or the state's family planning waiver programs, Family Planning Only and Family Planning Only – Pregnancy Related. Formerly known as Take Charge, Family Planning Only is a 1115 demonstration waiver program that serves low-income (up to 260% of FPL) uninsured male and female clients seeking to prevent unintended pregnancy, and teens and domestic violence victims who need confidential family planning services. The Family Planning Only – Pregnancy Related program (previously known as the Family Planning Only extension) provides services to recently pregnant women who lose Medicaid coverage 60 days post-pregnancy. The Washington Medicaid program serves 1.8 million members and includes 5 MCOs; about 85% of WA HCA's clients were enrolled in managed care. A CMCS MIHI grantee, WA HCA has annually calculated and publicly reported NQF #2903 at the health plan level for the past six years. Approximately 196,568 eligible women ages 15-44 comprise the NQF #2903 denominator in 2019.

6. NewYork-Presbyterian Hospital (NYP)/Columbia University Irving Medical Center Ambulatory Care Network (ACN)

In 2018, NYP ACN consisted of 14 primary care sites, 7 school-based facilities, 13 mental health school-based programs, and over 60 specialty practices serving New York City and its surrounding communities. Since 2016, NYP ACN has computed this measure annually among female clients aged 15-44 who received primary care health services from 8 NYP outpatient locations; within these 8 ACN locations are 31 facilities. NQF #2903 results are calculated at the level of facility for internal quality improvement, and about 31,084 women ages 15-44 comprise the NQF #2903 denominator in 2018.

7. Planned Parenthood Federation of America (PPFA)

PPFA comprised 49 independently incorporated affiliates, operating approximately 600 facilities in the United States, and providing reproductive health care to nearly 2.4 million patients in 2019. Through its Clinical Quality Improvement (CQI) Department, PPFA coordinates a federation-wide clinical quality improvement program for its Affiliates. A set of core reports built within PPFA's health information technology infrastructure assess this measure and other key measures of contraceptive services, quality of care, and health outcomes. Since 2012, nearly 70% of the affiliates collaborate with the PPFA CQI Department to receive quarterly quality reports on NQF #2903 and other important clinical measures, plus technical assistance for quality improvement activities. Affiliates vary in size and can cover geographic service areas that range from several counties within a single state, to an entire state population, up to multiple states; thus, an affiliate can be considered representative of a U.S state. PPFA calculates measure scores at the levels of health facility and affiliate. In 2014, about 30% of clients served by 25 PPFA affiliates were women ages 15-44. For the application, 123,978 women who visited 56 PPFA facilities in 2019 were included in the analysis.

8. Title X Family Planning Program

In 2019, the Title X Family Planning program funded 100 grantees that support a network of 3,825 family planning service sites, which in turn served 3.1 million clients. The program helped to pilot this measure through quality improvement initiatives and measure testing. In 2015-2016, OPA conducted a Performance Measure Learning Collaborative (PMLC) to support Title X grantees to improve the quality of their family planning services through use of this measure alongside adoption of strategies documented in an evidence-based change package. However, the measure is calculated and interpreted somewhat differently than the NQF #2903 specifications (e.g., the denominator is comprised of women seeking care from the reproductive health clinics). Based on the Institute of Healthcare Improvement's Breakthrough Series model, PMLC involved coaching and supporting the members through the plan, do, study, act cycle for selected change package strategies. The collaborative also convened an online community to facilitate peer exchange and learning. Ten of twelve PMLC sites (83%) experienced an increase in percentage of clients using a most or moderately effective method after employing a combination of the following strategies to improve the quality of contraceptive care: ensuring access to a broad range of contraceptive methods, providing patient-centered counseling to support reproductive life planning, developing same-day contraceptive provision systems for all methods, and utilizing diverse payment options to reduce cost as a barrier [2]. To aid PMLC sites in calculating measure scores, OPA designed and deployed an online contraceptive measures calculator. This tool allows calculation of this measure and the access to long-acting reversible contraceptive (LARC) measure using Family Planning Annual Report (FPAR) data. After completion of PMLC, the OPA-funded Reproductive Health National Training Center published on its website the change package documents and online calculator for all Title X grantees. Currently, the program uses NQF #2903 for internal quality improvement; approximately 2.5 million women ages 15-44 visited a Title X service site in 2019 and were included in the measure calculation.
In addition, OPA aims to calculate this measure and NQF #2904 (as well as related measure NQF #3543) within its grantee network using FPAR 2.0, an interoperable, standards-based reporting system that will collect a set of defined data elements from all Title X service sites. FPAR 2.0 will enable participants to improve the way they send and receive health-related data for analysis and annual reports. Currently in development, OPA has defined the FPAR 2.0 set of data elements to support the interoperability standards and is working to map each data element and response option to standardized value sets, utilizing LOINC, SNOMED CT, and RxNorm code systems. Title X grantees will collaborate with new stakeholders and technical experts to pilot and test FPAR 2.0 across the Title X network with the goal of utilization at all service sites [3].

References

[1] Louisiana Department of Health. (n.d.). Medicaid Managed Care Quality Dashboard. Retrieved December 22, 2020 from https://qualitydashboard.ldh.la.gov/

[2] Loyola Briceno, A. C., Kawatu, J., Saul, K., DeAngelis, K., Frederiksen, B., Moskosky, S. B., & Gavin, L. (2017). From theory to application: using performance measures for contraceptive care in the Title X family planning program. Contraception, 96(3), 166–174. https://doi.org/10.1016/j.contraception.2017.06.009

[3] Office of Population Affairs. Family Planning Annual Report (FPAR) 2.0. U.S. Department of Health and Human Services, Office of the Assistant Secretary for Health. Retrieved December 22, 2020 from https://opa.hhs.gov/evaluation-research/title-x-services-research/family-planning-annual-report/family-planning-1

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?*)

Not applicable.

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.*)

Not applicable.

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.

Following NQF's 2016 endorsement of #2903, OPA co-authored multiple articles in peer-reviewed journals to inform professionals delivering care in public and private settings (e.g., commercial health plans, Medicaid, community health centers, free-standing reproductive health clinics) about the new measure. These publications outline our conceptual framework for developing #2903 alongside its two complementary measures (NQF #2902 and #2904) and describe appropriate measure implementation and use. Furthermore, OPA highlighted systematic reviews which indicate that effective contraceptive method use increases the interbirth interval and reduces adolescent and unintended pregnancies. This association between use of most and moderately effective methods and positive reproductive health outcomes demonstrates the importance of contraceptive care measures to health care quality (https://doi.org/10.1016/j.contraception.2017.05.013, https://doi.org/10.1016/j.contraception.2017.06.001, https://doi.org/10.1097/AOG.0000000000002314).

To promote and support use of NQF #2903, HHS Office of Population Affairs (OPA) publishes detailed information on measure specifications and calculation on its public website (https://opa.hhs.gov/evaluation-research/title-x-services-research/contraceptive-care-measures). NQF #2903 has its own page with details on the limitations of claims data, appropriate utilization and interpretation, measure specifications, and links to programming code and code sets needed to calculate the measure (https://opa.hhs.gov/evaluation-research/title-x-services-research/contraceptive-care-measures/most-or-moderately). The latest specification available is for measurement year 2019. OPA also provides National Survey of Family Growth (NSFG) estimates needed for users to adjust their measure scores in the general Medicaid population. OPA updates its measure pages after annually updating the measure specification, code sets, and syntax.

Users can submit questions to OPA about NQF #2903 and the contraceptive care measures via two email addresses posted on the OPA website. One address goes to a general mailbox; the other is for a single point of contact for the measures at OPA. With assistance from its statistical support contractor, Far Harbor, OPA responds to technical assistance requests sent to both email addresses. Users submit inquiries related to all aspects of measure calculation, including preparing an analysis claims dataset, troubleshooting programming code, code sets used to define the measure numerator and denominator, and interpretation of scores. Some questions ask OPA for guidance on how to calculate the measure by client characteristics (e.g., benefit type, health condition) or setting (e.g., health plan, facility). The Centers for Medicaid & Medicare Services' (CMS) Health Care Quality Measures Program and the National Committee for Quality Assurance (NCQA) also forward inquiries they receive on NQF #2903 to OPA to respond directly to users needing help with measure calculation and interpretation. In FY 2020, over half of the technical assistance requests submitted to OPA were related to NQF #2903. Most requests came from state Medicaid programs reporting measure scores for CMS Adult and Child Core Sets of Health Care Quality Measures. A California public hospital system participating in the California Department of Health Care Services (DHCS) Quality Incentive Program (QIP) also asked for assistance in implementing the measure.

Starting in 2016, OPA has provided technical assistance to state Medicaid programs calculating NQF #2903. First implemented among 13 Maternal and Infant Health Initiative (MIHI) grantees during 2015 – 2018 for development and testing, the CMS Adult and Child Core Sets of Health Care Quality Measures incorporated the measure in 2017. Thus, states in addition to MIHI grantees could calculate their respective NQF #2903 scores by year to report CMS. Measure specifications, code sets, interpretation guidance, and other reporting resources are published annually for measured entities at CMS's Adult and Child Core Set website (https://www.medicaid.gov/medicaid/quality-of-care/performance-measurement/adult-and-child-health-care-quality-measures/index.html). CMS's technical assistance contractor, Mathematica Policy Research, collects feedback and questions from users on code sets, specifications, and interpretation of scores for NQF #2903 and the Health Care Quality Measures through its coordination of yearly Core Set measures' updates. Mathematica manages the requests from states computing and reviewing the measure and provides

requestors the responses from OPA. During the FFY 2018 and 2019 annual updates, OPA responded to ten technical assistance requests submitted to Mathematica by state Medicaid programs and managed care organizations.

Most MIHI grantees also participated in the Association of State and Territorial Health Officials (ASTHO) Increasing Access to Contraception Learning Community from 2015-2018, which also utilized NQF #2903 for outcome evaluation. Along with CDC and CMS, OPA supported ASTHO in dissemination of strategies and best practices to implement policies and programs to increase access to the full range of contraceptive options. OPA also presented information to the group about NQF #2903's calculation, importance, and appropriate use and implementation.

To connect with other measure users, OPA participated in the National Contraceptive Measures Workgroup, led by Planned Parenthood Federation of America (PPFA). The workgroup focused on ensuring appropriate use of NQF #2903 and contraceptive care measures and discussed efforts by health systems to implement the measure. An Implementation Subgroup supported the translation of the measures to the front lines of service delivery to minimize misunderstanding about the contraceptive care measures' purpose and intended use in the field and was coordinated by the National Family Planning & Reproductive Health Association (NFPHRA). They have developed a brief with key messages for health facility staff who want to use NQF #2903 and OPA's contraceptive care measures (https://www.nationalfamilyplanning.org/file/Onepager_Contraceptive-Measures_-Messages-for-Health-Care-Settings.pdf).

Planned Parenthood Federation of America's (PPFA) Clinical Quality Improvement (CQI) team works with its affiliates to use NQF #2903 and NQF #2904 for internal quality improvement initiatives. OPA shared with PPFA the measure specifications and code sets to utilize in CQI projects. PPFA's 2016 CQI cohort focused on contraceptive care and consisted of 35 Planned Parenthood affiliates operating 439 health centers. A total of 1,322,660 women ages 15-44 were identified with at least one health center visits in 2016 at one of those 35 affiliates. From September 2016 – June 2017, PPFA led a second cohort with 20 affiliates that aimed to improve quality and increase access to contraceptive care. Currently, PPFA CQI can review this measure's quarterly rate alongside other quality measures in an internal EHR performance measure dashboard. All CQI reports and initiatives focus on system-level strategies and honor patient choice and autonomy.

To support the implementation of the contraceptive provision measures, PPFA created a Data Stratification Guide that helps entities look at the contraceptive provision measures by different stratifications (e.g., delivery site location, payer type, patient demographics, visit type, method type) to identify subgroups where there may still be access barriers to contraception and allow entities to better understand trends and variations.

OPA worked closely with and shared feedback with its partners who contributed data for NQF #2903 reliability and validity testing (e.g., PPFA, NewYork-Presbyterian Hospital, Iowa Department of Public Health Title X grantee, and state Medicaid programs for Iowa, Louisiana, Massachusetts, Washington). To ensure correct calculation of measure numerators and denominators for analyses, OPA and its statistical support contractor Far Harbor provided the partners with a summary data request and technical assistance via email and online meeting. Partners received programming syntax to calculate measure scores and aggregate data for analysis as needed. OPA and Far Harbor reviewed the datasets and aggregate tables and met with the data partners to confirm that the results contained the correct measure numerators and denominators by age group. Once prepared, data was analyzed and summarized to submit for NQF maintenance endorsement. Descriptive statistics were computed for each dataset and included in this application. Each partner will receive a detailed summary report with an overview of methods and full reliability and/or validity results at the levels of analysis available.

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.

To assist states in calculating NQF #2903 for public reporting, CMS relies on OPA to provide annually the latest measure code sets, specifications, and programming syntax for measure calculation. CMS also offers several resources to assist state Medicaid programs in computing the measure. As CMS technical assistance

contractor, Mathematica Policy Research conducts quality assurance on the measure data submitted and works with states to resolve any issues with the data reported. The code sets and specifications are published by CMS in its Technical Specifications and Resource Manual for the Child and Adult Core Sets (https://www.medicaid.gov/medicaid/quality-of-care/downloads/medicaid-and-chip-child-core-set-manual.pdf, https://www.medicaid.gov/medicaid/quality-of-care/downloads/medicaid-adult-core-set-manual.pdf). The latest manual provides reporting resources for measurement year 2020, which also includes an interpretation guide for NQF #2903 to help states understand their measure scores. This interpretation guide was developed by OPA and is posted on OPA's website as well (https://opa.hhs.gov/sites/default/files/2020-07/interpreting-rates-for-contraceptive-care-measures.pdf). CMS and Mathematica also conduct regular technical assistance webinars (about two per year) for Core Set users to hear how states are using the measures, including the contraceptive care measures, and to answer any questions states have about calculating and reporting on the measures.

CMS' Center for Medicaid & CHIP Services (CMCS) annually releases Adult and Child Core Set data for measures that were reported by at least 25 states and met its internal standards for data quality. For Federal Fiscal Year (FFY) 2018, NQF #2902, NQF #2903, and NQF #2904 met CMCS's threshold for public reporting of state-specific results, and thus CMS publicly reported these rates for the first time. In FFY 2019, the number of states reporting NQF #2903 in ages 15-20 increased from 26 to 28; Alabama, Alaska, California, Colorado, Connecticut, Delaware, Florida, Illinois, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Michigan, Missouri, New Hampshire, New York, North Carolina, North Dakota, Oklahoma, Pennsylvania, South Carolina, Tennessee, Texas, Vermont, Washington, West Virginia, Wyoming all reported their scores at the state level. NQF #2903 and NQF #2904 rates for ages 15-20 by state are available online in the State Medicaid & CHIP Profiles (https://www.medicaid.gov/state-overviews/index.html). Only 23 states reported NQF #2903 and NQF #2904 for ages 21-44, so CMS did not publish these state-specific measure scores. For an overview of Child and Adult Core Set Reporting for FFY 2019, CMCS also published a Fact Sheet online (https://www.medicaid.gov/medicaid/quality-of-care/downloads/ffy-2019-core-set-reporting.pdf).

In addition to its public-facing web pages for the contraceptive care measures, OPA annually reports NQF #2903 and NQF #2904 among women seeking care from each Title X Family Planning Program grantee state and territory in the Title X Family Planning Annual Report National Summary (https://opa.hhs.gov/evaluation-research/title-x-services-research/family-planning-annual-report). OPA also disseminates The Contraceptive Access Change via its Reproductive Health National Training Center website to support Title X grantees' performance improvement on NQF #2903 and NQF #2904 (https://rhntc.org/resources/contraceptive-access-change-package). This evidence-based change package was refined through a Title X grantee Performance Measure Learning Collaborative (PMLC). Ten of twelve PMLC sites (83%) experienced an increase in percentage of clients using a most or moderately effective method after employing a combination of the following strategies to improve the quality of contraceptive care: ensuring access to a broad range of contraceptive methods, providing patient-centered counseling to support reproductive life planning, developing same-day contraceptive provision systems for all methods, and utilizing diverse payment options to reduce cost as a barrier (https://doi.org/10.1016/j.contraception.2017.06.009). The four best practices identified in the Contraceptive Access Change Package were:

a. Stock a broad range of contraceptive methods;
b. Discuss pregnancy intention and support patients through evidence-informed, patient-centered counseling;
c. Develop systems for same-visit provision of all contraceptive methods, at all visit types;
d. Utilize diverse payment options to reduce cost as a barrier for the facility and the patient.

In addition, OPA aims to calculate this measure and NQF #2904 (as well as related measure NQF #3543) within its grantee network using FPAR 2.0, an interoperable, standards-based reporting system that will collect a set of defined data elements from all Title X service sites. FPAR 2.0 will enable participants to improve the way they send and receive health-related data for analysis and annual reports. Currently in development, OPA has

defined the FPAR 2.0 set of data elements to support the interoperability standards and is working to map each data element and response option to standardized value sets, utilizing LOINC, SNOMED CT, and RxNorm code systems. Title X grantees will collaborate with new stakeholders and technical experts to pilot and test FPAR 2.0 across the Title X network with the goal of utilization at all service sites (https://opa.hhs.gov/evaluation-research/title-x-services-research/family-planning-annual-report/family-planning-1).

To strengthen performance measurement capacity and support quality improvement initiatives, PPFA's Clinical Quality Improvement (CQI) team provides quarterly clinical quality measure dashboards to a subset of its affiliates via a shared electronic health record (EHR) system. PPFA completed two CQI cohorts of affiliates which implemented NQF #2903 and NQF #2904 in its quality measure dashboards. The cohorts aimed to improve quality, increase access to contraceptive care, and remove barriers for patients when they wish to receive a contraceptive method of their choice. Participating teams reviewed their performance on NQF #2903 and NQF #2904 monthly to determine where barriers might exist and created improvement plans. Teams shared successful strategies and lessons learned around clinic workflow, payment and reimbursement, patient education, and staff training. Data were automatically uploaded from the EHR into a data warehouse where the report logic is configured. The dashboards display breakdowns of the measures across health centers, visit types, and by providers allowing health centers to identify performance strengths, variations, and opportunities for improvement. As a result, NQF #2903 and #2904 became main components of PPFA's performance measurement. PPFA continues to track NQF #2903 and NQF #2904 scores quarterly within each affiliate and across the federation through its CQI dashboard. This allows PPFA providers to assess how well patient needs are being met and identify opportunities to strengthen service provision.

In addition to convening the National Contraceptive Measures Workgroup to support appropriate contraceptive care measure use, PPFA released a policy paper with Manatt Health in October 2019 that helps state policymakers and payers implement contraceptive care quality measures to improve access to all forms of contraception. The paper, "Measuring Quality Contraceptive Care in a Value-Based System," serves as a tool for policymakers, detailing how to incorporate contraceptive care quality measures (NQF #2902, NQF #2903, and NQF #2904) in Value Based Payment (VBP) initiatives to both ensure agency in women's contraceptive choices and develop strategies to improve people's access to contraception (https://www.plannedparenthood.org/uploads/filer_public/7e/90/7e90b4cb-4b3d-499f-8c6c-f31ab865b621/ppfa-manatt_measuring_quality_contraceptive_care.pdf).

PPFA's current CQI focus related to NQF #2903 and NQF #2904 is to pilot these measures' tandem use in facilities with the Person-Centered Contraceptive Counseling (PCCC) measure (NQF #3543) developed by University of California San Francisco (UCSF). PPFA has conducted webinars and briefings on NQF #2904 and NQF #3543 in tandem use for its affiliates, which can also request individual coaching sessions with the CQI team. These resources build upon the joint PPFA-Manatt Health policy paper and encourages affiliates to collaborate with its state agency counterparts to appropriately utilize NQF #2903 by implementing the measures in pay-for-reporting settings and minimizing risk of patient coercion.

NewYork-Presbyterian Hospital (NYP)/Columbia University Irving Medical Center Ambulatory Care Network (ACN) began testing NQF #2903 and NQF #2904 in 2016. Calculating the measures by year, age group, and facility, NYP ACN began building the infrastructure to create annual reports for external reporting as well as quarterly reports for internal quality improvement. Although paused for implementation of a new EHR system, this project has been well received by departmental leadership and hospital-wide quality leadership. NYP ACN aims to include NQF #2903 and NQF #2904 as part of the quality bundles evaluating departments, facilities, and providers on client-centered contraceptive care.

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.

Since 2015, OPA has been the recipient of on-going feedback on NQF #2903 through CMS. CMS has a contract with Mathematica Policy Research to provide technical assistance (TA) on states reporting NQF #2902, NQF

#2903, and NQF #2904 for the CMS Adult and Child Core sets. Mathematica manages a TA email inbox that states use to provide feedback on the measures and receive technical assistance. Mathematica forwards messages on NQF #2903 from the TA box to OPA as needed, who then drafts responses to requestors.

OPA has also received feedback on NQF #2902, NQF #2903, and NQF #2904 via the e-mail addresses posted on its public-facing website. Multiple organizations (e.g., state Medicaid programs, public hospital systems, universities, and public health agencies) which are implementing and computing the measures send or forward their questions this way; OPA replies via email.

OPA convenes an expert panel to discuss the appropriate use and interpretation of this measure in different health systems (e.g., programs with a reproductive health services focus compared to general health care providers). On September 9 and 11, 2020, OPA held an online Expert User Group Meeting on the Contraceptive Care Performance Measures, which included current and future measure users. One purpose of this conference was to gather feedback on the contraceptive care measures. During 15-minute discussion sessions at the conference, we asked expert users to describe their current or planned use of the contraceptive care measures, how the measures have helped improve the quality of care to date, and how the measures can be improved. In addition, two states that received CMS' MIHI funding presented to the panel a summary of their experiences implementing NQF #2903. A meeting facilitator recorded input from attendees in a summary document.

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

Measure users, including states reporting NQF #2903 scores to CMS and reproductive health organizations utilizing this measure for quality improvement, shared the following input this year:

- Using the Generic Product Identifier (GPI) code system to identify contraceptive medications for the numerator has advantages over FDA's National Drug Code (NDC) system. New NDCs are created frequently for new products and identify over one thousand oral pills available for contraceptive use. The repositories containing NDCs for prescription contraceptive medications are difficult to utilize and search for valid codes. GPI uses fewer codes to identify oral contraceptive pills and may simplify the measure code sets and numerator calculation.
- Consider state-specific policies for coding administrative claims for prescription contraceptive medications in measure specifications. One state described its coding guidelines for requiring modifiers indicating family planning use to flag CPT codes 11981, 11982, 11983 as related to contraceptive implants (which is a method counted in the NQF #2903 numerator) and the HCPCS code S4993 to only denote emergency contraception (which is excluded from the NQF #2903 numerator).
- As described in 3c.1, multiple states stated that the calculation of NQF #2903 was complex and time-consuming, even with OPA's published SAS programming code. While the syntax has been simplified since NQF #2903's original endorsement, other barriers related to measure calculation may exist for states. One state reported that the available syntax did not mesh well with its existing data systems, requiring their analysts to develop syntax from scratch.
- PPFA reported that affiliates participating in its CQI cohorts using the measures found it challenging to interpret performance on NQF #2903 and NQF #2904 while considering client preferences. PPFA noted that utilization does not directly measure access, and cohort teams were not sure how to set improvement targets. Along with the National Family Planning & Reproductive Health Association (NFPRHA), PPFA re-iterated that NQF #2903 should be calculated by geography, health plan (e.g., Medicaid managed care organization), and other patient attributes (e.g. race, ethnicity, benefit type, etc.) to examine disparities in access and to establish stratified baseline measure scores for future quality improvement initiatives. Another recommendation is for health systems to report overall and stratified NQF #2903 scores publicly for analysis and discussion.
- OPA continues to receive feedback on appropriate interpretation of the measure, as health systems naturally want to increase their measure scores on a performance measure. It is hypothesized that

some providers may therefore use a non-client-centered manner during contraceptive care. As stated on our website, we emphasize that OPA has not yet set a specific benchmark for NQF #2903, but "does not expect it to reach 100%, as some women will make informed decisions to choose methods in the lower tier of efficacy even when offered the full range of methods and all logistical or financial barriers to access are removed."(https://opa.hhs.gov/evaluation-research/title-x-services-research/contraceptive-care-measures/most-or-moderately) OPA encourages states to use NQF #2903 in tandem with the Person-Centered Contraceptive Counseling (PCCC) measure developed by University of California San Francisco (UCSF) or another measure of client experience to ensure contraceptive care is provided in a patient-centered manner. Recently endorsed in November 2020 by NQF's Consensus Standards Approval Committee as NQF #3543, research has started to identify ways to operationalize the 'tandem use' of NQF #2903 with the new PCCC measure.

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

A measure user pointed out that the current edition of the clinical reference Contraceptive Technology (http://www.contraceptivetechnology.org/the-book/take-a-peek/contraceptive-efficacy/) classified diaphragm as a less effective method of contraception because of increased estimated typical use failure rates. This user asked if the NQF #2903 numerator had been updated to consider these new failure rates.

Other users of the measures have provided feedback on CPT codes for hysterectomy and oophorectomy that were not included in the measure specifications to indicate sterilization for non-contraceptive reasons and determine a woman is not at risk for unintended pregnancy. These codes are:

- 58550 Laparoscopy, surgical, with vaginal hysterectomy, for uterus 250 g or less
- 58553 Laparoscopy, surgical, with vaginal hysterectomy, for uterus greater than 250 g
- 58575 Laparoscopy, surgical, total hysterectomy for resection of malignancy (tumor debulking), with omentectomy including salpingo-oophorectomy, unilateral or bilateral, when performed
- 59120 Surgical treatment of ectopic pregnancy; tubal or ovarian, requiring salpingectomy and/or oophorectomy, abdominal or vaginal approach
- 59151 Laparoscopic treatment of ectopic pregnancy; with salpingectomy and/or oophorectomy
- 59135 Surgical treatment of ectopic pregnancy; tubal or ovarian, requiring total hysterectomy

OPA received inquiries asking if this measure has a lookback period for women who underwent a sterilization procedure or obtained a LARC method prior to the measurement year. These users also asked if providers should offer contraception after sterilization and wondered if it makes sense to only count clients receiving a most or moderately effective method during the year.

Another user suggested that codes related to bilateral salpingectomy should be added to indicate use of female sterilization as contraception because the procedure is an increasingly common surgical method for sterilization. These CPT and ICD-10-PCS codes include:

- 0U570ZZ Destruction of Bilateral Fallopian Tubes, Open Approach
- 0U573ZZ Destruction of Bilateral Fallopian Tubes, Percutaneous Approach
- 0U577ZZ Destruction of Bilateral Fallopian Tubes, Via Natural or Artificial Opening
- 0UL70CZ Occlusion of Bilateral Fallopian Tubes with Extraluminal Device, Open Approach
- 0UL70DZ Occlusion of Bilateral Fallopian Tubes with Intraluminal Device, Open Approach
- 0UL70ZZ Occlusion of Bilateral Fallopian Tubes, Open Approach
- 0UL73CZ Occlusion of Bilateral Fallopian Tubes with Extraluminal Device, Percutaneous Approach
- 0UL73DZ Occlusion of Bilateral Fallopian Tubes with Intraluminal Device, Percutaneous Approach
- 0UL73ZZ Occlusion of Bilateral Fallopian Tubes, Percutaneous Approach

- 0UL77DZ Occlusion of Bilateral Fallopian Tubes with Intraluminal Device, Via Natural or Artificial Opening
- 0UL77ZZ Occlusion of Bilateral Fallopian Tubes, Via Natural or Artificial Opening
- 0UT70ZZ Resection of Bilateral Fallopian Tubes, Open Approach
- 0UT74ZZ Resection of Bilateral Fallopian Tubes, Percutaneous Endoscopic Approach
- 0UT77ZZ Resection of Bilateral Fallopian Tubes, Via Natural or Artificial Opening
- 0UT78ZZ Resection of Bilateral Fallopian Tubes, Via Natural or Artificial Opening Endoscopic
- 0UT7FZZ Resection of Bilateral Fallopian Tubes, Via Natural or Artificial Opening With Percutaneous Endoscopic Assistance

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.

To align the measure numerator with the latest edition of Contraceptive Technology, CDC, and WHO publications on contraceptive effectiveness, we changed the measure specifications to exclude diaphragm from the NQF #2903 numerator. Sixty-three codes were removed from the code sets as a result.

The Generic Product Identifier (GPI) code system requires a license fee to utilize, which may not be possible for all states calculating NQF #2903 and the contraceptive care measures. OPA will continue to only utilize NDC codes to identify medications for the measure numerator for now, even though it has frequent updates and is time-consuming to search.

Regarding the use of S4993 only for emergency contraception, OPA will investigate the various state-specific policies and examine data for this procedure code in administrative claims. While one state uses it only for emergency contraception, another state requires a specific modifier for it to be used for the same reimbursement. This code will remain in the NQF #2903 sets for numerator compilation for the next measurement year.

Regarding the suggestion to include additional CPT codes for hysterectomy and oophorectomy to indicate sterilization for non-contraceptive reasons and determine a woman is not at risk for unintended pregnancy, additional CPT and ICD-10-PCS procedure codes were included for measurement year 2020 in CCW-A, Codes Indicating Sterilization for Non-Contraceptive Reasons (e.g., hysterectomy, oophorectomy, or menopause). Previous measurement years did not utilize ICD-10-PCS codes in CCW-A. The following 4 CPT codes and 19 ICD-10-PCS codes were added:

- 58550 Laparoscopy, surgical, with vaginal hysterectomy, for uterus 250 g or less
- 58553 Laparoscopy, surgical, with vaginal hysterectomy, for uterus greater than 250 g
- 58575 Laparoscopy, surgical, total hysterectomy for resection of malignancy (tumor debulking), with omentectomy including salpingo-oophorectomy, unilateral or bilateral, when performed
- 59135 Surgical treatment of ectopic pregnancy; tubal or ovarian, requiring total hysterectomy
- 0U520ZZ Destruction of Bilateral Ovaries, Open Approach
- 0U523ZZ Destruction of Bilateral Ovaries, Percutaneous Approach
- 0U524ZZ Destruction of Bilateral Ovaries, Percutaneous Endoscopic Approach
- 0U528ZZ Destruction of Bilateral Ovaries, Via Natural or Artificial Opening Endoscopic
- 0UT20ZZ Resection of Bilateral Ovaries, Open Approach
- 0UT24ZZ Resection of Bilateral Ovaries, Percutaneous Endoscopic Approach
- 0UT27ZZ Resection of Bilateral Ovaries, Via Natural or Artificial Opening
- 0UT28ZZ Resection of Bilateral Ovaries, Via Natural or Artificial Opening Endoscopic

- 0UT2FZZ Resection of Bilateral Ovaries, Via Natural or Artificial Opening With Percutaneous
- 0UT90ZL Resection of Uterus, Supracervical, Open Approach
- 0UT90ZZ Resection of Uterus, Open Approach
- 0UT94ZL Resection of Uterus, Supracervical, Percutaneous Endoscopic Approach
- 0UT94ZZ Resection of Uterus, Percutaneous Endoscopic Approach
- 0UT97ZL Resection of Uterus, Supracervical, Via Natural or Artificial Opening
- 0UT97ZZ Resection of Uterus, Via Natural or Artificial Opening
- 0UT98ZL Resection of Uterus, Supracervical, Via Natural or Artificial Opening Endoscopic
- 0UT98ZZ Resection of Uterus, Via Natural or Artificial Opening Endoscopic
- 0UT9FZL Resection of Uterus, Supracervical, Via Natural or Artificial Opening With Percutaneous Endoscopic Assistance
- 0UT9FZZ Resection of Uterus, Via Natural or Artificial Opening With Percutaneous Endoscopic Assistance

The following 2 codes were included in CCW-A for the 2019 measurement year. After re-evaluation for the 2020 measurement year, they were removed from CCW-A in part because they could indicate unilateral salpingectomy or oophorectomy, which might still allow women to become pregnant. These codes are:

- 59120 Surgical treatment of ectopic pregnancy; tubal or ovarian, requiring salpingectomy and/or oophorectomy, abdominal or vaginal approach
- 59151 Laparoscopic treatment of ectopic pregnancy; with salpingectomy and/or oophorectomy

For measurement year 2020, we decided to augment Table CCW-B Codes Indicating a Pregnancy by adding 21 ICD-10-CM codes for maternal care for abnormalities of the fetal heart rate or rhythm and 1 new pregnancy code. These codes are:

- O36.8310 Maternal care for abnormalities of the fetal heart rate or rhythm, first trimester, not applicable or unspecified
- O36.8311 Maternal care for abnormalities of the fetal heart rate or rhythm, first trimester, fetus 1
- O36.8312 Maternal care for abnormalities of the fetal heart rate or rhythm, first trimester, fetus 2
- O36.8313 Maternal care for abnormalities of the fetal heart rate or rhythm, first trimester, fetus 3
- O36.8314 Maternal care for abnormalities of the fetal heart rate or rhythm, first trimester, fetus 4
- O36.8315 Maternal care for abnormalities of the fetal heart rate or rhythm, first trimester, fetus 5
- O36.8319 Maternal care for abnormalities of the fetal heart rate or rhythm, first trimester, other fetus
- O36.8320 Maternal care for abnormalities of the fetal heart rate or rhythm, second trimester, not applicable or unspecified
- O36.8321 Maternal care for abnormalities of the fetal heart rate or rhythm, second trimester, fetus 1
- O36.8322 Maternal care for abnormalities of the fetal heart rate or rhythm, second trimester, fetus 2
- O36.8323 Maternal care for abnormalities of the fetal heart rate or rhythm, second trimester, fetus 3
- O36.8324 Maternal care for abnormalities of the fetal heart rate or rhythm, second trimester, fetus 4
- O36.8325 Maternal care for abnormalities of the fetal heart rate or rhythm, second trimester, fetus 5
- O36.8329 Maternal care for abnormalities of the fetal heart rate or rhythm, second trimester, other fetus
- O36.8330 Maternal care for abnormalities of the fetal heart rate or rhythm, third trimester, not applicable or unspecified

- O36.8331 Maternal care for abnormalities of the fetal heart rate or rhythm, third trimester, fetus 1
- O36.8332 Maternal care for abnormalities of the fetal heart rate or rhythm, third trimester, fetus 2
- O36.8333 Maternal care for abnormalities of the fetal heart rate or rhythm, third trimester, fetus 3
- O36.8334 Maternal care for abnormalities of the fetal heart rate or rhythm, third trimester, fetus 4
- O36.8335 Maternal care for abnormalities of the fetal heart rate or rhythm, third trimester, fetus 5
- O36.8339 Maternal care for abnormalities of the fetal heart rate or rhythm, third trimester, other fetus
- O99.891 Other specified diseases and conditions complicating pregnancy

After confirming the existence of these codes in CPT and ICD-10-PCS (https://www.cms.gov/Medicare/Coding/ICD10/index), we added the following 5 procedure codes in Table CCW-C:

- 59151 Laparoscopic treatment of ectopic pregnancy; with salpingectomy and/or oophorectomy
- 10D20ZZ Extraction of Products of Conception, Ectopic, Open Approach
- 10D24ZZ Extraction of Products of Conception, Ectopic, Percutaneous Endoscopic Approach
- 10D27ZZ Extraction of Products of Conception, Ectopic, Via Natural or Artificial Opening
- 10D28ZZ Extraction of Products of Conception, Ectopic, Via Natural or Artificial Opening Endoscopic

We added 17 procedure codes to CCW-E Codes Used to Identify Provision of a Most or Moderately Effective Contraceptive Method for measurement year 2020 to indicate female sterilization, including 16 codes for bilateral salpingectomy. These codes are:

- 0567T Blockage of fallopian tubes with implants inserted through cervix
- 0U570ZZ Destruction of Bilateral Fallopian Tubes, Open Approach
- 0U573ZZ Destruction of Bilateral Fallopian Tubes, Percutaneous Approach
- 0U577ZZ Destruction of Bilateral Fallopian Tubes, Via Natural or Artificial Opening
- 0UL70CZ Occlusion of Bilateral Fallopian Tubes with Extraluminal Device, Open Approach
- 0UL70DZ Occlusion of Bilateral Fallopian Tubes with Intraluminal Device, Open Approach
- 0UL70ZZ Occlusion of Bilateral Fallopian Tubes, Open Approach
- 0UL73CZ Occlusion of Bilateral Fallopian Tubes with Extraluminal Device, Percutaneous Approach
- 0UL73DZ Occlusion of Bilateral Fallopian Tubes with Intraluminal Device, Percutaneous Approach
- 0UL73ZZ Occlusion of Bilateral Fallopian Tubes, Percutaneous Approach
- 0UL77DZ Occlusion of Bilateral Fallopian Tubes with Intraluminal Device, Via Natural or Artificial Opening
- 0UL77ZZ Occlusion of Bilateral Fallopian Tubes, Via Natural or Artificial Opening
- 0UT70ZZ Resection of Bilateral Fallopian Tubes, Open Approach
- 0UT74ZZ Resection of Bilateral Fallopian Tubes, Percutaneous Endoscopic Approach
- 0UT77ZZ Resection of Bilateral Fallopian Tubes, Via Natural or Artificial Opening
- 0UT78ZZ Resection of Bilateral Fallopian Tubes, Via Natural or Artificial Opening Endoscopic
- 0UT7FZZ Resection of Bilateral Fallopian Tubes, Via Natural or Artificial Opening with Percutaneous Endoscopic Assistance

We responded to users with questions about a lookback period by explaining that measure does not count contraception that is "ever provided". It looks only within the measurement year to assess contraception provided during that period (i.e., annual provision). These rates are expected to be lower than contraception "ever provided", but they will be consistently lower when comparing across clinics, and it enables year over

year comparisons. Thus, for the purposes of identifying lowest performing clinics that could use a quality improvement intervention, the current specification is appropriate. Women who already use a most or moderately effective method can be included in the numerator if they have a claim with a diagnosis surveillance code during the measurement year. These diagnosis surveillance codes denote when a health care provider assesses a woman's current method and are among the codes used to define the numerator.

For this application, OPA calculated NQF #2903 at several levels of analysis: facility, clinician group/practice, health plan, public health region, and state to test the measure's reliability and validity. In this form's 1b.4, measure scores were examined by race/ethnicity (and over time, where available) in multiple datasets to examine differences in access. OPA agrees with the importance of stratifying NQF #2903 scores by client characteristics to monitor quality improvement initiatives and better understand contraceptive provision among women wishing to use most or moderately effective methods. To address the concerns around appropriate measure implementation and interpretation, OPA will continue to promote use of NQF #2903 in tandem with the Person-Centered Contraceptive Counseling (PCCC) measure developed by University of California San Francisco or another measure of client experience to ensure contraceptive care is provided in a patient-centered manner. Recently endorsed in November 2020 by NQF's Consensus Standards Approval Committee, research is currently under way to identify ways to operationalize the 'tandem use' of NQF #2903 with the new PCCC measure (NQF #3543).

Improvement

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

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

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

The results from federal and state Medicaid programs, and the NYP ACN, indicate that approximately one-third of women were provided a most or moderately effective method of contraception. These estimates are higher than the NQF #2903 measure scores reported in a recent study conducted using data from community health centers (e.g., federally qualified health centers, rural health centers, county health departments), which reported that the provision of most or moderately effective methods to be about 24% in states with Medicaid expansion and 20% in non-expansion states [9]. When NQF #2903 scores presented in this application are adjusted to approximate the use rate among women that are at risk of unintended pregnancy (i.e., because they have had sex, are neither pregnant nor seeking pregnancy, and are fecund) using the National Survey of Family Grown (NSFG) data in Appendix B, the adjusted estimate suggests that approximately 65% were using a most or moderately effective method.

This leaves up to a 35-percentage point opportunity for improvement. However, as the measure steward, we have noted that: "No specific benchmark has been set for this measure, but the Office of Population Affairs (OPA) does not expect it to reach 100%, as some women will make informed decisions to choose methods in the lower tier of efficacy even when offered the full range of methods and all logistical or financial barriers to access are removed."[1] Hence, we recommend using a more conservative estimate, e.g., a 15-20 percentage point opportunity for improvement.

The measure scores from programs that focus on the delivery of reproductive health services (e.g., Iowa´s state-funded Family Planning Program, Planned Parenthood, and Title X) do not need to be adjusted with NSFG data. This is because most clients who receive services from these programs are seeking contraceptive services and should therefore be considered at risk of unintended pregnancy. It is noteworthy that the estimates from

the Title X program are similar to the rates reported from the other programs, after adjustment for risk of unintended pregnancy. There are also Title X grantees that had measure scores of 0%, which should be investigated further and may be a result of the Final Rule, which resulted in Title X regulations that de-emphasized CDC's and OPA's Providing Quality Family Planning Services Recommendations [2] that promote offering a full range of contraceptive methods for persons seeking to prevent pregnancy.

Some IME clinician group/practices had measure scores of 100%. While these were likely entities with small numbers of patients, it is important to ensure patient-centered contraceptive counseling is being provided and women are not being coerced into receiving most and moderately effective methods. A range of contraceptive preferences is expected, and it is vital that women who wish to use contraception have the full range of methods available to them.

Data on changes in performance over time show that trends have increased very slightly (e.g., in Iowa) or have remained stable over time (e.g., in the Title X program). An important piece of context is that the past 3-4 years of measure use (2016-2020) have coincided with a presidential administration that restricted efforts to expand access to contraceptive care. For example, systems change efforts such as CMS' Maternal and Infant Health initiative were not renewed, and regulatory changes to the Title X program decimated the service delivery system (the number of family planning users seen by the Title X program dropped from 3.6 million in 2016 to 2.7 million in 2019, a decline of 25%) [3-6]. This likely led to the slight decrease in mean measure scores from 2018 to 2019 in the Title X program. The experience in Oregon demonstrates that when a state makes a concerted effort to improve performance on the measure, it is possible to do so.

However, improvements and strategies such as those employed in Oregon must be weighed against the potential risk of coercive practices and highlights the need for a 'balancing' measure to ensure access to contraception is offered in a client-centered manner. A measure of client experience with contraceptive care has just been endorsed by NQF: the Person-Centered Contraceptive Counseling (PCCC) measure (NQF #3543). The PCCC is a patient-reported outcome performance measure (PRO-PM) that assesses the patient-centeredness of contraceptive counseling [7]. Research is currently under way to identify ways to operationalize the 'tandem use' of NQF #2903 with the new PCCC measure.

In sum, we believe that the measures should be re-endorsed given that there remains substantial room for improvement and that the isolated and/or relatively modest improvements of the past 3-4 years are due to contextual influences, which will be ameliorated moving forward. Investigators at UCSF are currently conducting research that will allow NQF #2903 to be used together with NQF #3543, and to ensure women are offered contraceptive care that is client-centered.

References

[1] Office of Population Affairs. (n.d.). Most or Moderately Effective Contraceptive Methods. https://opa.hhs.gov/evaluation-research/title-x-services-research/contraceptive-care-measures/most-or-moderately. Accessed 22 Dec 2020.

[2] Gavin, L., Moskosky, S., Carter, M., Curtis, K., Glass, E., Godfrey, E., Marcell, A., Mautone-Smith, N., Pazol, K., Tepper, N., Zapata, L., & Centers for Disease Control and Prevention (CDC) (2014). Providing quality family planning services: Recommendations of CDC and the U.S. Office of Population Affairs. MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 63(RR-04), 1–54.

[3] Fowler, C. I., Gable, J., Lasater, B., & Asman, K. (2020, September). Family Planning Annual Report: 2019 National Summary. Washington, DC: Office of Population Affairs, Office of the Assistant Secretary for Health, Department of Health and Human Services. https://opa.hhs.gov/sites/default/files/2020-09/title-x-fpar-2019-national-summary.pdf

[4] Fowler, C. I., Gable, J., Wang, J., & Lasater, B. (2017, August). Family Planning Annual Report: 2016 national summary. Research Triangle Park, NC: RTI International. https://opa.hhs.gov/sites/default/files/2020-07/title-x-fpar-2016-national.pdf

[5] Gold, R.B., Cross, L. The Title X Gag Rule Is Wreaking Havoc—Just as Trump Intended. Guttmacher Institute. Retrieved December 22, 2020 from https://www.guttmacher.org/article/2019/08/title-x-gag-rule-wreaking-havoc-just-trump-intended

[6] Frederiksen B, Salganicoff A, Ramaswamy A, Gomez I. Current Status of the Title X Network and the Path Forward. Kaiser Family Foundation. Retrieved December 22, 2020 from https://www.kff.org/womens-health-policy/issue-brief/current-status-of-the-title-x-network-and-the-path-forward/. Accessed 22 Dec 2020.

[7] University of California San Francisco. The Person-Centered Contraceptive Counseling Measure. Retrieved December 22, 2020 from https://pcccmeasure.ucsf.edu/. Accessed 22 Dec 2020.

[8] University of California San Francisco. Christine Dehlendorf, MD, MAS. Retrieved December 22, 2020 from https://pcccmeasure.ucsf.edu/people/christine-dehlendorf-md-mas

[9] Darney, B. G., Jacob, R. L., Hoopes, M., Rodriguez, M. I., Hatch, B., Marino, M., Templeton, A., Oakley, J., & Cottrell, E. K. (2020). Evaluation of Medicaid Expansion Under the Affordable Care Act and Contraceptive Care in US Community Health Centers. JAMA network open, 3(6), e206874. https://doi.org/10.1001/jamanetworkopen.2020.6874

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.

No unintended negative consequences were identified. The one issue that remains a potential concern is that the measure may lead to coercive practices in which women are not offered a free choice of methods and are pressured to use most or moderately effective contraception [1-3]. We reaffirm our commitment to client-centered care through the following actions taken during development and testing of NQF #2903.

Although existing research [4, 5] show a high percentage of women will choose LARC when given the opportunity, OPA has deliberately not set a benchmark for this measure. We explicitly state this on the measure website and provide specific guidance on how the contraceptive care measures should be used. This should remove pressure on providers to improperly push all women to use a most or moderately effective method. We also designed NQF #2903 so that seven methods of contraception are included in the numerator, which are treated as being of equal value during measure calculation. Hence, the numerator represents a wide range of methods from which clients can choose. We hope this encourages providers to deliver family planning care in a fully client-centered, non-coercive manner.

In partnership with CDC, OPA also co-authored detailed recommendations on providing client-centered contraceptive counseling [6]. To deliver provider education on this topic, we sponsored multiple online training modules. OPA published its first online client-centered contraceptive counseling training module, "Quality Contraceptive Counseling and Education: A Client-Centered Conversation eLearning and Explaining Contraception for Healthcare Providers eLearning" in 2017. This OPA-sponsored training was updated to a new module in September 2020, "Contraceptive Counseling and Education eLearning", which is available to all providers at the OPA's Reproductive Health National Training Center website [7].

The OPA team and our partners involved in measure development anticipated that utilization of the contraceptive care measures could unintentionally result in incentivizing providers to impel patients to use more effective methods. During the NQF endorsement process for the contraceptive care measures, stakeholders echoed this concern during the public comment period and suggested an accompanying measure of patient experience with contraceptive care. The National Partnership for Women & Families described this balancing measure further by stating, "Such a measure can be expected to help identify and/or check inappropriate pressure from the health care system." After NQF endorsed the contraceptive care measures, OPA acted on this shared concern by providing funding to the University of California San Francisco to support

initial development of a patient-reported outcome performance measure (PRO-PM). Following the first year of funding, UCSF secured private funding to continue the project. Recently endorsed by NQF in November 2020 as the Person-Centered Contraceptive Counseling measure (PCCC), it facilitates proper use of the provision measures by allowing organizations to observe variations in patient experience that occur with changes in provision of most or moderately effective contraception. Health care providers can then ensure that increases in provision are not associated with worse patient experience; ideally, improved provision would be linked to better patient experience. The UCSF team has started research to operationalize the 'tandem use' of NQF #2903 with the new PCCC measure.

References

[1] Dehlendorf, C., Bellanca, H., & Policar, M. (2015). Performance measures for contraceptive care: what are we actually trying to measure?. Contraception, 91(6), 433–437. https://doi.org/10.1016/j.contraception.2015.02.002

[2] Gold, R.B. (2014). Guarding Against Coercion While Ensuring Access: A Delicate Balance. Guttmacher Policy Review, 17(3), 8-14.

[3] Sonfield, A. (2017). Why family planning policy and practice must guarantee a true choice of contraceptive methods. Guttmacher Policy Review, 20, 103–107.

[4] Harper, C. C., Rocca, C. H., Thompson, K. M., Morfesis, J., Goodman, S., Darney, P. D., Westhoff, C. L., & Speidel, J. J. (2015). Reductions in pregnancy rates in the USA with long-acting reversible contraception: a cluster randomised trial. Lancet (London, England), 386(9993), 562–568. https://doi.org/10.1016/S0140-6736(14)62460-0

[5] Winner, B., Peipert, J. F., Zhao, Q., Buckel, C., Madden, T., Allsworth, J. E., & Secura, G. M. (2012). Effectiveness of long-acting reversible contraception. The New England journal of medicine, 366(21), 1998–2007. https://doi.org/10.1056/NEJMoa1110855

[6] Gavin, L., Moskosky, S., Carter, M., Curtis, K., Glass, E., Godfrey, E., Marcell, A., Mautone-Smith, N., Pazol, K., Tepper, N., Zapata, L., & Centers for Disease Control and Prevention (CDC) (2014). Providing quality family planning services: Recommendations of CDC and the U.S. Office of Population Affairs. MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports, 63(RR-04), 1–54.

[7] Reproductive Health National Training Center. (n.d.). Contraceptive Counseling and Education eLearning. U.S. Department of Health and Human Services, Office of the Assistant Secretary for Health. Retrieved December 22, 2020 https://rhntc.org/resources/contraceptive-counseling-and-education-elearning

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

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)

2902 : Contraceptive Care - Postpartum

2904 : Contraceptive Care - Access to LARC

3543 : Person-Centered Contraceptive Counseling (PCCC) measure

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?

Yes

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

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.)

OPA is submitting two other applications for NQF maintenance endorsement, which are complementary to this application. One of the applications is for NQF #2902 and focuses on use of most and moderately effective contraceptive methods in a key sub-population of women at risk of unintended pregnancy: postpartum women. The other application is for NQF #2904 and focuses on use of a sub-set of contraceptive methods, i.e., use of long-acting reversible contraception (LARC); the goal of this measure to monitor whether women have access to LARC methods as determined by whether any units report very low levels of LARC use (e.g., less than 1-2 percent).

We also wish to acknowledge another measure with conceptual overlap to this measure: the Person-Centered Contraceptive Counseling (PCCC) measure (NQF #3543). Since 2017, OPA has met with an expert panel three times to discuss the appropriate use and interpretation of this measure in different health systems (e.g., programs with a reproductive health services focus compared to general health care providers). To ensure that healthcare systems employ a client-centered approach to implementation, the expert panel has recommended using this measure with a patient-reported outcome performance measure (PRO-PM) for contraceptive counseling.

OPA and our partners have not set a specific target for this measure with the purpose of discouraging coercion into use of contraception or a certain contraceptive method. We do not expect measure scores to reach 100% because some women will make informed decisions to choose less effective contraception, even when offered the full range of methods and with financial or logistical barriers to access removed. After NQF endorsed the contraceptive provision measures, OPA demonstrated its commitment to patient-centered contraceptive care by providing funding to the University of California San Francisco (UCSF) to develop a PRO-PM as a 'balancing measure' to support proper utilization of all contraceptive provision measures, and to enable health facilities and systems to assess patient experience in its own right. Following the initial year of support, UCSF secured private funding to continue the project.

Recently endorsed in November 2020 by NQF's Consensus Standards Approval Committee as NQF #3543, the Person-Centered Contraceptive Counseling (PCCC) measure is a four-item PRO-PM designed to specifically evaluate the patient-centeredness of contraceptive counseling at the individual clinician/provider and facility

levels of analysis. The PCCC's target population intersects with this measure's target population (e.g., ages 15-45 and assigned female at birth), but the PCCC is visit-specific. It is given to patients who have been identified as having received contraceptive counseling during their visit. A multi-organization partnership led by UCSF and the National Association of Community Health Centers (NACHC) has started research to test the PCCC and NQF #2903 in tandem use.

We share UCSF's hypothesis that the PCCC will serve as a balancing measure for the provision measures. After implementing the PCCC, organizations can observe any fluctuations in PCCC scores that occur with variations in provision scores. Ideally, increased contraceptive provision would be linked with improved patient experience. PCCC scores used in tandem with this measure allow groups to ensure that any increased contraceptive provision does not come at the cost of patient experience. Use of these two types of measures together can result in a more complete understanding of contraceptive care quality and help health care organizations to provide both access to a range of contraceptive methods and patient-centered counseling without coercion.

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.

Attachment **Attachment:** Appendices_for_2903_2021-04-27-final.docx

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): HHS Office of Population Affairs

Co.2 Point of Contact: Jamie, Kim, Jamie.Kim@hhs.gov, 240-453-2817-

Co.3 Measure Developer if different from Measure Steward: HHS Office of Population Affairs

Co.4 Point of Contact: Jamie, Kim, Jamie.Kim@hhs.gov, 240-453-2817-

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.

OPA convenes an expert work group (EWG) for the three contraceptive care measures: NQF #2902, NQF #2903, NQF #2904. The EWG represents several organizations and helps to develop the measure. EWG members' roles included calculating measure numerators and denominators, describing their organizations' activities supporting access to client-centered contraceptive care, and providing input on the measure implementation, interpretation, specifications, and code sets. EWG members over the past three years have included the following organizations and their staff:

HHS Office of Population Affairs: Amy F. Farb PhD, Diane Foley MD FAAP

HHS Centers for Disease Control and Prevention (CDC) Division of Reproductive Health: Jiajia Chen PhD, Shanna Cox MSPH, Ekwutosi Okoroh, MD MPH, Antoinette Nguyen MD MPH FACOG, Lisa Romero PhD

HHS CDC National Center for Health Statistics: Gladys Martinez PhD, Kimberly Daniels PhD

HHS Health Resources and Services Administration: Rui Li PhD

Iowa Department of Public Health and Iowa Medicaid Enterprise: Debra Kane PhD, Lindsey Jones MHA, Mark McMahon, Robert Schlueter, Kelly Garcia MPA, Gerd Clabaugh (retired)

Planned Parenthood Federation of America: Monika Grzeniewski MPH, Mark Bronstein

NewYork-Presbyterian Hospital/Columbia University Irving Medical Center: Nancy Fang MD, Carolyn Westhoff MD MSc

Washington State Department of Human Services and Health Care Authority: Dorothy Lyons, Joyce Fan PhD, Amanda Avalos MPA

Massachusetts Department of Public Health and MassHealth: Paul B. Kirby MA, Linda C. Shaughnessy MBA, Monica Le MD MPH, Susan E. Manning MD MPH

Louisiana Department of Health and Louisiana Medicaid: Lyn Kieltyka PhD MPH, Kolynda Parker MHS MLS(ASCP)CM CPHQ CLSSGB, Marcus Bachhuber, Larry Humble, Eddy Meyers, Amanda Dumas

HHS Centers for Medicaid & Medicare Services, Center for Medicaid and CHIP Services:

Renee E. Fox MD FAAP

Lekisha Daniel-Robinson MPH, IBM Watson Health

Elizabeth Jones MPA, National Family Planning & Reproductive Health Association

Research Triangle Institute: Christina I. Fowler PhD, Julia Gable, Beth Lasater, Kat Asman MSPH

Mathematica Policy Research: Emily N. Hoe MPA PMP; Margo Rosenbach PhD

University of Michigan Department of Obstetrics and Gynecology: Michelle H. Moniz MD MSc

University of California San Francisco Person-Centered Reproductive Health Program: Christine E Dehlendorf MD MAS, Ilana Silverstein

National Contraceptive Quality Measures Workgroup

OPA's statistical support contractor, Far Harbor LLC, completed reliability, data element and score level validity analyses for the application. Far Harbor's team includes Philip A. Hastings PhD, Fei Dong PhD, Antonio Garcia PhD, Ella d. Puga MPH, and Denise Wheeler MS.

Along with UCSF representatives, the following original measure developers also reviewed and offered suggestions on the NQF application: Brittni N. Frederiksen MPH PhD, Emily J. Decker MPH, Lorretta E. Gavin PhD.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2016

Ad.3 Month and Year of most recent revision: 03, 2020

Ad.4 What is your frequency for review/update of this measure? Every 3 years for Maintenance Endorsement

Ad.5 When is the next scheduled review/update for this measure? 04, 2021

Ad.6 Copyright statement: Not applicable.

Ad.7 Disclaimers: Not applicable.

Ad.8 Additional Information/Comments: Not applicable.