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 #: 2904

Corresponding Measures:

De.2. Measure Title: Contraceptive Care - Access to LARC

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

De.3. Brief Description of Measure: Percentage of women aged 15-44 years at risk of unintended pregnancy that is provided a long-acting reversible method of contraception (i.e., implants, intrauterine devices or systems (IUD/IUS)).

It is an access measure because it is intended to identify very low rates (less than 1-2%) of long-acting reversible methods of contraception (LARC), which may signal barriers to LARC provision.

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 [3-6]. 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 [26]. Not using any method at all has a typical failure rate of 85% [4].

Existing research shows that many women will select LARC methods if given the opportunity. Studies indicate that younger women who prefer LARC methods are not using them, signaling unmet demand [16, 17]; another analysis of the National Survey for Family Growth noted that women ages 15-24 and 25-34 were more likely to report cost as a barrier to use of their preferred contraceptive method [18]. In one large prospective study, 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 [5, 19-20]. High rates of LARC use were also found in a cluster randomized trial of a contraceptive counseling intervention, with more women enrolled in the intervention choosing a LARC method than those in the comparison group (28% vs 17%) [6].

Despite their effectiveness at preventing pregnancy and many women's preference for them, provider-related barriers to LARC access persist. A recent national survey of obstetrics-gynecology residents found that 41% had low long-acting reversible contraception insertion experience (i.e., zero implants inserted and/or 10 or fewer IUDs placed), although experience increased with more years of residency completed [24]. Another survey of obstetricians and gynecologists found while most respondents provide IUDs, only 29% offered same-day placement, and less than 25% offered immediate postpartum LARC to clients, which are not in-line with current clinical guidelines [25].

Although LARC methods are safe and effective, special concerns are present that affect how this performance measure should be implemented. The United States has a long history of coercive practices with regard to contraception, in which disadvantaged and minority women were forced to use sterilization and/or long-acting methods of contraception [22, 23]. Setting a high benchmark for a clinical performance measure for LARC methods could cause great harm by incentivizing providers to overly promote the use of LARC over other methods and discourage use of the client-centered counseling approach jointly recommended by the Centers for Disease Control and Prevention (CDC) and Office of Population Affairs (OPA) [11].

After NQF endorsed #2904 in 2016, OPA published multiple articles in peer-reviewed journals to inform providers 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 #2904 alongside its two complementary measures (NQF #2902 and #2903) and emphasize 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 LARC use and positive reproductive health outcomes demonstrates the importance of contraceptive care measures to health care quality [27-29]. As measure steward, OPA recommends that the performance measure focus on low (rather than high) rates of use to evaluate women's LARC access. For example, if a reporting entity has no or very few women using LARC (e.g., less than 2%), barriers restricting LARC access might be present and should be investigated. Another way to identify potential obstacles is to compare performance across several reporting units and consider whether barriers to access are present among the units with LARC use rates of less than 2%.

We emphasize that NQF #2904 should not be used to encourage high LARC utilization rates nor in a pay-for-performance context. The goal of providing contraception should never be to recommend any one method or class of methods over women's individual choices. Women who wish to delay or prevent pregnancy should have access to a wide variety of contraceptive methods, including LARC. 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 American Academy of Pediatrics (AAP), ACOG, and CDC, and OPA recommendations [7-12, 21] 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] Potter, J. E., Hopkins, K., Aiken, A. R., Hubert, C., Stevenson, A. J., White, K., & Grossman, D. (2014). Unmet demand for highly effective postpartum contraception in Texas. Contraception, 90(5), 488–495. https://doi.org/10.1016/j.contraception.2014.06.039

[17] Potter, J. E., Coleman-Minahan, K., White, K., Powers, D. A., Dillaway, C., Stevenson, A. J., Hopkins, K., & Grossman, D. (2017). Contraception After Delivery Among Publicly Insured Women in Texas: Use Compared With Preference. Obstetrics and gynecology, 130(2), 393–402. https://doi.org/10.1097/AOG.0000000000002136

[18] Burke, K. L., Potter, J. E., & White, K. (2020). Unsatisfied contraceptive preferences due to cost among women in the United States. Contraception: X, 2, 100032. https://doi.org/10.1016/j.conx.2020.100032

[19] Diedrich, J.T., et al., Three-year continuation of reversible contraception. Am J Obstet Gynecol, 2015. 213(5): p. 662 e1-8.

[20] O´Neil-Callahan, M., et al., Twenty-four-month continuation of reversible contraception. Obstet Gynecol, 2013. 122(5): p. 1083-91.

[21] Committee on Practice Bulletins-Gynecology, Long-Acting Reversible Contraception Work Group (2017). Practice Bulletin No. 186: Long-Acting Reversible Contraception: Implants and Intrauterine Devices. Obstetrics and gynecology, 130(5), e251–e269. https://doi.org/10.1097/AOG.0000000000002400

[22] Gold, J., Guarding Against Coercion While Ensuring Access: A Delicate Balance. Guttmacher Policy Review, 2014. 17(3).

[23] 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

[24] Maples, J. M., Espey, E., Evans, M. L., Breeze, J. L., Ogburn, T., & Zite, N. B. (2020). Obstetrics-gynecology resident long-acting reversible contraception training: the role of resident and program characteristics. American journal of obstetrics and gynecology, 222(4S), S923.e1–S923.e8. https://doi.org/10.1016/j.ajog.2019.12.007

[25] Castleberry, N. M., Stark, L., Schulkin, J., & Grossman, D. (2019). Implementing best practices for the provision of long-acting reversible contraception: a survey of obstetrician-gynecologists. Contraception, 100(2), 123–127. https://doi.org/10.1016/j.contraception.2019.03.053

[26] 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

[27] 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

[28] 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

[29] 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 were provided a long-acting reversible method of contraception (LARC), i.e., intrauterine device or implant.

S.6. Denominator Statement: Women ages 15-44 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) women who had a live birth in the last 2 months of the measurement year; or (3) women who were still pregnant or their pregnancy outcome was unknown at the end of the measurement year.

De.1. Measure Type: Structure

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:

- Provision of most and moderately effective methods of contraception – The primary goal of this intermediate outcome measure is to monitor the percentage of women of reproductive age who are at risk of unintended pregnancy that are provided the most and moderately effective methods of contraception.
- Postpartum women – this is a very important sub-population of all women at risk of unintended pregnancy. Contraceptive care 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.

Preliminary Analysis: Maintenance of Endorsement

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

Criteria 1: Importance to Measure and Report

1a. Evidence

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

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

The developer provides the following evidence for this measure:

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

Evidence Summary or **Summary of prior review in [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. The developer reported that the evidence showed support for the most effective or long acting reversible contraceptive (LARC) methods 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 an updated and robust number of guidelines and a conceptual framework in support of the measure. These included guidelines from the Centers for Disease Control

(CDC), the HHS Office of Population Affairs (OPA), American College of Obstetricians and Gynecologists (ACOG), and the Health Resources and Services Administration (HRSA).

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 #2904, 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

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 eight 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: 4.2 and Ages 21-44: 4.8
        - FFY 2017 Median Measure Scores: Ages 15-20: 4.9 and Ages 21-44: 5.7
        - FFY 2018 Measure Scores Ages 15-20 Median: 5.4, Range: 1.0 - 11.3
        - FFY 2019 Measure Scores Ages 15-20 Median: 4.8, Range: 1.1 - 20.0

- Performance scores are not reported by moderate and most, rather as overall median or mean performance. Although #2904 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, #2904 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

- As measure steward, OPA states on our website that NQF #2904 "should be used as an access measure; very low rates (less than 1-2%) may signal barriers to LARC provision that should be addressed through training … [and] and quality improvement processes".
- 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. The results showed indicated that no race/ethnicity group had measure scores less than 2%, suggesting that PPFA clients may have access to LARC - African American: 9.51, Alaskan Native: 26.15, Asian: 14.83, Hispanic: 16.90, Multi-racial: 17.37, Native American: 16.78, Pacific Islander: 15.27, White: 16.84, and Other race: 15.14.
- Washington State Health Care Authority (WA HCA) from 2014-2018 scores for female clients ages 15-44 by age group and race/ethnicity (https://www.hca.wa.gov/assets/program/ccw-contraceptive-care.pdf). Among adults ages 21-44, all race/ethnicity groups had LARC provision rates greater than 2% during this five-year period. Except for 2014, LARC provision rates in clients ages 15-20 were also more than 2% for all race/ethnicity groups.
- For 2018, WA HCA found aged 15-20 females with receiving LARC were Hispanic: 5.2, White: 7.0, Asian: 3.1, Black: 4.5, American Indian/Alaska Native: 7.3, Hawaiian/Pacific Islander: 4.2, More than One Race: 6.9, and Other/Unknown: 4.0. For ages 21-44, Hispanic: 8.8, White: 5.9, Asian: 5.8, Black: 6.4, American Indian/Alaska Native: 5.7, Hawaiian/Pacific Islander: 5.3, More than One Race: 6.5, and Other/Unknown: 5.9.

Questions for the Committee:

- Was the performance gaps and disparities data available by 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
- As is laid out in the NQF measure evaluation worksheet, the data and evidence are robust. I do think some questions laid out are important for us to discuss as a committee: "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 #2904, or patients with live or not live births in the last two months of the measurement period where contraceptives may be applicable?" I think these patients should be included. I agree that diaphragm should be removed from "moderately effectively." "If derived from patient report, does the target population value the measured process or structure and find it meaningful?" I would like to discuss this last question as a committee. I don't think this is "patient report" because the measure looks at claims data in the numerator. Can we just change to 2% since the data given in 1b show that all race/ethnicity subsets had much higher LARC rates? Although I think for consistency, just sticking to 1% would be fine too.
- evidence acceptable
- Maintenance measure updated evidence was submitted with high evidence.

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
- Yes - very clear need for this measure as outlined in the measure evaluation worksheet.
- yes gap demonstrated
- Less than half the states report data and disparities are noted. Population subgroup data was provided. Data could indicate access to care opportunities.

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 *#2903* Contraceptive Care – Most & Moderately Effective Methods 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), New York 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

- 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 stated the correlation magnitude may be weak for cervical cancer screening and chlamydia screening 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 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.

- 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, "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* and *#2903* 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 re-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
- If we go with 1%, shouldn't a clinic have at least a denominator of 100 to capture at least 1%? That would be the minimum sample size too, right? I would like to hear from the measure developer on this, and for the committee to discuss/vote on/recommend to the developer?
- ok
- Elements are clearly defined. No concerns.

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

- no concerns
- I would like to hear from the developer about minimum sample size.
- no
- Some concerns

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

- no concerns
- I think lowering to 1% (instead of 1-2%) would allow health systems to feel like this measure is valid.
- no
- Some concerns. Not risk-adjusted

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
- No response from a health system would mean that they would get 0% of patients who use LARC, which might not be valid - but it could also be valid. We wouldn't know without the data.
- none
- Screening frequency differences and possible weak correlations

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-

- I don't think that people who were pregnant in the past year should be excluded - they should also all be offered LARC, and if only 1% uptake LARC, then they have been offered the choice, which is what we want/the data show we should be doing as clinicians for best patient outcomes..
- not sure why need to exclude deliveries other than live births
- Consistent with the evidence

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
    - I have no concerns about feasibility. I'm glad to know that there is ongoing work to develop an eCQM version of this measure
    - ok
    - No concerns

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 Maternal and Infant Health Initiative, https://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Quality-of-Care/Maternal-and-Infant-Health-Care-Quality.htmlIowa 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
- Washington State Health Care Authority, https://www.hca.wa.gov/about-hca/reproductive-health
- OPA Title X Family Planning Program, https://rhntc.org/resources/contraceptive-access-change-package and
- OPA 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

The developer states the measure is included the NQF Core Quality Measure Collaborative (CQMC) project led by CMS/AHIP at the clinician/group level in outpatient settings. http://www.qualityforum.org/cqmc/. 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

- 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.
- The Generic Product Identifier (GPI) code system requires a license fee to utilize, which may not be possible for all states calculating NQF #2904 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.

- 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 [Impact/trends over time/improvement]

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 LARC 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**:*

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

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

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

4a1. Use - Accountability and Transparency: 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!
- "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." I think it should be reported as either ">1%" or "does not meet the measurement criteria (<1%) so that clinics/health systems/affiliates are not focused on the numbers above 1%
- ok
- Is currently publicly reported.

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
- "Although not yet tested in pregnant patients," But if these are from claims data, you can't put a LARC into a pregnant patient. As is, for this measure, benefits outweigh harms. I think going down to 1% would make this easier.
- concern re overzealous prescribing
- No unintended consequences.

Criterion 5: Related and Competing Measures

Related or competing measures

- 1517: Prenatal & Postpartum Care (PPC)
- 2902: Contraceptive Care – Postpartum
- 2903: Contraceptive Care – Most & Moderately Effective Methods
- 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 #2903 are complementary to this measure.
    - #2902 focuses on most or moderately effective contraceptive provision in all women of who had a live birth
    - #2903 focuses on moderate and most effective contraceptive provisions for 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
- As written out.
- yes other BC measures
- none

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: 2904

Measure Title: Contraceptive Care - Access to LARC

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 2: Although the developer provided some testing results at group/practice level, the measure is not specified for use at group/practice level in the testing form. This is sensible but needs to be clear to measure users.

Panel Member 3: Data dictionary not available: NQF_2904_Codes_2021.xlsx. It was not clear to me how the calculated rates were used to compute the measure score. In the rationale, a 2% threshold was recommended. Was this the threshold used, i.e., less than 2% flagged a negative performance? This needs to be clarified.

Panel Member 4: No concerns.

Panel Member 5: No Concerns

Panel Member 6: None

Panel Member 7: How reliably can one identify "at risk of unintended pregnancy" - a rhetorical question or an empirical one?

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 3: Chart abstract from clinical records for data element validity testing
Panel Member 5: Chart abstract from clinical records for data element validity testing
Panel Member 7: Chart abstract 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 method. Calculated a signal-to-noise ratio for each level of analysis.

   Panel Member 2: The developer estimated measure score reliability via beta-binomial model using parametric empirical Bayes methods.

   Panel Member 3: No concerns

   Panel Member 4: Used 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 used for testing, Planned Parenthood Federation of America, Iowa Medicaid, Enterprise, Iowa Department of Public Health, New York Presbyterian/Columbia University, Washington State Health Care Authority, Massachusetts MassHealth, and Louisiana Medicaid. Testing was performed at the facility, public health region, group billing provider and health plan levels. Reliability for this measure as a signal to noise approach was estimated from a beta-binomial model applied to each level

   Panel Member 7: 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. The distribution of reliability across entities was not described. Note: "The measure steward, OPA recommends that the performance measure focus on low (rather than high) rates of use to evaluate women's LARC access. For example, if a reporting entity has no or very few women using LARC (e.g., less than 2%), barriers restricting LARC access might be present and should be investigated." No analysis was conducted on the reliability of being classified as a low outlier.
7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

Panel Member 1: The tested reliability is greater than .70 at the facility and health plan levels and consistently greater than .90 at the public health region level. At the group level, estimates were above 0.70 if the measure is restricted to practices with >75 patients.

Panel Member 2: Measure reliability scores were in general very high at facility, public health region, and health plan level. At group provider level, measure reliability scores were low for provider with low case counts; when a 75 unit size limit was imposed, measure reliability scores improved substantially.

Panel Member 3: No concerns

Panel Member 4: Adequate.

Panel Member 5: Reliability is greater than .70 at the facility and health plan levels and consistently greater than .90 at the public health region level for the 15-44 age group, showing adequate to high reliability

Panel Member 6: At the facility level, reliability was more than .70 up to .978 for most levels at all age groups. It was less than .2 for group billing providers at all age groups. From two health plans, Louisiana Medicaid and Washington, it was between .4 and .6 for three age groups. The utilization of the unit size greater than 75 generally increased the reliability estimate and greatly increased it to more than .70 for the group billing provider level.

Panel Member 7: Broadly, yes, for level with N greater than provider.

Panel Member 8: In general, and especially for entities with >75 women, average reliabilities were high. No analysis of the reliability of low outlier classification.

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 score-level reliability. The statistics were high (>0.70) for all levels, including clinician groups (if restricted to groups with 75+ patients).

Panel Member 2: Measure reliability scores were very high except at group provider level. As this measure is not specified for use at group level, this is not critical.

Panel Member 4: Consistent results at the facility, heath plan and regional levels.

Panel Member 6: The reliability estimate was high for most of the levels and age groups, but moderate for some others.

Panel Member 8: Overall reliability appears to be excellent but no description of the distribution of reliabilities was provided nor the reliability out low outlier status (<2%).

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: Used an 8 person independent expert panel to assess whether the measure would reflect the quality of contraceptive services. Empirical validity: Tested whether performance on this measure was correlated with other measures of women's health services.

    Panel Member 2: The developer tested data element validity by evaluating if critical data elements can be captured by codes accurately. The developer assessed measure score validity by correlating this measure with other similar quality measures including cervical cancer screening, chlamydia screening, encounter for gynecological exam, and encounter for contraceptive counseling.

Panel Member 3: No concerns

Panel Member 4: Reasonable approach. Acceptable.

Panel Member 5: Tested for convergent validity of the most or moderately effective contraceptive measure by exploring whether it was correlated with other similar quality measures. Hypothesized that facilities/groups that perform well on contraceptive care should perform well on cervical cancer screening, chlamydia screening, contraceptive counseling, and gynecological exams.

Panel Member 6: Validity testing was performed by correlation with other quality measures, specifically cervical cancer screening, chlamydia screening, encounter for contraceptive counseling, and encounter for gynecologic exam. It is hypothesized that a provider who performs well on these will also perform well on this measure. Pearson correlation coefficient was performed and in addition, to mitigate the effect of non-linearity, a logic transformation of the binomial proportions was performed. At the facility level, Pearson correlation coefficient ranged from .23 to .78 across all age groups, the highest correlation was with a gynecological examination. Using a multilevel correlation estimation, the range across all groups was .78 to .98. At the group provider billing level, Pearson correlation coefficient ranged from .08 to .67 across all age groups, the highest correlation was with a contraceptive counseling. Using a multilevel correlation estimation, the range across all groups was .06 to .67. Critical data elements validity testing was performed and were assessed for sensitivity, specificity, PPV, NPV, %agreement, and kappa. The Kappa statistic ranged from .567 to 1.000.

Panel Member 7: Pearson with over GYB/OB items (cervical cancer screening, chlamydia screening, contraceptive counseling, and gynecological exams). For Data, (For each of the 6 Iowa Department of Public Health Title X Grantee 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)

Panel Member 8: Correlation analyses (both standard and improved) of the measure with similar measures. They hypothesized that facilities/groups that perform well on contraceptive care should perform well on cervical cancer screening, chlamydia screening, contraceptive counseling, and gynecological exams.

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, whereas specificity, PPV, and NPV were above 0.9 for all data elements. Percent agreement was consistently over 95%. They also observed statistically significant Kappa above 0.6 for all data elements, indicating moderate to almost perfect agreement between the claims records and the patient charts. Face validity: The mean rating for this measure was 4.33 with a median of 4.5 (between Agree and Strongly Agree), range 3-5. Empirical validity: Found weak to moderate positive correlations with other measures of women's health services.

    Panel Member 2: There are positive correlations between this measure and 4 related quality measures. Given than this measure is more designed to identify very low rate, higher rate may not necessarily be better, it is not easy to interpret the results. The results of critical data elements testing are in general good, although sensitivity for two critical data elements is somewhat low. For example, for live birth data element, for age 21-33 group, the sensitivity is only 0.40.

Panel Member 3: Data element validity results were very good. Empirical validity results were as expected, i.e., with weak to strong 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: The results of validity testing demonstrated a wide range of validity testing, generally in the low to moderate range. The utilization of the alternative multilevel correlation estimation generally improved the correlation results.

Panel Member 7: OK - poor sensitivity and kappa for "live birth in the past 2 mos." and "currently pregnant or unknown pregnancy outcome" although not clear to me whether this makes a big difference given reliability statistics.

Panel Member 8: The results generally support these hypotheses. More validity analysis could be done with the intended use of the measure (to identify low outliers.)

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 exclusions are logical (e.g., patient is infecund) and/or operationally needed (had live birth in the last 2 months of measurement year).

Panel Member 2: Low sensitivity for live birth data element is concerning as it is used to establish one exclusion criterion. Using WA HCA health plans 2019 data, live birth exclusion was around 0.9%, using IME public health regions 2018 data, live birth exclusion was around 4.6%. Using PPFA health center 2019 data, live birth exclusion was 0%. Potentially these differences may reflect the reliability of that data element.

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: No concerns.

Panel Member 6: None

19. **Risk Adjustment**

Submission Document: Testing attachment, section 2b3

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

Panel Member 4: Recommend by age group for 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 developer believe any variation is due to modifiable clinical and programmatic considerations and are not reflective of patient-level factors (race, age, SES, underlying health status).

Panel Member 2: The developer clearly articulated 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: Justification provided for no risk adjustment and no evidence contrary 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: The measure steward do not believe that risk adjustment is justified because, although variation exists for socio-demographic perspective, these are due to systematic structural issues, not biologic characteristics.

Panel Member 7: I still do not understand rationale for no SES/risk adjustment.

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 concern. There is variation in performance across the different levels, with the least variation in the health plans and population levels (which makes sense).

    Panel Member 2: Given the emphasis on using this measure to identify very low rate to uncover potential barriers for access to LARC, it is not clear how to interpret the rate differences among entities when rates were not lower than 2%.

    Panel Member 3: No concerns

    Panel Member 4: No concerns.

    Panel Member 5: Ability to identify differences needs to be taken in context of the units measured. As noted by Authors; 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

    Panel Member 6: None

    Panel Member 7: Differences may be due to patient choice (systematic by region) but this may not be a SMP issue.

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: The developer should provide clear guidance to measure users on how to interpret the results, particularly when they may intend to compare rates across settings. For example, mean rate for facility in PPFA was 0.135 while mean rate for facility in NYP was 0.072. For this measure, typical better or worse than average performance may not be an appropriate reporting method.

    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. As the measure is based on claims data, there is minimal missing data.

    Panel Member 3: No concerns

    Panel Member 4: No analysis of missing data.

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):**

    Panel Member 6: None.

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: Conducted both data element and score-level validity testing. The data element validity testing indicated high levels of agreement between the claim and chart. The empirical score-level validity indicated that the measure has a weak to moderate relationship with other measures of women's health services.

    Panel Member 2: Measure score validity results are not as clear cut given the nature of this measure. The sensitivity for live birth data element is somewhat concerning.

    Panel Member 3: The treats to validity that were identified drove the moderate rating.

    Panel Member 4: Appropriate methods but no analysis of missing data.

    Panel Member 6: The overall correlation estimates are a very wide range and a combination of low to significantly higher estimates, depending on the age group and the related measure chosen. They are increased by the use of an alternative estimation model.

    Panel Member 7: Similar concerns as with 2902 and 2903.

    Panel Member 8: Data element validity is very good. The entity-level results generally support the hypotheses.

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: No concerns.

Panel Member 6: None

Developer Submission

NQF #: 2904

Corresponding Measures:

De.2. Measure Title: Contraceptive Care - Access to LARC

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

De.3. Brief Description of Measure: Percentage of women aged 15-44 years at risk of unintended pregnancy that is provided a long-acting reversible method of contraception (i.e., implants, intrauterine devices or systems (IUD/IUS)).

It is an access measure because it is intended to identify very low rates (less than 1-2%) of long-acting reversible methods of contraception (LARC), which may signal barriers to LARC provision.

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 [3-6]. 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 [26]. Not using any method at all has a typical failure rate of 85% [4].

Existing research shows that many women will select LARC methods if given the opportunity. Studies indicate that younger women who prefer LARC methods are not using them, signaling unmet demand [16, 17]; another analysis of the National Survey for Family Growth noted that women ages 15-24 and 25-34 were more likely to report cost as a barrier to use of their preferred contraceptive method [18]. In one large prospective study, 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 [5, 19-20]. High rates of LARC use were also found in a cluster randomized trial of a contraceptive counseling intervention, with more women enrolled in the intervention choosing a LARC method than those in the comparison group (28% vs 17%) [6].

Despite their effectiveness at preventing pregnancy and many women's preference for them, provider-related barriers to LARC access persist. A recent national survey of obstetrics-gynecology residents found that 41% had low long-acting reversible contraception insertion experience (i.e., zero implants inserted and/or 10 or fewer IUDs placed), although experience increased with more years of residency completed [24]. Another survey of obstetricians and gynecologists found while most respondents provide IUDs, only 29% offered same-day placement, and less than 25% offered immediate postpartum LARC to clients, which are not in-line with current clinical guidelines [25].

Although LARC methods are safe and effective, special concerns are present that affect how this performance measure should be implemented. The United States has a long history of coercive practices with regard to contraception, in which disadvantaged and minority women were forced to use sterilization and/or long-acting methods of contraception [22, 23]. Setting a high benchmark for a clinical performance measure for LARC methods could cause great harm by incentivizing providers to overly promote the use of LARC over other methods and discourage use of the client-centered counseling approach jointly recommended by the Centers for Disease Control and Prevention (CDC) and Office of Population Affairs (OPA) [11].

After NQF endorsed #2904 in 2016, OPA published multiple articles in peer-reviewed journals to inform providers 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 #2904 alongside its two complementary measures (NQF #2902 and #2903) and emphasize 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 LARC use and positive reproductive health outcomes demonstrates the importance of contraceptive care measures to health care quality [27-29]. As measure steward, OPA recommends that the performance measure focus on low (rather than high) rates of use to evaluate women's LARC access. For example, if a reporting entity has no or very few women using LARC (e.g., less than 2%), barriers restricting LARC access might be present and should be investigated. Another way to identify potential obstacles is to compare performance across several reporting units and consider whether barriers to access are present among the units with LARC use rates of less than 2%.

We emphasize that NQF #2904 should not be used to encourage high LARC utilization rates nor in a pay-for-performance context. The goal of providing contraception should never be to recommend any one method or class of methods over women's individual choices. Women who wish to delay or prevent pregnancy should have access to a wide variety of contraceptive methods, including LARC. 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 American Academy of Pediatrics (AAP), ACOG, and CDC, and OPA recommendations [7-12, 21] 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] Potter, J. E., Hopkins, K., Aiken, A. R., Hubert, C., Stevenson, A. J., White, K., & Grossman, D. (2014). Unmet demand for highly effective postpartum contraception in Texas. Contraception, 90(5), 488–495. https://doi.org/10.1016/j.contraception.2014.06.039

[17] Potter, J. E., Coleman-Minahan, K., White, K., Powers, D. A., Dillaway, C., Stevenson, A. J., Hopkins, K., & Grossman, D. (2017). Contraception After Delivery Among Publicly Insured Women in Texas: Use Compared With Preference. Obstetrics and gynecology, 130(2), 393–402. https://doi.org/10.1097/AOG.0000000000002136

[18] Burke, K. L., Potter, J. E., & White, K. (2020). Unsatisfied contraceptive preferences due to cost among women in the United States. Contraception: X, 2, 100032. https://doi.org/10.1016/j.conx.2020.100032

[19] Diedrich, J.T., et al., Three-year continuation of reversible contraception. Am J Obstet Gynecol, 2015. 213(5): p. 662 e1-8.

[20] O´Neil-Callahan, M., et al., Twenty-four-month continuation of reversible contraception. Obstet Gynecol, 2013. 122(5): p. 1083-91.

[21] Committee on Practice Bulletins-Gynecology, Long-Acting Reversible Contraception Work Group (2017). Practice Bulletin No. 186: Long-Acting Reversible Contraception: Implants and Intrauterine Devices. Obstetrics and gynecology, 130(5), e251–e269. https://doi.org/10.1097/AOG.0000000000002400

[22] Gold, J., Guarding Against Coercion While Ensuring Access: A Delicate Balance. Guttmacher Policy Review, 2014. 17(3).

[23] 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

[24] Maples, J. M., Espey, E., Evans, M. L., Breeze, J. L., Ogburn, T., & Zite, N. B. (2020). Obstetrics-gynecology resident long-acting reversible contraception training: the role of resident and program characteristics. American journal of obstetrics and gynecology, 222(4S), S923.e1–S923.e8. https://doi.org/10.1016/j.ajog.2019.12.007

[25] Castleberry, N. M., Stark, L., Schulkin, J., & Grossman, D. (2019). Implementing best practices for the provision of long-acting reversible contraception: a survey of obstetrician-gynecologists. Contraception, 100(2), 123–127. https://doi.org/10.1016/j.contraception.2019.03.053

[26] 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

[27] 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

[28] 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

[29] 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 were provided a long-acting reversible method of contraception (LARC), i.e., intrauterine device or implant.

S.6. Denominator Statement: Women ages 15-44 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) women who had a live birth in the last 2 months of the measurement year; or (3) women who were still pregnant or their pregnancy outcome was unknown at the end of the measurement year.

De.1. Measure Type: Structure

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:

- Provision of most and moderately effective methods of contraception – The primary goal of this intermediate outcome measure is to monitor the percentage of women of reproductive age who are at risk of unintended pregnancy that are provided the most and moderately effective methods of contraception.
- Postpartum women – this is a very important sub-population of all women at risk of unintended pregnancy. Contraceptive care 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.

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

LARC_2904_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*)**:** #2904

Measure Title: Contraceptive Care – Access to LARC

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.

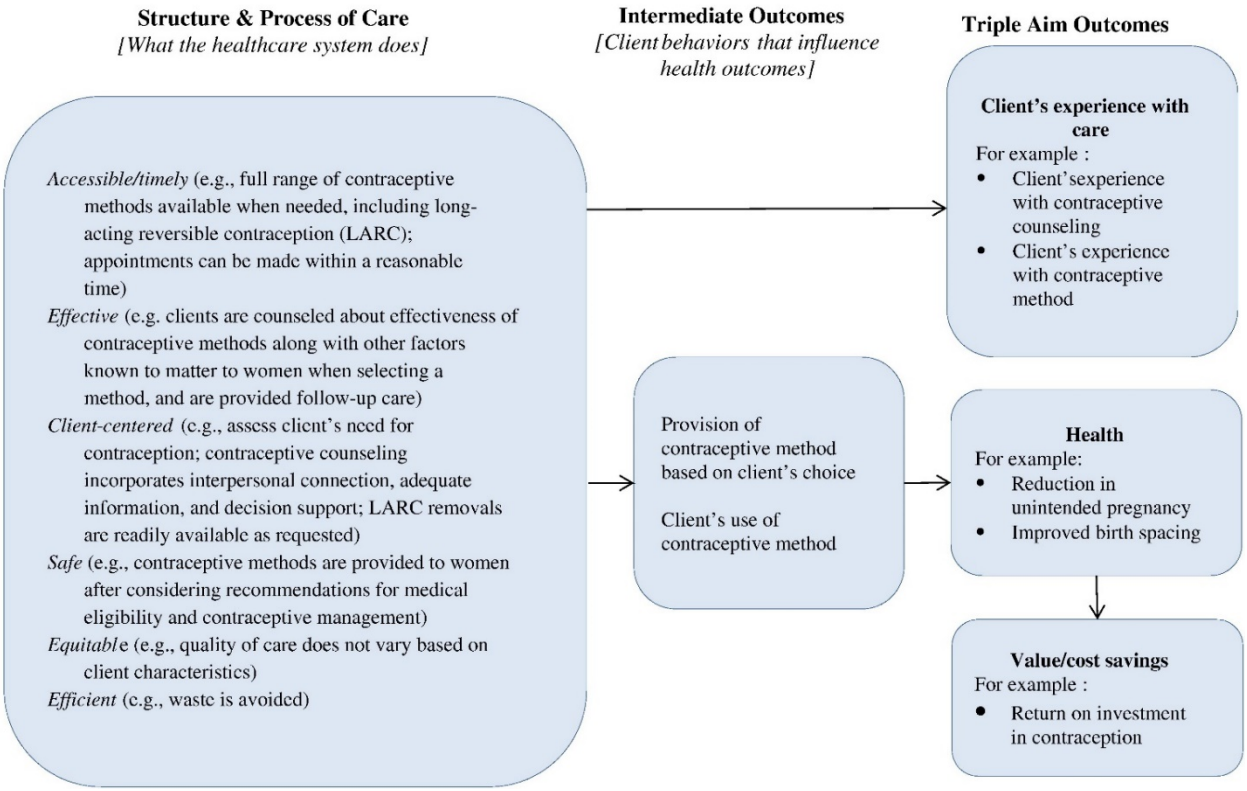
2021 Submission

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*)

☐ Othe

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

Systematic Review	Evidence
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. 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.

Systematic Review	Evidence
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 • 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%.

Systematic Review	Evidence
	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 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**

Systematic Review	Evidence
	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%). 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.

Systematic Review	Evidence
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	Evidece
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
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

Systematic Review	Evidece
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
	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
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:

Systematic Review	Evidece
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
	"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

Systematic Review	Evidece
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
	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
Systamtic Reviewq	**Evidence**
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 penicillin treatment in the 1940s) could also be regarded as this type of evidence.

Systematic Review	Evidece
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
	III Opinions of respected authorities, based on clinical experience, descriptive studies and case reports, or reports of expert committees GRADE The implications of a strong recommendation are: • For patients—most people in your situation would want the recommended course of action and only a small proportion would not; request discussion if the intervention is not offered • For clinicians—most patients should receive the recommended course of action • For policy makers—the recommendation can be adopted as a policy in most situations. The implications of a weak recommendation are: • For patients—most people in your situation would want the recommended course of action, but many would not • For clinicians—you should recognize that different choices will be appropriate for different patients and that you must help each patient to arrive at a management decision consistent with her or his values and preferences

Systematic Review	Evidece
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
	• For policy makers—policy making will require substantial debate and involvement of many stakeholders.
Grade assigned to the **recommendation** with definition of the grade	A multistage process was used to develop the recommendations that drew on established procedures for developing clinical guidelines. First, an Expert Work Group was formed comprising family planning clinical providers, program administrators, and representatives from relevant federal agencies and professional medical associations to help define the scope of the recommendations. Next, literature about three priority topics (i.e., counseling and education, serving adolescents, and quality improvement) was reviewed by using the USPSTF methodology for conducting systematic reviews. The results were presented to three technical panels comprising subject matter experts (one panel for each priority topic) who considered the quality of the evidence and made suggestions for what recommendations might be supported on the basis of the evidence. In a separate process, existing clinical recommendations on women's and men's preventive services were compiled from more than 35 federal and professional medical associations, and these results were presented to two technical panels of subject matter experts, one that addressed women's clinical services and one that addressed men's clinical services. The panels provided individual feedback about which clinical preventive services should be offered in a family planning setting and which

Systematic Review	Evidece
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
	clinical recommendations should receive the highest consideration. CDC and OPA used the input from the subject matter experts to develop a set of core recommendations and asked the Expert Work Group to review them. The members of the Expert Work Group were more familiar with the family planning service delivery context than the members of the Technical Panel and thus could better comment on the feasibility and appropriateness of the recommendations, as well as the supporting evidence. The Expert Work Group considered the core recommendations by using the following criteria: 1) the quality of the evidence; 2) the positive and negative consequences of implementing the recommendations on health outcomes, costs or cost-savings, and implementation challenges; and 3) the relative importance of these consequences, (e.g., the likelihood that implementation of the recommendation will have a substantial effect on health outcomes might be considered more than the logistical challenges of implementing it). In certain cases, when the evidence from the literature reviews was inconclusive or incomplete, recommendations were made on the basis of expert opinion. Finally, CDC and OPA staff considered the individual feedback from Expert Work Group

Systematic Review	Evidece
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
	members when finalizing the core recommendations and writing the recommendations document. Summary can be found in Appendix B of the 2014 QFP (p. 35-44).
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.
Body of evidence: • Quantity – how many 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.

Systematic Review	Evidece
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
• Quality – what type of studies?	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*

Systematic Review	Evidece
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
	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.
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	Evidece
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
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

Systematic Review	Evidece
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
	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 • 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: (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.

Systematic Review	Evidence
	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. 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-

Systematic Review	Evidence
	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. 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.

Systematic Review	Evidence
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 • 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

Systematic Review	Evidence
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 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.
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.

Systematic Review	Evidence
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).
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*)**:** Click here to enter NQF number

Measure Title: Contraceptive Care – Postpartum

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: **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.*

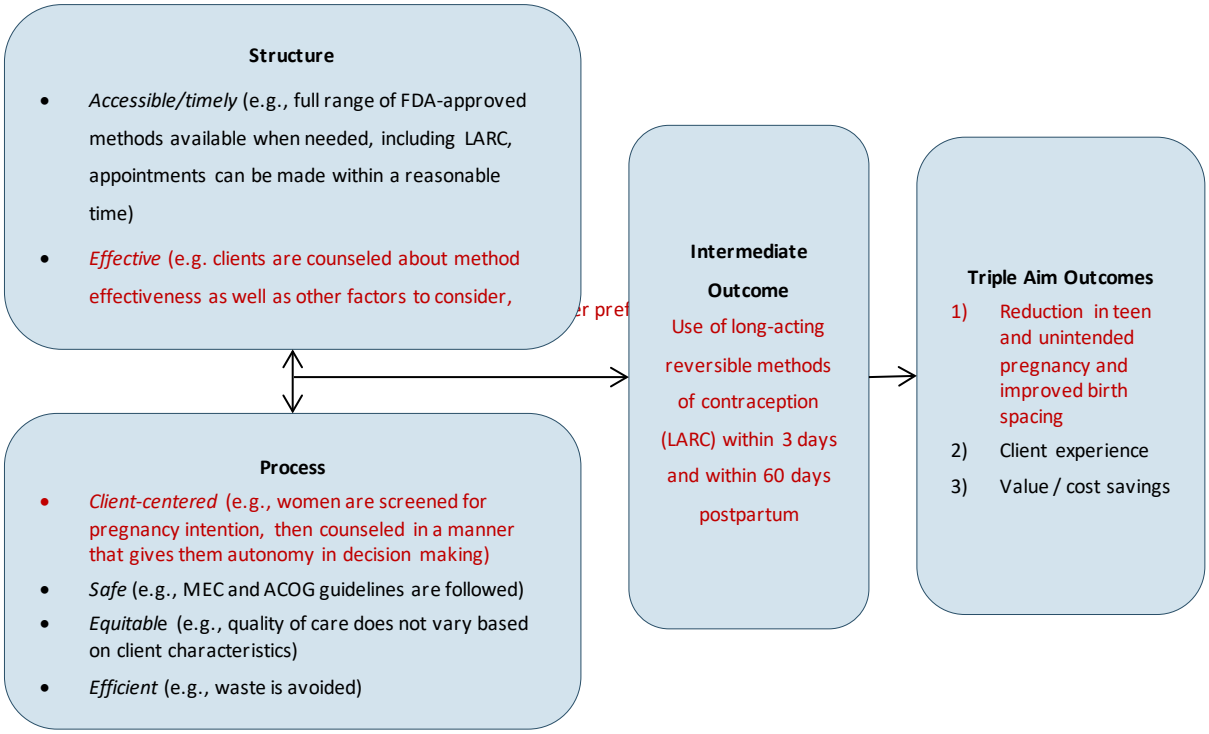
INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

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 [3-6]. 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 [26]. Not using any method at all has a typical failure rate of 85% [4].

Existing research shows that many women will select LARC methods if given the opportunity. Studies indicate that younger women who prefer LARC methods are not using them, signaling unmet demand [16, 17]; another analysis of the National Survey for Family Growth noted that women ages 15-24 and 25-34 were more likely to report cost as a barrier to use of their preferred contraceptive method [18]. In one large prospective study, 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 [5, 19-20]. High rates of LARC use were also found in a cluster randomized trial of a contraceptive counseling intervention, with more women enrolled in the intervention choosing a LARC method than those in the comparison group (28% vs 17%) [6].

Despite their effectiveness at preventing pregnancy and many women's preference for them, provider-related barriers to LARC access persist. A recent national survey of obstetrics-gynecology residents found that 41% had low long-acting reversible contraception insertion experience (i.e., zero implants inserted and/or 10 or fewer IUDs placed), although experience increased with more years of residency completed [24]. Another survey of obstetricians and gynecologists found while most respondents provide IUDs, only 29% offered same-day placement, and less than 25% offered immediate postpartum LARC to clients, which are not in-line with current clinical guidelines [25].

Although LARC methods are safe and effective, special concerns are present that affect how this performance measure should be implemented. The United States has a long history of coercive practices with regard to contraception, in which disadvantaged and minority women were forced to use

sterilization and/or long-acting methods of contraception [22, 23]. Setting a high benchmark for a clinical performance measure for LARC methods could cause great harm by incentivizing providers to overly promote the use of LARC over other methods and discourage use of the client-centered counseling approach jointly recommended by the Centers for Disease Control and Prevention (CDC) and Office of Population Affairs (OPA) [11].

After NQF endorsed #2904 in 2016, OPA published multiple articles in peer-reviewed journals to inform providers 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 #2904 alongside its two complementary measures (NQF #2902 and #2903) and emphasize 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 LARC use and positive reproductive health outcomes demonstrates the importance of contraceptive care measures to health care quality [27-29]. As measure steward, OPA recommends that the performance measure focus on low (rather than high) rates of use to evaluate women's LARC access. For example, if a reporting entity has no or very few women using LARC (e.g., less than 2%), barriers restricting LARC access might be present and should be investigated. Another way to identify potential obstacles is to compare performance across several reporting units and consider whether barriers to access are present among the units with LARC use rates of less than 2%.

We emphasize that NQF #2904 should not be used to encourage high LARC utilization rates nor in a pay-for-performance context. The goal of providing contraception should never be to recommend any one method or class of methods over women's individual choices. Women who wish to delay or prevent pregnancy should have access to a wide variety of contraceptive methods, including LARC. 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 American Academy of Pediatrics (AAP), ACOG, and CDC, and OPA recommendations [7-12, 21] 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] Potter, J. E., Hopkins, K., Aiken, A. R., Hubert, C., Stevenson, A. J., White, K., & Grossman, D. (2014). Unmet demand for highly effective postpartum contraception in Texas. Contraception, 90(5), 488–495. https://doi.org/10.1016/j.contraception.2014.06.039

[17] Potter, J. E., Coleman-Minahan, K., White, K., Powers, D. A., Dillaway, C., Stevenson, A. J., Hopkins, K., & Grossman, D. (2017). Contraception After Delivery Among Publicly Insured Women in Texas: Use

Compared With Preference. Obstetrics and gynecology, 130(2), 393–402. https://doi.org/10.1097/AOG.0000000000002136

[18] Burke, K. L., Potter, J. E., & White, K. (2020). Unsatisfied contraceptive preferences due to cost among women in the United States. Contraception: X, 2, 100032. https://doi.org/10.1016/j.conx.2020.100032

[19] Diedrich, J.T., et al., Three-year continuation of reversible contraception. Am J Obstet Gynecol, 2015. 213(5): p. 662 e1-8.

[20] O´Neil-Callahan, M., et al., Twenty-four-month continuation of reversible contraception. Obstet Gynecol, 2013. 122(5): p. 1083-91.

[21] Committee on Practice Bulletins-Gynecology, Long-Acting Reversible Contraception Work Group (2017). Practice Bulletin No. 186: Long-Acting Reversible Contraception: Implants and Intrauterine Devices. Obstetrics and gynecology, 130(5), e251–e269. https://doi.org/10.1097/AOG.0000000000002400

[22] Gold, J., Guarding Against Coercion While Ensuring Access: A Delicate Balance. Guttmacher Policy Review, 2014. 17(3).

[23] 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

[24] Maples, J. M., Espey, E., Evans, M. L., Breeze, J. L., Ogburn, T., & Zite, N. B. (2020). Obstetrics-gynecology resident long-acting reversible contraception training: the role of resident and program characteristics. American journal of obstetrics and gynecology, 222(4S), S923.e1–S923.e8. https://doi.org/10.1016/j.ajog.2019.12.007

[25] Castleberry, N. M., Stark, L., Schulkin, J., & Grossman, D. (2019). Implementing best practices for the provision of long-acting reversible contraception: a survey of obstetrician-gynecologists. Contraception, 100(2), 123–127. https://doi.org/10.1016/j.contraception.2019.03.053

[26] 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

[27] 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

[28] 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

[29] 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 are presented for eight programs: federal Medicaid efforts to support state use of the measures; four state Medicaid programs (i.e., the Iowa Medicaid Enterprise, the Washington State Health Care Authority, Louisiana Medicaid, 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); 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 #2904 at the following levels: Clinician group/practice, Facility, Health Plan, State, and Public Health Region. When data was available, we also examined trends over time, starting in 2016, the year that NQF#2904 was initially endorsed. We include descriptive statistics for each program and level of analysis below. For more details, see the attached Testing Attachment.

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

Federal Medicaid's use of NQF #2904 is demonstrated in two ways: first, as part of the Center for Medicaid and CHIP Services' (CMCS) Maternal and Infant Health Initiative from 2015 to 2018; and second, through the inclusion of the measure in the Adult and Child Core set.

Although CMCS' Maternal and Infant Health Initiative was implemented from 2015 to 2018, the overall measure scores were reported only for Federal Fiscal Year (FFY) 2016 and 2017.

FFY 2016 Median Measure Scores

Ages 15-20: 4.2

Ages 21-44: 4.8

FFY 2017 Median Measure Scores

Ages 15-20: 4.9

Ages 21-44: 5.7

Measure #2904 has been adopted into CMS' Adult and Child Core Set. However, 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 and FFY 2019, #2904 was reported in the Child Core Set for women ages 15-20. The measure score decreased slightly over this time period. Alabama and Nevada CHIP reported LARC rates less than 2% in FFY 2018. In FFY 2019, Alabama CHIP and Illinois Medicaid and CHIP reported LARC rates less than 2%, which may signal access barriers and should be investigated further.

FFY 2018 Median Measure Scores

Ages 15-20: 5.4

Range (minimum - maximum): 1.0 - 11.3

FFY 2019 Median Measure Scores

Ages 15-20: 4.8

Range (minimum - maximum): 1.1 - 20.0

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 5.50% of clients ages 15-44 were provided a LARC method. There was variation by public health region (n = 6) and clinician group/facility (n=3,081). For more details, see the Testing Attachment.

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

Number of included women ages 15-44: 116,892

Dates included: January 1 through December 31, 2018

Mean performance score: 3.69

Standard deviation: 10.45

Range (minimum – maximum): 0.00 – 100.00

Percentiles:

25th: 0.00

50th: 0.00

75th: 3.77

Scores by decile

0 – 10: 2778

11 – 20: 190

21 – 30: 41

31 – 40: 30

41 – 50: 22

51 – 60: 0

61 – 70: 0

71 – 80: 0

81 – 90: 0

91 – 100: 20

Number of entities with <2%: 2,175

Percentage of measured entities with <2%: 70.6%

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

Mean performance score: 5.55

Standard deviation: 0.67

Range: 4.64 – 6.57

Number of entities with <2%: 0

When analyzed over time, the percentage of women provided a LARC increased from 5.0% in 2017 to 6.3% in 2019. Among adolescents, the percentage increased from 5.2% to 7.0%. Among adults, the percentage increased from 5.0% to 6.0%.

Overall Measure Scores for IME (State)

2015

Ages 15-44: 5.0

Ages 15-20: 5.2

Ages 21-44: 5.0

2016

Ages 15-44: 6.0

Ages 15-20: 4.7

Ages 21-44: 4.9

2017

Ages 15-44: 6.0

Ages 15-20: 6.4

Ages 21-44: 5.9

2018

Ages 15-44: 5.5

Ages 15-20: 6.0

Ages 21-44: 5.3

2019

Ages 15-44: 6.3

Ages 15-20: 7.0

Ages 21-44: 6.0

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 3.97% of clients aged 15-44 years were provided a LARC method; the measure scores varied by health plan. For more details, see the Testing Attachment.

Number of measured entities: 5 Health Plans

Number of included women ages 15-44: 279,100

Dates included: January 1 through December 31, 2019

Mean performance score: 4.0

Range: 3.8 – 4.2

Number of entities with <2%: 0

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 6.5% of clients aged 15-44 years were provided a LARC method; the measure scores varied by health plan. For more details, see the Testing Attachment.

Number of measured entities: 5 Health Plans

Number of included women ages 15-44: 196,568

Dates included: January 1 through December 31, 2019

Mean performance score: 6.4

Range: 5.9 – 6.8

Number of entities with <2%: 0

5. Massachusetts Medicaid (MassHealth)

The 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 care organizations. About 6.0% of clients aged 15-44 years were provided a LARC method; the measure scores varied by health plan. For more details, see the Testing Attachment.

Number of measured entities: 21 Health Plans

Number of included women ages 15-44: 197,529

Dates included: January 1 through December 31, 2019

Mean performance score: 6.0

Range: 2.8 – 8.0

Number of entities with <2%: 0

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

The NYP ACN analysis included 31,084 female clients ages 15-44 who in calendar year 2018 received services from 31 NYP ACN facilities. Approximately 7.2% of clients ages 15-44 received a LARC method, and the measure scores varied across 31 facilities. For more details, see the Testing Attachment.

Number of measured entities: 31 facilities

Number of included women ages 15-44: 31,084

Dates included: January 1 through December 31, 2018

Mean performance score: 3.47

Standard deviation: 4.32

Range: 0.00 – 17.00

Percentiles:

25th: 0.6

50th: 1.8

75th: 5.65

Number of entities with <2%: 18

Percentage of measured entities with <2%: 58.1%

7. 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 ##% of clients ages 15-44 were provided a long-acting reversible contraceptive (LARC) method; variation existed across 56 facilities. For more details, see the Testing Attachment.

Number of measured entities: 56 facilities

Number of included women ages 15-44: 123,978

Dates included: January 1 through December 31, 2019

Mean performance score: 12.51

Standard deviation: 7.32

Range (minimum – maximum): 0.00 – 24.18

Percentiles:

25th: 9.00

50th: 13.00

75th: 17.00

Scores by decile

0 – 10: 20

11 – 20: 29

21 – 30: 7

31 – 40: 0

41 – 50: 0

51 – 60: 0

61 – 70: 0

71 – 80: 0

81 – 90: 0

91 – 100: 0

Number of entities with <2%: 12

Percentage of measured entities with <2%: 21.4%

8. 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, 15.9% of clients ages 15-19 and 17.2% of clients ages 20-44 were provided a LARC method; variation by grantee existed (e.g., from 0 to 43.2% for adolescent clients, and from 0 to 33.9% 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: 16.2

Standard deviation: 0.10

Range (minimum – maximum): 0.00 – 62.5

Percentiles:

25th: 9.2

50th: 14.7

75th: 22.7

Scores by decile

0 – 10: 25

11 – 20: 40

21 – 30: 21

31 – 40: 6

41 – 50: 1

51 – 60: 0

61 – 70: 1

71 – 80: 0

81 – 90: 0

91 – 100: 0

Number of patients: 563,474

Number of grantees with rates <2%: 4

Percent of grantees with rates <2%: 4.3

Ages 20-44

Mean performance score: 17.3

Standard deviation: 0.07

Range (minimum – maximum): 0.00 – 39.6

Percentiles:

25th: 12.7

50th: 18.1

75th: 22.2

Scores by decile

0 – 10: 16

11 – 20: 46

21 – 30: 31

31 – 40: 2

41 – 50: 0

51 – 60: 0

61 – 70: 0

71 – 80: 0

81 – 90: 0

91 – 100: 0

Number of patients: 2,642,038

Number of grantees with rates <2%: 2

Percent of grantees with rates <2%: 2.0

Number of measured entities: 100 grantees

FPAR 2019

Ages 15-19

Mean performance score: 15.9

Standard deviation: 0.09

Range (minimum – maximum): 0.00 – 43.2

Percentiles:

25th: 9.7

50th: 15.2

75th: 22.0

Scores by decile

0 – 10: 26

11 – 20: 44

21 – 30: 22

31 – 40: 6

41 – 50: 1

51 – 60: 0

61 – 70: 0

71 – 80: 0

81 – 90: 0

91 – 100: 0

Number of patients: 429,112

Number of grantees with rates <2%: 5

Percent of grantees with rates <2%: 5.1

Ages 20-44

Mean performance score: 17.2

Standard deviation: 0.08

Range (minimum – maximum): 0.00 – 33.9

Percentiles:

25th: 12.4

50th: 17.9

75th: 22.5

Scores by decile

0 – 10: 17

11 – 20: 43

21 – 30: 36

31 – 40: 3

41 – 50: 0

51 – 60: 0

61 – 70: 0

71 – 80: 0

81 – 90: 0

91 – 100: 0

Number of patients: 2,059,301

Number of grantees with rates <2%: 4

Percent of grantees with rates <2%: 4.0

From 2016 through 2019, the percentage of all Title X family planning users provided a LARC method of contraception remained stable during this period [1-4]:

2016: 14%

2017: 16%

2018: 16%

2019: 16%

References

[1] 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

[2] 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

[3] 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

[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

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.

Not applicable.

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.*

As measure steward, OPA states on our website that NQF #2904 "should be used as an access measure; very low rates (less than 1-2%) may signal barriers to LARC provision that should be addressed through training … [and] and quality improvement processes" [1]. To evaluate access to LARC methods by race/ethnicity, we calculated NQF #2904 measure scores by these sub-populations in two datasets: 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 indicated that no race/ethnicity group had measure scores less than 2%, suggesting that PPFA clients may have access to LARC methods:

African American: 9.51

Alaskan Native: 26.15

Asian: 14.83

Hispanic: 16.90

Multi-racial: 17.37

Native American: 16.78

Pacific Islander: 15.27

White: 16.84

Other race: 15.14

For 2014-2018, WA HCA reported NQF #2904 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). Among adults ages 21-44, all race/ethnicity groups had LARC provision rates greater than 2% during this five-year period. Except for 2014, LARC provision rates in clients ages 15-20 were also more than 2% for all race/ethnicity groups.

In 2018, the NQF #2904 measure scores for ages 15-20 were as follows (note that race/ethnicity categories other than "Hispanic" report ethnicity as "Not Hispanic" or "Unknown"):

Hispanic: 5.2

White: 7.0

Asian: 3.1

Black: 4.5

American Indian/Alaska Native: 7.3

Hawaiian/Pacific Islander: 4.2

More than One Race: 6.9

Other/Unknown: 4.0

The 2018 LARC provision rates for ages 21-44 by race/ethnicity reported were:

Hispanic: 8.8

White: 5.9

Asian: 5.8

Black: 6.4

American Indian/Alaska Native: 5.7

Hawaiian/Pacific Islander: 5.3

More than One Race: 6.5

Other/Unknown: 5.9

In these two health systems, all race/ethnicity groups appear to have access to LARC methods. It is important to note that OPA emphasizes that the measure should be used only to monitor access to LARC; and that it could be harmful to set a high benchmark for this measure, because doing so may incentivize coercive practices [2-3]. Contraceptive services must be offered in a client-centered manner, as recommended by CDC and OPA [4].

References

[1] Office of Population Affairs. (n.d.). Long-Acting Reversible Contraceptive Methods. 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/contraceptive-care-measures/long-acting-reversible.

[2] 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

[3] Gold, R.B. (2014). Guarding Against Coercion While Ensuring Access: A Delicate Balance. Guttmacher Policy Review, 17(3), 8-14.

[4] 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.

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 use of long-acting reversible contraception (LARC) methods, a special analysis of data from the National Survey of Family Growth (NSFG) 2015-2017 was conducted. The

current analysis suggests that no significant differences exist by age group, race/ethnicity, marital status, and poverty level. For more details, 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):*

Access to Care, 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/long-acting-reversible

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_2904_Codes_2021.xlsx

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

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

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

Not an instrument-based measure

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

No

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

There have not been any important changes to the measure specifications since the last measure update.

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 were provided a long-acting reversible method of contraception (LARC), i.e., intrauterine device or implant.

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 were provided a long-acting reversible method of contraception (LARC). To identify the numerator, follow these steps:

Step 1 Define the numerator by identifying women who used a long-acting reversible method of contraception (LARC) in the measurement year. To do this, use the codes in Table CCW-F.

Step 2 Calculate the rates by dividing the number of women who used a LARC 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 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) women who had a live birth in the last 2 months of the measurement year; or (3) women 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_2904_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 obtained 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 2020. 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 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 Postnatal 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; or
- 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 long-acting reversible contraceptive (LARC) methods are safe and recommended for teen and nulliparous populations who wish to use them, the American Academy of Pediatrics (AAP), ACOG, 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. Thus, it is important to monitor NQF #2904 measure scores for adolescents and adults to identify reporting units with very low LARC provision (less than 2%). We utilize age groups that are consistent with Center for Medicaid and CHIP Services (CMCS) reporting requirements; adolescents are defined as 15-20 years of age 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 = Score within a defined interval

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 aged 15-44 years of age 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 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 a long-acting reversible method of contraception (LARC), i.e., IUD or implant.

Step 4 Calculate the rates by dividing the number who were provided or continued use of a long-acting reversible method of contraception (LARC) by the number of women in the denominator. Calculate the rates for all women aged 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, paid, suspending, pending, and denied claims with diagnosis codes (ICD-10-CM) and procedures codes (HCPCS, CPT, and ICD-10-PCS), as well as NDC codes.

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.)*

Validity – See attached Measure Testing Submission Form

LARC_2904_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*)**:** 2904
Measure Title: Contraceptive Care – Access to LARC
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 abstract 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 60 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 facilities within affiliates. Affiliates cover geographic service areas that range from several counties within a state, a state population, and 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 2018 IME dataset included all female Medicaid clients aged 15-44 years who resided in 99 counties and 6 public health regions and participated in either the general Medicaid program or the state-sponsored Family Planning Program (FPP). To be eligible for FPP services, the following guidelines apply: 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 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 (MCO), 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 years 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. Female clients aged 15-44 years who received services in 2018 from 8 NYP outpatient locations comprise the NYP dataset. 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

Datasets 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 the 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 the affiliate be considered a reasonable proxy for a U.S. state, for purposes of this application.

(2) **The Iowa Medicaid enterprise** (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. 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). Due to the small percentage of clients in Iowa who were enrolled in MCOs, we did not conduct reliability testing at this level in Iowa.

(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

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 Grantee 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
Health center	363	PPFA
Benefit type	2	IME
Public health region	6	IME
Health plan (Medicaid managed care/HMO)	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
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)	*	*	*

Level of analysis	Number of patients: 15 – 20 years	Number of patients: 21-44 years	Number of patients: 15-44 years
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)	*	*	*

Level of analysis	Number of patients: 15 - 20 years	Number of patients: 21 -44 years	Number of patients: 15 - 44 years
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
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	269	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 analyzed 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 initiative and public health interventions. We utilized the age group categories developed by the Center for Medicaid and CHIP Services (CMCS). CMCS define 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 are identified 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 LARC 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 managed care 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)	15-44	2,915	**.978**	2,929	**.995**
*	21-44	2,180	.974	2,201	.992
*	15-20	604	.974	644	.986
Facility (NYP)	15-44	597	**.953**	597	**.953**
*	21-44	560.5	.811	629	.953
*	15-20	87.5	.743	145	.885
Public health region (IME)	15-44	14,955	**.901**	14,955	**.901**
*	21-44	11,346	.727	11,346	.727

Level	Age group	Results: Median N (all units)	Results: Reliability (all units)	Results: Median N (unit size ≥ 75)	Results: Reliability (unit size ≥ 75)
*	15-20	3609	.886	3609	.886
Group billing provider (IME)	15-44	5	**.178**	148.5	**.749**
*	21-44	3	.159	148	.733
*	15-20	1	.149	129	.788
Health plan (WA HCA)	15-44	26,460	**.811**	26,460	**.811**
*	21-44	17,867.5	.843	17,867.5	.843
*	15-20	8,707.5	.527	8,707.5	.527
Health plan (MA)	15-44	7,362.5	**.919**	7,362.5	**.919**
*	21-44	5,320	.825	5,320	.825
*	15-20	1,683	.832	1,683	.832
Health plan (LA Medicaid)	15-44	65,818.5	**.708**	65,818.5	**.708**
*	21-44	48,151	.497	48,151	.497
*	15-20	17,667.5	.532	17,667.5	.532

*cell intentionally left blank

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)	15-20	4,839	.0673	.9971
*	21-44	11,648	.0675	.9988
*	15-44	16,590	.0617	.9991
Health centers (estimated within each affiliate) (PPFA)	15-20	366	.0649 (median)	.6688 - .9949
*	21-44	1,016	.0401 (median)	.7775 - .9994

Level	Age group	Results: Median N	Results: ICC	Results: Reliability
*	15-44	1,379	.0488 (median)	.8329 - .9994
Public health region (IME)	15-20	1,438	.0055	.8887
*	21-44	3,767	.0017	.8666
*	15-44	5,205	.0022	.9197
Benefit type (IME)	15-20	5,850	.0682	.9977
*	21-44	16,526	.0537	.9989
*	15-44	22,375	.0585	.9993
Region by benefit type (IME)	15-20	716	.0716	.9822
*	21-44	2,325	.0512	.9921
*	15-44	2,954	.0574	.9945
Health plan (WMP)	15-20	1,231	.0043	.8414
*	21-44	3,936	.0082	.9702
*	15-44	5,167	.0067	.9721

*cell intentionally left blank

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 22 of 25 affiliate groupings of health centers, for benefit type, for region by benefit type, and 2 of 3 age groups at the health plan level; ICCs at these levels were moderately high, ranging from 4-8%. Of note, reliability did decline <.90 at three levels, i.e., for three of the 25 affiliate groupings of health centers, among public health region and for the age group of 15-20 for health plan. However, two of the three affiliate groupings with lower reliability had only a single health center and thus no reliability estimation was possible. The ICC for public health region was also below .01; yet due to the larger number of cases for region, reliabilities remained above .70.

It is commonly advised that reliability should be ≥ .90 for making decisions, and ≥ .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 ≥ .90 for making high-stakes decisions, and ≥ .70 for general reporting/monitoring (Eijkenaar, 2013; Adams, 2010). Our tested reliability is greater than .70 at the facility and health plan levels and consistently greater than .90 at the public health region level for the 15-44 age group, showing adequate to high reliability. This was mostly driven by the large number of patients per unit at these levels. At the health plan level, the reliability was above .70 among 15-20 and 21-44 age groups for MA but below .70 for WA HCA and LA Medicaid among these stratified age groups, likely due to the combination of low LARC provision rates and the limited variation of provision rates across health plans in these age groups in WA and LA. With only 5 health plans having very similar rates, the ability to distinguish among health plans by measure performance is limited.

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 .70 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. At the affiliate level, at some health centers, for benefit type, for region by benefit type, and for 2 of 3 age groups at the health plan level, we found reliabilities well above the commonly accepted .90 reliability threshold for reporting and decision-making. Of those that were below .90, only a few three were below .70, and two of those were due to having only one health center inside the affiliate.

High reliability was largely driven by two factors. First, the data exhibited adequate variation in the rates of LARC 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 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. This hypothesis is based on the assumption that the provision of LARC and most of these well women-related measures requires direct provider interaction in a medical facility, and women visiting a facility for one service will likely be offered other related services. Therefore, these related measures should be positively correlated to the contraceptive care measure. We also hypothesize that the correlation may be weaker for chlamydia screening due to the difference between its target population and that of 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 Grantee 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 5 data elements in total, including 2 LARC methods (Implantable and IUD) 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 Vanessa Dalton, MD, MPH, Associate Professor, Director, Program on Women's Health Care Effectiveness Research, University of Michigan
3. Anne Dunlop, MD, MPH, Program Director, Preventive Medicine Division, Emory University School of Medicine
4. Daryn Eikner, MS, Vice President of Health Care Delivery, National Family Planning & Reproductive Health Association
5. Jan Engstrom, PhD, RN, CNM, WHNP-BC, Professor & Acting Chairperson, Department of Women, Children and Family Nursing, College of Nursing, Armour Academic Center
6. Mark Hathaway, MD, MPH, Senior Technical Advisor, Jhpiego – Johns Hopkins University
7. Michael Policar, MD, MPH, Clinical Professor of Obstetrics, Gynecology, and Reproductive Sciences, UCSF School of Medicine
8. 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 listed 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 Puerperium (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 years). 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

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	15-44	3,075	**.26***	**.80***	**(.70, .87)**
- -	21-44	2,266	.27*	.80*	(.68, .86)
- -	15-20	623	.23	.78*	(.66, .85)
Gynecological Examination	15-44	3,075	**.78***	**.98***	**(.97, .99)**
- -	21-44	2,266	.74*	.98*	(.97, .99)
- -	15-20	623	.41*	.93*	(.83, .97)
Cervical Cancer Screening†	21-44	2,266	**.66***	**.83***	**(.73, .89)**
Chlamydia Screening†	16-24	1,233	**.63***	**.88***	**(.80, .92)**
- -	21-24	657	.60*	.85*	(.76, .90)
- -	16-20	503	.63*	.89*	(.81, .93)

s*statistically significant at $p < 0.05$

†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

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	15-44	150.5	**.26***	**.39***	**(.26, .51)**
- -	21-44	157	.23*	.37*	(.20, .51)
- -	15-20	130	.66*	.67*	(.42, .82)
Gynecological Examination	15-44	150.5	**.18***	**.20***	**(.04, .34)**
- -	21-44	157	.08	.11	(-.07, .28)

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)
- -	15-20	130	.26*	.53*	(.16, .76)
Cervical Cancer Screening†	21-44	156	**.24***	**.29***	**(.12, .45)**
Chlamydia Screening†	16-24	150	**.22**	**.18**	**(-.15, .46)**
- -	15-44	150.5	**.26***	**.39***	**(.26, .51)**
- -	21-44	157	.23*	.37*	(.20, .51)

*statistically significant at $p < 0.05$

†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

Critical data elements

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

Data element validity test results, Iowa title X clinics, 2019

Data elements	Age group	Sensitivity	Specificity	PPV	NPV	% agreement	Kappa	95% CI
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
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**

Data elements	Age group	Sensitivity	Specificity	PPV	NPV	% agreement	Kappa	95% CI
*	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

*cell intentionally left blank

2016 Submission

The mean rating for this measure was 4.33 with a median of 4.5 (between Agree and Strongly Agree), range 3-5. There were 44.4% (n = 4) of respondents who strongly agreed, 44.4% (n = 4) of respondents who agreed, and 11.1% (n = 1) of respondents who neither agreed nor disagreed that the scores obtained from this measures, 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 measures 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." Finally, one respondent said this measure "provides a good metric for access, not necessarily quality."

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 statistically significant moderate to strong positive correlations between the contraceptive care measure and all four related measures at the facility level among all age groups. In comparison, Pearson's correlation test showed weak to moderate positive correlations with these measures. At the group billing provider level, multilevel correlation estimation showed weak to moderate positive associations between the contraceptive care measure and all the related measures except chlamydia screening among the 15-44 age group women (21-44 for cervical cancer screening). Similarly, Pearson's correlation test demonstrated weak positive correlations with the same related measures. We didn't find any association between the contraceptive care measure and chlamydia screening.

While both methods showed statistically significant correlations, the magnitude of correlation was 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 much better suited for binary outcomes and smaller units of analysis.

Overall, we observed statistically significant positive correlations between the contraceptive care measure and those services that (in theory) should be closely related (contraceptive counseling, gynecological examination, and cervical cancer screening); these were highly consistent with our hypotheses, provide good evidence for validity of the contraceptive care measure at the score level. We also observed no association for chlamydia screening at the group billing provider level. This is expected and we speculate that the absence of association may be due to the 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 that limit the variation of the measure and the comparability of the target populations. It's 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, whereas specificity, PPV, and NPV were above 0.9 for all data elements. Percent agreement was consistently over 95%. We also observed statistically significant Kappa above 0.6 for all data elements, 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 2904 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 Submission

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

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	*	*	*

*cell intentionally left blank

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	3,793	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 more measured entities would have very low percentages of LARC provision, making it challenging to identify units with truly low (i.e., less than 2%) measure scores, which may signify barriers to access. 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 2b5.

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 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 use of long-acting reversible contraception (LARC) methods, a special analysis of data from the National Survey of Family Growth (NSFG) 2015-2017 was conducted. The current analysis suggests that no significant differences exist by age group, race/ethnicity, marital status, and poverty level.

Percentage of women 15-44 years of age at risk of unintended pregnancy* that used a long-acting reversible method of contraception (LARC), National Survey of Family Growth, 2015-2017

Measures	Frequency	Weighted Frequency	Percent	95% Confidence Limits
Age	- -	- -	- -	- -
15-19	54	738,481	19.41	11.25 – 27.58
20-29	223	2,794,796	17.00	13.48 – 20.53
30-44	249	3,397,855	12.44	10.39 – 14.49
Race/ethnicity	- -	- -	- -	- -
Hispanic	122	1,684,627	16.82	13.00 – 20.64
Non-Hispanic White	262	4,044,862	15.45	12.86 – 18.04
Non-Hispanic Black	100	813,875	12.73	9.84 – 15.62

Measures	Frequency	Weighted Frequency	Percent	95% Confidence Limits
Marital status	- -	- -	- -	- -
Married	182	2,873,639	14.45	11.86 – 17.05
Cohabitating	93	1,372,168	18.88	13.90 – 24.17
Widowed/divorced/separated	55	577,608	12.54	7.98 – 17.11
Never married	196	2,107,718	13.34	10.31 – 16.37
Federal poverty level	- -	- -	- -	- -
<100	149	1,650,226	15.22	11.92 – 18.52
100-199	116	1,460,401	13.05	9.38 – 16.73
200-399	139	1,923,664	14.84	12.06 – 17.62
400-499	36	471,308	12.2	7.15 – 17.25
500+	86	1,425,534	16.38	11.10 – 21.67

* 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

- - cell intentionally left blank

2016 Submission

We do not believe that risk adjustment is justified. Although there are [possible] 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 long-acting reversible methods of contraception (see table below). This analysis suggests that there are significant differences by age (for adolescents compared to adult women) and for women who were never married (compared to women of other marital status). 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 long-acting reversible method of contraception (LARC), National Survey of Family Growth, 2011-2013

Measures	Frequency	Weighted Frequency	Percent	95% Confidence Limits for Row Percent
Age	- -	- -	- -	- -
15-19	15	128,000	3.21	0.67 - 5.75
20-29	243	2,038,000	12.36	10.03 - 14.69
30-44	193	2,340,000	9.06	7.12 - 11.01
Race/ethnicity	- -	- -	- -	- -
Hispanic	140	1,060,000	11.62	8.31 - 14.93
NH White, Single race	204	2,699,000	10.05	8.01 - 12.08
NH Black, Single race	80	414,000	6.40	4.63 - 8.18
Marital status	- -	- -	- -	- -
Married	177	2,331,000	12.13	9.52 - 14.73
Cohab	92	851,000	11.86	8.81 - 14.91
Wid/div/sep	56	529,000	10.35	4.98 - 15.72
Never married	126	796,000	5.38	4.03 - 6.73
Federal poverty level	- -	- -	- -	- -
<100	166	1,310,000	10.29	8.01 - 12.56
100-199	107	1,035,000	10.42	7.43 - 13.42
200-399	112	1,265,000	9.37	6.75 - 11.99
400-499	22	293,000	7.91	4.23 - 11.60
500+	44	604,000	9.37	5.86 - 12.88

* 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.

- - cell intentionally left blank

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 contraceptive counseling recommendations specifically for this population (Gavin 2014, ACOG 2017

Committee Opinion 710, Menon 2020). Historically, LARCs have been more difficult to access for all age groups because they require clinicians to have specialized training in implant and IUD placement and removal, but they have been particularly difficult for teens to access due to outdated clinical guidance around which populations are eligible for LARCs (Kumar 2016, Pritt 2017, Smith 2017). Though current guidance notes that LARCs are safe and recommended for teen and nulliparous populations, it can still be difficult for these populations to access these highly effective contraceptive methods (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 LARC access 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 LARC among women who wish to prevent pregnancy. The literature is summarized in section 2b3.3a above.

2016 Submission

Not applicable.

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 clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

2021 Submission

We excluded 2 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 LARC methods rate (%) in 54 PPFA facilities, 2019

Mean	SD	Minimum	10th percentile	25th percentile	Median	75th percentile	90th percentile	Maximum
12.3	7.1	0.0	0.1	9.3	12.9	17.7	20.9	24.2

Using the approach described in 2b4.1, 25 facilities (46.3%) of 54 PPFA facilities were rated as higher than the mean (i.e. the lower limit of facility's 95% confidence interval was > 12.3) and 20 facilities (37.0%) were identified as lower than the mean (i.e. the upper limit of facility's 95% confidence interval was < 12.3). Another 9 facilities were either higher or lower than the mean (12.3) 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 LARC methods

Level	Age group	Rate: Mean (range)
Facility (PPFA), n=56	15-44	.135 (.000-.282)
*	21-44	.127 (.000-.268)
*	15-20	.164 (.000-.406)
Facility (NYP), n=31	15-44	.072 (.000-.170)
*	21-44	.068 (.000-.167)
*	15-20	.089 (.000-.190)
Public Health Region (IME), n=6	15-44	.055 (.046-.066)
*	21-44	.053 (.047-.062)
*	15-20	.060 (.044-.077)
Group Billing Provider (IME), n=3081	15-44	.060 (.000-1.00)
*	21-44	.059 (.000-1.00)

Level	Age group	Rate: Mean (range)
*	15-20	.066 (.000-1.00)
Health Plan (WA HCA), n=5	15-44	.065 (.059-.068)
*	21-44	.067 (.058-.069)
*	15-20	.062 (.055-.065)
Health Plan (MA), n=21	15-44	.060 (.028-.080)
*	21-44	.070 (.054-.091)
*	15-20	.032 (.010-.063)
Health Plan (LA Medicaid), n=5	15-44	.040 (.038-.042)
*	21-44	.039 (.038-.041)
*	15-20	.041 (.038-.048)

*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 1-4, which are appended at the end of this document.

Level	Age group	Rate (Provision of LARC methods)
Affiliate (PPFA), n=25 Mean (range)	15-20	.099 (.029-.210)
*	21-44	.119 (.028-.199)
*	15-44	.114 (.035-.202)
Health center (PPFA), n=363 Mean (range)	15-20	.102 (.000-.388)
*	21-44	.114 (.000-.312)
*	15-44	.110 (.000-.347)
Public health region (IME) Mean (range)	15-20	.085 (.074-.104)
*	21-44	.096 (.087-.113)
*	15-44	.093 (.087-.111)
Benefit type (IME) Mean (range)	15-20	.085 (.047-.116)
*	21-44	.096 (.051-.114)
*	15-44	.093 (.050-.114)
PH Region by benefit type (IME) Mean (range)	15-20	.085 (.034-.139)
*	21-44	.096 (.048-.129)
*	15-44	.093 (.048-.131)
Health plan/HMO (WMP) Mean (range)	15-20	.057 (.048-.075)
*	21-44	.077 (.058-.122)
*	15-44	.072 (.061-.109)

*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 7% (range: 2% - 12%) lower or 6% (range: 2% - 12%) 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.

The primary intent of the measure is to evaluate access to LARC methods, and very low rates (less than 1-2%) may signal existing barriers to LARC provision. These barriers include client and physician lack of knowledge, financial constraints, and logistical issues, which all are well-documented but can be addressed. In our analysis, measure rates vary widely across all reporting units, but it is more important to focus on units with scores less than 2%. For example, about 19% of PPFA facilities reported less than 2% LARC provision; seven facilities provided zero LARC methods. LARC use among NYP facilities was lower than in the PPFA facilities. Within the IME program, LARC provision percentages at the group billing provider and public health region levels also varied.

These differences suggest that identifying meaningful differences in performance across measured entities is possible. The PPFA and NYP facilities with LARC use below 2% could be evaluated to detect and address preventable barriers to LARC provision. In Iowa, practices/groups and regions that show LARC use well below the median should be assessed to determine if clients' access could be improved. Barriers to LARC provision within health systems might be removed through training and quality improvement processes. However, access to LARC does not mean that its use is aligned with patient goals.

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 using this contraceptive provision measure and the PCCC in tandem.

2016 Submission

There are very large and meaningful differences in rates across all reporting units. For example, the provision of LARC across affiliates ranged from approximately 3% to 20% within the 15-44 year age group. Among health centers, the range was 0% to almost 40% within the 15-44 year age group; four health centers had 0% LARC use and 24 had LARC use that was less than 2%. Within the IME program, the differences between LARC provision in the general Medicaid program was substantially lower than in the waiver program, i.e., from approximately 5% to 11.5%. The ranges across regions were more narrow but still notable, i.e., from approximately 5-8% on the lower end and up to 13% on the higher end of the range. In Wisconsin, the rates across health plans ranged from 4.8% to 12.2%.

These differences suggest that it will be possible to use these measures to identify meaningful differences in performance across measured entities. For example, the PPFA health centers with LARC use that is below 2% could be assessed to identify avoidable barriers to LARC access, and steps could be taken to remove those barriers. In Iowa, it may be useful to explore why LARC provision is so much lower in the general Medicaid program than in the family planning waiver program; and regions that are well below the median should be similarly assessed to see if steps can be taken to improve clients' access to LARC. Similarly, in Wisconsin, health plans with LARC provision rates that are below the mean could be assessed to determine if there are barriers that could be removed.

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

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

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

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

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. Thus, it is difficult to determine when claims data is missing.

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.

Smith, A., Harney, K. F., Singh, T., & Hurwitz, A. G. (2017). Provider and Health System Factors Associated with Usage of Long-Acting Reversible Contraception in Adolescents. *Journal of pediatric and adolescent gynecology*, *30*(6), 609–614. https://doi.org/10.1016/j.jpag.2017.05.001

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 the provision of LARC methods by facility, NewYork Presbyterian Hospital system, 2018.

Facility ID	LARCmeasure: 15 to <21 Years (LARC Provision)	LARCmeasure: 15 to <21 Years (Total N)	LARCmeasure: 15 to <21 Years (Rate)	LARCmeasure: 15 to <21 Years (Reliability (all units))	LARCmeasure: 15 to <21 Years (Reliability (unit size≥75))	LARCmeasure: 21 to 44 years (LARC Provision)	LARCmeasure: 21 to 44 years (Total N)	LARCmeasure: 21 to 44 years (Rate)	LARCmeasure: 21 to 44 years (Reliability (all units))	LARCmeasure: 21 to 44 years (Reliability (unit size≥75))	LARCmeasure: all age groups (LARC Provision)	LARCmeasure: all age groups (Total N)	LARCmeasure: all age groups (Rate)	LARCmeasure: all age groups (Reliability (all units))	LARCmeasure: all age groups (Reliability (unit size≥75))
101	0	2	0.000	0.093	*	0	90	0.000	0.795	0.848	0	92	0.000	0.848	0.848
102	0	12	0.000	0.380	*	0	436	0.000	0.949	0.965	0	448	0.000	0.965	0.965
103	0	38	0.000	0.660	*	0	1083	0.000	0.979	0.986	0	1121	0.000	0.986	0.986
201	26	614	0.042	0.969	0.969	0	17	0.000	0.422	0.975	26	631	0.041	0.975	0.975
202	2	88	0.023	0.818	0.819	27	1032	0.026	0.978	0.985	29	1120	0.026	0.985	0.985
203	1	114	0.009	0.854	0.854	13	1184	0.011	0.981	0.987	14	1298	0.011	0.987	0.987
204	0	35	0.000	0.642	*	4	522	0.008	0.957	0.971	4	557	0.007	0.971	0.971
205	298	1568	0.190	0.988	0.988	1477	8894	0.166	0.997	0.998	1775	10462	0.170	0.998	0.998
301	0	7	0.000	0.264	*	0	489	0.000	0.955	0.968	0	496	0.000	0.968	0.968
302	0	271	0.000	0.933	0.933	0	2	0.000	0.079	0.943	0	273	0.000	0.943	0.943
303	6	228	0.026	0.921	0.921	41	2161	0.019	0.989	0.993	47	2389	0.020	0.993	0.993
304	5	165	0.030	0.894	0.895	48	1709	0.028	0.987	0.991	53	1874	0.028	0.991	0.991
401	0	14	0.000	0.417	*	0	127	0.000	0.845	0.895	0	141	0.000	0.895	0.895
402	3	129	0.023	0.868	0.869	24	1289	0.019	0.982	0.989	27	1418	0.019	0.989	0.989
403	0	19	0.000	0.493	*	5	281	0.018	0.924	0.948	5	300	0.017	0.948	0.948
404	0	42	0.000	0.682	*	1	407	0.002	0.946	0.965	1	449	0.002	0.965	0.965
405	0	53	0.000	0.731	*	2	460	0.004	0.952	0.969	2	513	0.004	0.969	0.969
501	0	23	0.000	0.541	*	3	540	0.006	0.959	0.972	3	563	0.005	0.972	0.972
502	1	166	0.006	0.895	0.895	0	4	0.000	0.147	0.912	1	170	0.006	0.912	0.912
503	1	83	0.012	0.809	0.810	4	684	0.006	0.967	0.979	5	767	0.007	0.979	0.979
504	2	48	0.042	0.711	*	21	677	0.074	0.967	0.978	23	725	0.069	0.978	0.978
601	0	67	0.000	0.774	*	1	902	0.001	0.975	0.983	1	969	0.001	0.983	0.983
602	1	109	0.009	0.848	0.849	5	965	0.005	0.976	0.985	6	1074	0.006	0.985	0.985
603	1	54	0.019	0.734	*	21	836	0.025	0.973	0.982	22	890	0.025	0.982	0.982
701	2	81	0.025	0.806	0.807	33	581	0.057	0.961	0.976	35	662	0.053	0.976	0.976
801	97	710	0.137	0.973	0.973	1	6	0.167	0.205	0.977	98	716	0.137	0.977	0.977
802	31	518	0.060	0.964	0.964	0	0	*	*	0.969	31	518	0.060	0.969	0.969
803	10	93	0.108	0.826	0.827	0	0	*	*	0.849	10	93	0.108	0.849	0.849
804	16	161	0.099	0.892	0.892	0	1	0.000	0.041	0.908	16	162	0.099	0.908	0.908
805	1	87	0.011	0.817	0.817	0	0	*	*	0.841	1	87	0.011	0.841	0.841
806	4	106	0.038	0.844	0.845	0	0	*	*	0.865	4	106	0.038	0.865	0.865
Total or Mean	**508**	**5705**	**0.089**	*	*	**1731**	**25379**	**0.068**	*	*	**2239**	**31084**	**0.072**	*	*

Facility ID	LARC measure: 15 to <21 Years (LARC Provision)	LARC measure: 15 to <21 Years (Total N)	LARC measure: 15 to <21 Years (Rate)	LARC measure: 15 to <21 Years (Reliability (all units))	LARC measure: 15 to <21 Years (Reliability (unit size≥75))	LARC measure: 21 to 44 years (LARC Provision)	LARC measure: 21 to 44 years (Total N)	LARC measure: 21 to 44 years (Rate)	LARC measure: 21 to 44 years (Reliability (all units))	LARC measure: 21 to 44 years (Reliability (unit size≥75))	LARC measure: all age groups (LARC Provision)	LARC measure: all age groups (Total N)	LARC measure: all age groups (Rate)	LARC measure: all age groups (Reliability (all units))	LARC measure: all age groups (Reliability (unit size≥75))
*	*	*	*	**Overall Reliability**	**Overall Reliability**	*	*	*	**Overall Reliability**	**Overall Reliability**	*	*	*	**Overall Reliability**	**Overall Reliability**
*	Median n	87.5	*	**0.743**	**0.885**	Median n	560.5	*	**0.811**	**0.953**	Median n	597	*	**0.953**	**0.953**
*	Min n	2	*	*	*	Min n	0	*	*	*	Min n	87	*	*	*

*cell intentionally left blank

Table 2. Rates and reliabilities for the provision of LARC methods by facility, 56 PPFA Facilities, 2019.

Facility ID	LARCmeasure: 15 to <21 Years LARC Provision	LARCmeasure: 15 to <21 Years Total N	LARCmeasure: 15 to <21 Years Rate	LARCmeasure: 15 to <21 Years Reliability (all units)	LARCmeasure: 15 to <21 Years Reliability (unit size≥75)	LARCmeasure: 21 to 44 years LARC Provision	LARCmeasure: 21 to 44 Total N	LARCmeasure: 21 to 44 Rate	LARCmeasure: 21 to 44 Reliability (all units)	LARCmeasure: 21 to 44 Reliability (unit size≥75)	LARCmeasure: all age groups LARC Provision	LARCmeasure: all age groups Total N	LARCmeasure: all age groups Rate	LARCmeasure: all age groups Reliability (all units)	LARCmeasure: all age groups Reliability (unit size≥75)
1	65	422	0.154	0.986	0.985	83	1083	0.077	0.993	0.993	148	1505	0.098	0.995	0.995
2	227	737	0.308	0.992	0.991	1046	4999	0.209	0.998	0.998	1273	5736	0.222	0.999	0.999
3	134	747	0.179	0.992	0.991	379	2605	0.145	0.997	0.997	513	3352	0.153	0.998	0.998
4	222	884	0.251	0.993	0.993	650	3529	0.184	0.998	0.998	872	4413	0.198	0.998	0.998
5	235	1213	0.194	0.995	0.995	604	4002	0.151	0.998	0.998	839	5215	0.161	0.999	0.999
6	127	868	0.146	0.993	0.993	318	2923	0.109	0.997	0.997	445	3791	0.117	0.998	0.998
7	130	759	0.171	0.992	0.992	400	2951	0.136	0.997	0.997	530	3710	0.143	0.998	0.998
8	106	587	0.181	0.990	0.989	223	1612	0.138	0.995	0.995	329	2199	0.150	0.997	0.997
9	252	1400	0.180	0.996	0.995	628	4253	0.148	0.998	0.998	880	5653	0.156	0.999	0.999
10	219	812	0.270	0.993	0.992	523	2541	0.206	0.997	0.997	742	3353	0.221	0.998	0.998
11	30	225	0.133	0.974	0.972	43	668	0.064	0.988	0.988	73	893	0.082	0.992	0.992
12	159	721	0.221	0.992	0.991	380	2222	0.171	0.996	0.996	539	2943	0.183	0.998	0.998
13	1	803	0.001	0.992	0.992	4	3489	0.001	0.998	0.998	5	4292	0.001	0.998	0.998
14	178	886	0.201	0.993	0.993	488	3136	0.156	0.997	0.997	666	4022	0.166	0.998	0.998
15	231	781	0.296	0.992	0.992	687	3016	0.228	0.997	0.997	918	3797	0.242	0.998	0.998
16	36	167	0.216	0.965	0.963	66	438	0.151	0.982	0.982	102	605	0.169	0.989	0.988
17	59	220	0.268	0.973	0.971	137	719	0.191	0.989	0.989	196	939	0.209	0.993	0.993
18	131	769	0.170	0.992	0.992	275	2514	0.109	0.997	0.997	406	3283	0.124	0.998	0.998
19	224	979	0.229	0.994	0.993	387	2481	0.156	0.997	0.997	611	3460	0.177	0.998	0.998
20	76	306	0.248	0.980	0.979	138	724	0.191	0.989	0.989	214	1030	0.208	0.993	0.993
21	39	99	0.394	0.942	0.939	93	490	0.190	0.984	0.984	132	589	0.224	0.988	0.988
22	144	604	0.238	0.990	0.989	749	4050	0.185	0.998	0.998	893	4654	0.192	0.999	0.998
23	63	373	0.169	0.984	0.983	98	926	0.106	0.991	0.991	161	1299	0.124	0.995	0.995
24	73	236	0.309	0.975	0.973	318	1401	0.227	0.994	0.994	391	1637	0.239	0.996	0.996
25	175	833	0.210	0.993	0.992	363	2136	0.170	0.996	0.996	538	2969	0.181	0.998	0.998
26	0	612	0.000	0.990	0.990	4	2512	0.002	0.997	0.997	4	3124	0.001	0.998	0.998
27	103	510	0.202	0.988	0.987	352	2594	0.136	0.997	0.997	455	3104	0.147	0.998	0.998
28	151	772	0.196	0.992	0.992	436	3128	0.139	0.997	0.997	587	3900	0.151	0.998	0.998
29	56	268	0.209	0.978	0.976	126	989	0.127	0.992	0.992	182	1257	0.145	0.995	0.994
30	0	285	0.000	0.979	0.978	0	1053	0.000	0.992	0.992	0	1338	0.000	0.995	0.995
31	0	202	0.000	0.971	0.969	1	573	0.002	0.986	0.986	1	775	0.001	0.991	0.991
32	73	340	0.215	0.982	0.981	97	615	0.158	0.987	0.987	170	955	0.178	0.993	0.993
33	111	730	0.152	0.992	0.991	170	1581	0.108	0.995	0.995	281	2311	0.122	0.997	0.997
34	50	358	0.140	0.983	0.982	92	705	0.130	0.989	0.989	142	1063	0.134	0.994	0.993
35	44	478	0.092	0.987	0.987	101	1133	0.089	0.993	0.993	145	1611	0.090	0.996	0.996
36	0	11	0.000	0.643	*	0	48	0.000	0.857	*	0	59	0.000	0.895	*

37	92	774	0.119	0.992	0.992	146	1493	0.098	0.995	0.995	238	2267	0.105	0.997	0.997
38	31	340	0.091	0.982	0.981	88	918	0.096	0.991	0.991	119	1258	0.095	0.995	0.994
39	91	676	0.135	0.991	0.991	274	2739	0.100	0.997	0.997	365	3415	0.107	0.998	0.998
40	67	518	0.129	0.988	0.988	315	2397	0.131	0.997	0.997	382	2915	0.131	0.998	0.998

*cell intentionally left blank

Table 2. Rates and reliabilities for the provision of LARC methods by facility, 56 PPFA Facilities, 2019 (cont.)

Facility ID	LARCmeasure: 15 to <21 Years (LARC Provision)	LARCmeasure: 15 to <21 Years (Total N)	LARCmeasure: 15 to <21 Years (Rate)	LARCmeasure: 15 to <21 Years (Reliability (all units))	LARCmeasure: 15 to <21 Years (Reliability (unit size≥75))	LARCmeasure: 21 to 44 years (LARC Provision)	LARCmeasure: 21 to 44 years (Total N)	LARCmeasure: 21 to 44 years (Rate)	LARCmeasure: 21 to 44 years (Reliability (all units))	LARCmeasure: 21 to 44 years (Reliability (unit size≥75))	LARCmeasure: all age groups (LARC Provision)	LARCmeasure: all age groups (Total N)	LARCmeasure: all age groups (Rate)	LARCmeasure: all age groups (Reliability (all units))	LARCmeasure: all age groups (Reliability (unit size≥75))
41	46	327	0.141	0.982	0.981	78	713	0.109	0.989	0.989	124	1040	0.119	0.993	0.993
42	95	823	0.115	0.993	0.992	224	2180	0.103	0.996	0.996	319	3003	0.106	0.998	0.998
43	115	1258	0.091	0.995	0.995	316	3368	0.094	0.998	0.998	431	4626	0.093	0.999	0.998
44	0	0	*	*	*	0	1	0.000	0.111	*	0	1	0.000	0.126	*
45	96	800	0.120	0.992	0.992	240	2843	0.084	0.997	0.997	336	3643	0.092	0.998	0.998
46	167	1231	0.136	0.995	0.995	347	2840	0.122	0.997	0.997	514	4071	0.126	0.998	0.998
47	1	452	0.002	0.987	0.986	7	2441	0.003	0.997	0.997	8	2893	0.003	0.998	0.998
48	61	437	0.140	0.986	0.985	232	1967	0.118	0.996	0.996	293	2404	0.122	0.997	0.997
49	197	939	0.210	0.994	0.993	519	3332	0.156	0.998	0.998	716	4271	0.168	0.998	0.998
50	108	942	0.115	0.994	0.993	213	2257	0.094	0.996	0.996	321	3199	0.100	0.998	0.998
51	233	1040	0.224	0.994	0.994	576	3432	0.168	0.998	0.998	809	4472	0.181	0.998	0.998
52	3	288	0.010	0.979	0.978	18	1899	0.009	0.996	0.996	21	2187	0.010	0.997	0.997
53	0	342	0.000	0.982	0.981	1	1646	0.001	0.995	0.995	1	1988	0.001	0.997	0.996
54	0	16	0.000	0.724	*	0	79	0.000	0.908	0.907	0	95	0.000	0.932	0.931
55	0	59	0.000	0.906	*	0	213	0.000	0.964	0.963	0	272	0.000	0.975	0.975
56	0	199	0.000	0.970	0.969	0	463	0.000	0.983	0.983	0	662	0.000	0.990	0.989
Total or Mean	**5327**	**32458**	**0.164**	*	*	**14053**	**111060**	**0.127**	*	*	**19380**	**143518**	**0.135**	*	*
*	*	*	*	**Overall Reliability**	**Overall Reliability**	*	*	*	**Overall Reliability**	**Overall Reliability**	*	*	*	**Overall Reliability**	**Overall Reliability**
*	Median n	604	*	**0.974**	**0.986**	Median n	2180	*	**0.974**	**0.992**	Median n	2915	*	**0.978**	**0.995**
*	Min n	0	*	*	*	Min n	1	*	*	*	Min n	1	*	*	*

*cell intentionally left blank

Table 3. Rates and reliabilities for the provision of LARC methods by public health region, Iowa Medicaid Enterprise, 2018.

Public Health Region	LARC measure: 15 to <21 Years (LARC Provision)	LARC measure: 15 to <21 Years (Total N)	LARC measure: 15 to <21 Years (Rate)	LARC measure: 15 to <21 Years (Reliability)	LARC measure: 21 to 44 years (LARC Provision)	LARC measure: 21 to 44 years (Total N)	LARC measure: 21 to 44 years (Rate)	LARC measure: 21 to 44 years (Reliability)	LARC measure: all age groups (LARC Provision)	LARC measure: all age groups (Total N)	LARC measure: all age groups (Rate)	LARC measure: all age groups (Reliability)
1	417	8365	0.050	0.945	1275	25070	0.051	0.852	1692	33435	0.051	0.954
2	174	2247	0.077	0.822	394	6392	0.062	0.595	568	8639	0.066	0.841
3	140	3183	0.044	0.867	421	8915	0.047	0.672	561	12098	0.046	0.881
4	176	2824	0.062	0.853	422	8191	0.052	0.653	598	11015	0.054	0.871
5	246	3609	0.068	0.881	620	11346	0.055	0.723	866	14955	0.058	0.902
6	572	8409	0.068	0.945	1568	28341	0.055	0.867	2140	36750	0.058	0.958
Total or Mean	**1725**	**28637**	**0.060**	*	**4700**	**88255**	**0.053**	*	**6425**	**116892**	**0.055**	*
*	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**
*	Median n	3609	*	**0.886**	Median n	11346	*	**0.727**	Median n	14955	*	**0.901**
*	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 the provision of LARC methods by health plan, Washington State Health Care Authority, 2019.

Health Plan	LARC measure: 15 to <21 Years (LARC Provision)	LARC measure: 15 to <21 Years (Total N)	LARC measure: 15 to <21 Years (Rate)	LARC measure: 15 to <21 Years (Reliability)	LARC measure: 21 to 44 years (LARC Provision)	LARC measure: 21 to 44 years (Total N)	LARC measure: 21 to 44 years (Rate)	LARC measure: 21 to 44 years (Reliability)	LARC measure: all age groups (LARC Provision)	LARC measure: all age groups (Total N)	LARC measure: all age groups (Rate)	LARC measure: all age groups (Reliability)
MCO 1	259	4031	0.064	0.358	935	15357	0.061	0.807	1194	19388	0.062	0.753
MCO 2	535	9684	0.055	0.573	1376	20378	0.068	0.848	1911	30062	0.064	0.825
MCO 3	503	7731	0.065	0.517	1045	15127	0.069	0.805	1548	22858	0.068	0.782
MCO 4	2014	31628	0.064	0.814	5048	73240	0.069	0.952	7062	104868	0.067	0.943
MCO 5	256	4281	0.060	0.372	882	15111	0.058	0.805	1138	19392	0.059	0.753
Total or Mean	**3567**	**57355**	**0.062**	*	**9286**	**139213**	**0.067**	*	**12853**	**196568**	**0.065**	*
*	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**
*	Median n	8707.5	*	**0.527**	Median n	17867.5	*	**0.843**	Median n	26460	*	**0.811**
*	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 the provision of LARC methods by health plan, Massachusetts MassHealth, 2019.

Health Plan	LARC measure: 15 to <21 Years LARC Provision	LARC measure: 15 to <21 Years Total N	LARC measure: 15 to <21 Years Rate	LARC measure: 15 to <21 Years Reliability	LARC measure: 21 to 44 years LARC Provision	LARC measure: 21 to 44 years Total N	LARC measure: 21 to 44 years Rate	LARC measure: 21 to 44 years Reliability	LARC measure: all age groups LARC Provision	LARC measure: all age groups Total N	LARC measure: all age groups Rate	LARC measure: all age groups Reliability
ACOA 1	11	1057	0.010	0.771	331	5237	0.063	0.863	342	6294	0.054	0.938
ACOA 2	28	445	0.063	0.586	204	2934	0.070	0.779	232	3379	0.069	0.891
ACOA 3	15	634	0.024	0.669	228	3201	0.071	0.794	243	3835	0.063	0.902
ACOA 4	193	8036	0.024	0.962	59	858	0.069	0.508	252	8894	0.028	0.955
ACOA 5	46	1444	0.032	0.821	316	3944	0.080	0.826	362	5388	0.067	0.929
ACOA 6	84	1895	0.044	0.858	580	6354	0.091	0.884	664	8249	0.080	0.952
ACOA 7	68	2270	0.030	0.878	391	6408	0.061	0.885	459	8678	0.053	0.954
ACOA 8	25	534	0.047	0.630	213	3446	0.062	0.806	238	3980	0.060	0.906
ACOA 9	43	1557	0.028	0.832	295	4541	0.065	0.845	338	6098	0.055	0.936
ACOA 10	180	4720	0.038	0.938	1241	17779	0.070	0.955	1421	22499	0.063	0.982
ACOA 11	104	1809	0.057	0.852	300	4305	0.070	0.838	404	6114	0.066	0.937
ACOA 12	45	1073	0.042	0.774	450	5403	0.083	0.867	495	6476	0.076	0.940
ACOA 13	37	1235	0.030	0.797	196	3136	0.062	0.791	233	4371	0.053	0.913
ACOB 1	122	4690	0.026	0.937	1150	14398	0.080	0.945	1272	19088	0.067	0.979
ACOB 2	137	5243	0.026	0.943	820	15009	0.055	0.948	957	20252	0.047	0.980
ACOB 3	229	5450	0.042	0.945	1305	16436	0.079	0.952	1534	21886	0.070	0.981
Non-ACO 1	0	0	*	*	199	2884	0.069	0.777	199	2884	0.069	0.874
Non-ACO 2	52	1909	0.027	0.859	625	10238	0.061	0.925	677	12147	0.056	0.967
Non-ACO 3	189	5668	0.033	0.947	886	11833	0.075	0.934	1075	17501	0.061	0.977
Non-ACO 4	43	1265	0.034	0.801	452	7900	0.057	0.905	495	9165	0.054	0.957
Non-ACO 5	0	0	*	*	19	351	0.054	0.297	19	351	0.054	0.459
Total or Mean	**1651**	**50934**	**0.032**	*	**10260**	**146595**	**0.070**	*	**11911**	**197529**	**0.060**	*
*	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**
*	Median n	1683	*	**0.832**	Median n	5320	*	**0.825**	Median n	7362.5	*	**0.919**
*	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 the provision of LARC methods by health plan, Louisiana Medicaid, 2019.

Health Plan	LARC measure: 15 to <21 Years (LARC Provision)	LARC measure: 15 to <21 Years (Total N)	LARC measure: 15 to <21 Years (Rate)	LARC measure: 15 to <21 Years (Reliability)	LARC measure: 21 to 44 years (LARC Provision)	LARC measure: 21 to 44 years (Total N)	LARC measure: 21 to 44 years (Rate)	LARC measure: 21 to 44 years (Reliability)	LARC measure: all age groups (LARC Provision)	LARC measure: all age groups (Total N)	LARC measure: all age groups (Rate)	LARC measure: all age groups (Reliability)
MCO 1	144	3004	0.048	0.219	614	15174	0.040	0.295	758	18178	0.042	0.495
MCO 2	382	10115	0.038	0.486	1054	27867	0.038	0.435	1436	37982	0.038	0.672
MCO 3	522	12636	0.041	0.541	1480	37620	0.039	0.509	2002	50256	0.040	0.730
MCO 4	1290	29880	0.043	0.736	2511	61423	0.041	0.629	3801	91303	0.042	0.831
MCO 5	879	22699	0.039	0.680	2207	58682	0.038	0.618	3086	81381	0.038	0.814
Total or Mean	**3217**	**78334**	**0.041**	*	**7866**	**200766**	**0.039**	*	**11083**	**279100**	**0.040**	*
*	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**	*	*	*	**Overall Reliability**
*	Median n	17667.5	*	**0.532**	Median n	48151	*	**0.497**	Median n	65818.5	*	**0.708**
*	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 the provision of LARC methods 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	**.30***	**.47***	**(.37, .56)**	393	**.28***	**.43***	**(.31, .53)**	270	**.26***	**.39***	**(.26, .51)**
- -	21-44	525	.23*	.44*	(.32, .55)	297	.20*	.40*	(.26, .52)	201	.23*	.37*	(.20, .51)
- -	15-20	202	.38*	.58*	(.41, .72)	96	.54*	.61*	(.41, .76)	56	.66*	.67*	(.42, .82)
Gynecological Examination	15-44	633	**.20***	**.29***	**(.17, .40)**	393	**.18***	**.25***	**(.12, .37)**	270	**.18***	**.20***	**(.04, .34)**
- -	21-44	525	.15*	.20*	(.07, .33)	297	.10	.13	(-.02, .28)	201	.08	.11	(-.07, .28)
- -	15-20	202	.25*	.54*	(.34, .71)	96	.31*	.50*	(.23, .69)	56	.26*	.53*	(.16, .76)
Cervical Cancer Screening†	21-44	523	**.26***	**.37***	**(.24, .48)**	296	**.20***	**.26***	**(.11, .40)**	198	**.24***	**.29***	**(.12, .45)**
Chlamydia Screening†	16-24	186	**13**	**.14**	**(-.06, .32)**	87	**.09**	**.06**	**(-.21, .31)**	53	**.22**	**.18**	**(-.15, .46)**
- -	21-24	82	-.04	.08	(-.22, .36)	40	.10	.22	(-.20, .55)	24	.14	.10	(-.48, .58)
- -	16-20	99	.13	.21	(-.05, .44)	40	.31	.33	(-.05, .61)	27	.20	.06	(-.42, .50)

*statistically significant at p < .05

†Age range of the related measure differs from t of the contraceptive care measure and the analysis was conducted among the overlapping population onl

- - cell intentionally left blank

2016 Submission (All following tables are from 2016 submission)

Table 1. Rates and reliabilities for use of LARC methods, 25 PPFA affiliates, 2014

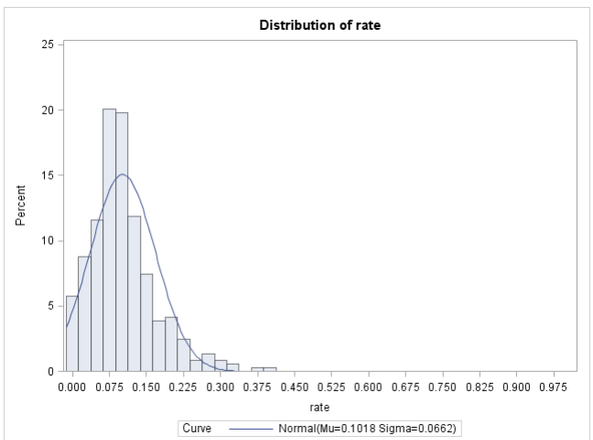
Affiliate ID	LARC measure: 15 to <21 Years	LARC measure: 15 to <21 Years (Used LARC)	LARC measure: 15 to <21 Years (Total N)	LARC measure: 15 to <21 Years (Rate)	LARC measure: 15 to <21 Years (HC Within Affiliate Reliability)	LARC measure: 21 to 45 years	LARC measure: 21 to 45 years (Used LARC)	LARC measure: 21 to 45 years (Total N)	LARC measure: 21 to 45 years (Rate)	LARC measure: 21 to 45 years (HC Within Affiliate Reliability)	LARC measure: all age groups	LARC measure: all age groups (Used LARC)	LARC measure: all age groups (Total N)	LARC measure: all age groups (Rate)	LARC measure: all age groups (HC Within Affiliate Reliability)
1	*	1516	7869	0.193	**0.9896**	*	5464	29638	0.184	**0.9976**	*	6980	37507	0.186	**0.9982**
3	*	1878	26591	0.071	**0.9855**	*	9687	88881	0.109	**0.9959**	*	11565	115472	0.100	**0.9970**
4	*	638	4147	0.154	**0.9832**	*	3073	21430	0.143	**0.9994**	*	3711	25577	0.145	**0.9994**
5	*	4979	42698	0.117	**0.9827**	*	18747	131187	0.143	**0.9937**	*	23726	173885	0.136	**0.9959**
6	*	273	2651	0.103	**0.9243**	*	907	7362	0.123	**0.9741**	*	1180	10013	0.118	**0.9825**
9	*	2035	25268	0.081	**0.9949**	*	10757	88455	0.122	**0.9978**	*	12792	113723	0.112	**0.9984**
10	*	1753	15188	0.115	**0.9408**	*	5839	47698	0.122	**0.9606**	*	7592	62886	0.121	**0.9773**
12	*	552	4839	0.114	**0.9557**	*	1181	10209	0.116	**0.9818**	*	1733	15048	0.115	**0.9861**
37	*	161	1965	0.082	**0.8686**	*	450	4194	0.107	**0.8263**	*	611	6159	0.099	**0.9194**
38	*	452	6093	0.074	**0.9484**	*	1307	10645	0.123	**0.9387**	*	1759	16738	0.105	**0.9685**
40	*	566	5030	0.113	**0.9356**	*	1336	10843	0.123	**0.9487**	*	1902	15873	0.120	**0.9690**
41	*	575	5466	0.105	**0.8980**	*	2448	17562	0.139	**0.9576**	*	3023	23028	0.131	**0.9639**
44	*	1102	11489	0.096	**0.9904**	*	3591	33620	0.107	**0.9952**	*	4693	45109	0.104	**0.9967**
47	*	349	5644	0.062	**0.9897**	*	876	16648	0.053	**0.9848**	*	1225	22292	0.055	**0.9931**
53	*	1032	8741	0.118	**0.9738**	*	2984	28791	0.104	**0.9852**	*	4016	37532	0.107	**0.9908**
54	*	656	3122	0.210	**0.8474**	*	1313	6614	0.199	**0.8879**	*	1969	9736	0.202	**0.9304**
59	*	371	3682	0.101	**0.6688**	*	906	9778	0.093	**0.8604**	*	1277	13460	0.095	**0.8329**
60	*	61	436	0.140	**0.0000**	*	163	1265	0.129	**0.0000**	*	224	1701	0.132	**0.0000**
70	*	145	4154	0.035	**0.9869**	*	454	12436	0.037	**0.9990**	*	599	16590	0.036	**0.9993**
73	*	57	996	0.057	**0.9469**	*	191	2825	0.068	**0.9704**	*	248	3821	0.065	**0.9784**
75	*	98	1171	0.084	**0.9895**	*	442	5070	0.087	**0.9930**	*	540	6241	0.087	**0.9950**
76	*	109	3817	0.029	**0.9762**	*	767	11648	0.066	**0.9803**	*	876	15465	0.057	**0.9876**
77	*	781	11359	0.069	**0.9553**	*	2329	31393	0.074	**0.9799**	*	3110	42752	0.073	**0.9885**
79	*	84	1260	0.067	**0.9760**	*	143	5149	0.028	**0.7775**	*	227	6409	0.035	**0.9796**
81	*	24	294	0.082	**0.0000**	*	134	1561	0.086	**0.0000**	*	158	1855	0.085	**0.0000**
Total or Mean	*	**20247**	**203970**	**0.099**	*	*	**75489**	**634902**	**0.119**	*	*	**95736**	**838872**	**0.114**	*
*	*	*	**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	**Median n**	**4839**	**0.2374**	**0.0673**	**0.9971**	**Median n**	**11648**	**0.2381**	**0.0675**	**0.9988**	**Median n**	**16590**	**0.2163**	**0.0617**	**0.9991**

Patient Volume															
Reliability using Minimum Patient Volume (Floor)	Min n	294	0.2374	0.0673	0.9550	Min n	1265	0.2381	0.0675	0.9892	Min n	1701	0.2163	0.0617	0.9911

*cell intentionally left blank

Table 2. Distributions of rates and ICCs among health centers (n=363) for use of LARC methods among 25 PPFA affiliates, 2014

LARCmeasure: 15 to <21 Years

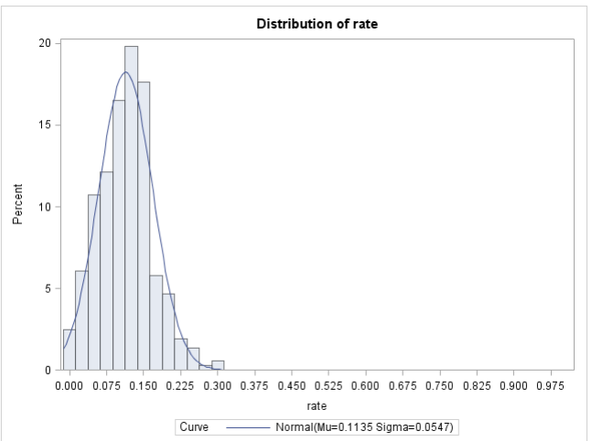


	cases (n)	rate
Mean	561.9	0.10
Median	366	0.09
SD	552.3	0.07
Variance	305,043	0.00
Range	2,976	0.39
Interquartile	629	0.07
Median ICC		0.06

HCs with rate <.02 = 32

Quantile	cases (n)	rate
100% Max	2,984	0.39
95%	1,766	0.23
90%	1,410	0.19
75% Q3	787	0.13

LARCmeasure: 21 to 45 years

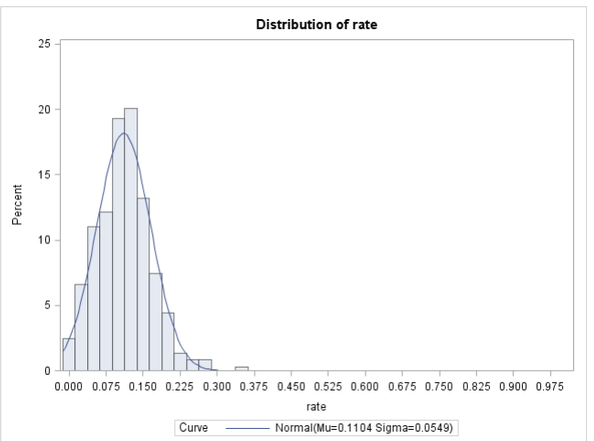


	cases (n)	rate
Mean	1,749.0	0.11
Median	1,016	0.12
SD	1,909	0.05
Variance	3,645,550	0.00
Range	11,360	0.31
Interquartile	2,145	0.07
Median ICC		0.04

HCs with rate <.02 = 11

Quantile	cases (n)	rate
100% Max	11,391	0.31
95%	5,489	0.21
90%	4,544	0.18
75% Q3	2,516	0.15

LARCmeasure: all age groups



	cases (n)	rate
Mean	2,310.9	0.11
Median	1,379	0.11
SD	2,424	0.05
Variance	5,875,321	0.00
Range	13,287	0.35
Interquartile	2,757	0.07
Median ICC		0.05

HCs with rate <.02 = 24

Quantile	cases (n)	rate
100% Max	13,335	0.35
95%	7,198	0.21
90%	5,872	0.18
75% Q3	3,315	0.14

50% Med	366	0.09		50% Med	1,016	0.12		50% Med	1,379	0.11								
25% Q1	158	0.06		25% Q1	371	0.08		25% Q1	558	0.07								
10%	83	0.02		10%	149	0.04		10%	240	0.04								
5%	53	0.01		5%	92	0.03		5%	141	0.03								
0% Min	8	0.00		0% Min	31	0.00		0% Min	48	0.00								

Table 3. Rates and reliabilities for use of LARC method, Iowa Medicaid Enterprise, 2013, by region and type of benefit

Public Health Region	LARCMeasure: 15 to <21 Years (Not Used)	LARCMeasure: 15 to <21 Years (Used LARC)	LARCMeasure: 15 to <21 Years (Total N)	LARCMeasure: 15 to <21 Years (Rate)	LARCMeasure: 15 to <21 Years	LARCMeasure: 15 to <21 Years	LARCMeasure: 21 to 45 years (Not Used)	LARCMeasure: 21 to 45 years (Used LARC)	LARCMeasure: 21 to 45 years (Total N)	LARCMeasure: 21 to 45 years (Rate)	LARCMeasure: 21 to 45 years	LARCMeasure: 21 to 45 years	LARCMeasure: all age groups (Not Used)	LARCMeasure: all age groups (Used LARC)	LARCMeasure: all age groups (Total N)	LARCMeasure: all age groups (Rate)	LARCMeasure: all age groups	LARCMeasure: all age groups
1	3204	256	3460	0.074	*	*	8715	873	9588	0.091	*	*	11919	1129	13048	0.087	*	*
2	1034	120	1154	0.104	*	*	2577	329	2906	0.113	*	*	3611	449	4060	0.111	*	*
3	1096	80	1176	0.068	*	*	2851	324	3175	0.102	*	*	3947	404	4351	0.093	*	*
4	992	95	1087	0.087	*	*	2635	252	2887	0.087	*	*	3627	347	3974	0.087	*	*
5	1566	135	1701	0.079	*	*	3966	393	4359	0.090	*	*	5532	528	6060	0.087	*	*
6	2815	306	3121	0.098	*	*	9132	1004	10136	0.099	*	*	11947	1310	13257	0.099	*	*
Total or Mean	**10707**	**992**	**11699**	**0.085**	*	*	**29876**	**3175**	**33051**	**0.096**	*	*	**40583**	**4167**	**44750**	**0.093**	*	*
*	*	*	*	**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	0.01827	0.0055	**0.8887**	*	Median n	3767	0.005674	0.0017	**0.8666**	*	Median n	5205.5	0.00724	0.0022	**0.9197**
Minimum Patient Volume (Floor)	*	Min n	1087	0.01827	0.0055	0.8579	*	Min n	2887	0.005674	0.0017	0.8327	*	Min n	3974	0.00724	0.0022	0.8974
Type of Benefit	LARCMeasure: 15 to <21 Years (Not Used)	LARCMeasure: 15 to <21 Years (Used LARC)	LARCMeasure: 15 to <21 Years (Total N)	LARCMeasure: 15 to <21 Years (Rate)	LARCMeasure: 15 to <21 Years	LARCMeasure: 15 to <21 Years	LARCMeasure: 21 to 45 years (Not Used)	LARCMeasure: 21 to 45 years (Used LARC)	LARCMeasure: 21 to 45 years (Total N)	LARCMeasure: 21 to 45 years (Rate)	LARCMeasure: 21 to 45 years	LARCMeasure: 21 to 45 years	LARCMeasure: all age groups (Not Used)	LARCMeasure: all age groups (Used LARC)	LARCMeasure: all age groups (Total N)	LARCMeasure: all age groups (Rate)	LARCMeasure: all age groups	LARCMeasure: all age groups
Family Planning Waiver	5698	747	6445	0.116	*	*	20880	2688	23568	0.114	*	*	26578	3435	30013	0.114	*	*
Non-Family Planning Waiver	5009	245	5254	0.047	*	*	8996	487	9483	0.051	*	*	14005	732	14737	0.050	*	*
Total or Mean	**10707**	**992**	**11699**	**0.085**	*	*	**29876**	**3175**	**33051**	**0.096**	*	*	**40583**	**4167**	**44750**	**0.093**	*	*
*	*	*	*	**VarL2**	**ICC**	Benefit type Reliability (Var L2)	*	*	*	**VarL2**	**ICC**	Benefit type Reliability (Var L2)	*	*	*	**VarL2**	**ICC**	Benefit type Reliability (Var L2)
Reliability Based on Median Patient Volume Among Health Centers	*	Median n	5849.5	0.2408	0.0682	**0.9977**	*	Median n	16525.5	0.1867	0.0537	**0.9989**	*	Median n	22375	0.2043	0.0585	**0.9993**

Public Health Region	LARCMeasure: 15 to <21 Years (Not Used)	LARCMeasure: 15 to <21 Years (Used LARC)	LARCMeasure: 15 to <21 Years (Total N)	LARCMeasure: 15 to <21 Years (Rate)	LARCMeasure: 15 to <21 Years	LARCMeasure: 15 to <21 Years	LARCMeasure: 21 to 45 years (Not Used)	LARCMeasure: 21 to 45 years (Used LARC)	LARCMeasure: 21 to 45 years (Total N)	LARCMeasure: 21 to 45 years (Rate)	LARCMeasure: 21 to 45 years	LARCMeasure: 21 to 45 years	LARCMeasure: all age groups (Not Used)	LARCMeasure: all age groups (Used LARC)	LARCMeasure: all age groups (Total N)	LARCMeasure: all age groups (Rate)	LARCMeasure: all age groups	LARCMeasure: all age groups
Calculated Based on Minimum Patient Volume (Floor)	*	Min n	5254	0.2408	0.0682	0.9974	*	Min n	9483	0.1867	0.0537	0.9981	*	Min n	14737	0.2043	0.0585	0.9989
Region 1/Family Planning Waiver	1464	72	1536	0.047	*	*	2421	123	2544	0.048	*	*	3885	195	4080	0.048	*	*
Region 1/Non-Family Planning Waiver	1740	184	1924	0.096	*	*	6294	750	7044	0.106	*	*	8034	934	8968	0.104	*	*
Region 2/Family Planning Waiver	438	25	463	0.054	*	*	617	38	655	0.058	*	*	1055	63	1118	0.056	*	*
Region 2/Non-Family Planning Waiver	596	95	691	0.137	*	*	1960	291	2251	0.129	*	*	2556	386	2942	0.131	*	*
Region 3/Family Planning Waiver	588	21	609	0.034	*	*	726	49	775	0.063	*	*	1314	70	1384	0.051	*	*
Region 3/Non-Family Planning Waiver	508	59	567	0.104	*	*	2125	275	2400	0.115	*	*	2633	334	2967	0.113	*	*
Region 4/Family Planning Waiver	511	25	536	0.047	*	*	857	44	901	0.049	*	*	1368	69	1437	0.048	*	*
Region 4/Non-Family Planning Waiver	481	70	551	0.127	*	*	1778	208	1986	0.105	*	*	2259	278	2537	0.110	*	*
Region 5/Family Planning Waiver	702	39	741	0.053	*	*	1423	74	1497	0.049	*	*	2125	113	2238	0.050	*	*
Region 5/Non-Family Planning Waiver	864	96	960	0.100	*	*	2543	319	2862	0.111	*	*	3407	415	3822	0.109	*	*
Region 6/Family Planning Waiver	1306	63	1369	0.046	*	*	2952	159	3111	0.051	*	*	4258	222	4480	0.050	*	*
Region 6/Non-Family Planning Waiver	1509	243	1752	0.139	*	*	6180	845	7025	0.120	*	*	7689	1088	8777	0.124	*	*
Total or Mean	**10707**	**992**	**11699**	**0.085**	*	*	**29876**	**3175**	**33051**	**0.096**	*	*	**40583**	**4167**	**44750**	**0.093**	*	*
*	*	*	*	**VarL2**	**ICC**	Region by benefit type Reliability (Var L2)	*	*	*	**VarL2**	**ICC**	Region by benefit type Reliability (Var L2)	*	*	*	**VarL2**	**ICC**	Region by benefit type Reliability (Var L2)
Reliability Based on Median Patient Volume Among Health Centers	*	Median n	716	0.2537	0.0716	**0.9822**	*	Median n	2325.5	0.1775	0.0512	**0.9921**	*	Median n	2954.5	0.2003	0.0574	**0.9945**
Calculated Based on Minimum Patient Volume (Floor)	*	Min n	463	0.2537	0.0716	0.9728	*	Min n	655	0.1775	0.0512	0.9725	*	Min n	1118	0.2003	0.0574	0.9855

*cell intentionally left blank

Table 4. Rates and reliabilities for use of LARC method, Wisconsin Medicaid, 2014, by health plan/HMO

HMO	LARCMeasure: 15 to <21 Years (Not Used)	LARCMeasure: 15 to <21 Years (Used LARC)	LARCMeasure: 15 to <21 Years (Total N)	LARCMeasure: 15 to <21 Years (Rate)	LARCMeasure: 15 to <21 Years	LARCMeasure: 15 to <21 Years	LARCMeasure: 21 to 45 years (Not Used)	LARCMeasure: 21 to 45 years (Used LARC)	LARCMeasure: 21 to 45 years (Total N)	LARCMeasure: 21 to 45 years (Rate)	LARCMeasure: 21 to 45 years	LARCMeasure: 21 to 45 years	LARCMeasure: all age groups (Not Used)	LARCMeasure: all age groups (Used LARC)	LARCMeasure: all age groups (Total N)	LARCMeasure: all age groups (Rate)	LARCMeasure: all age groups	LARCMeasure: all age groups
1	4598	234	4832	0.048	*	*	12894	1149	14043	0.082	*	*	17492	1383	18875	0.073	*	*
2	1742	96	1838	0.052	*	*	5314	374	5688	0.066	*	*	7056	470	7526	0.062	*	*
3	861	59	920	0.064	*	*	2633	229	2862	0.080	*	*	3494	288	3782	0.076	*	*
4	1682	113	1795	0.063	*	*	5188	493	5681	0.087	*	*	6870	606	7476	0.081	*	*
5	1147	84	1231	0.068	*	*	3673	263	3936	0.067	*	*	4820	347	5167	0.067	*	*
6	203	16	219	0.073	*	*	662	63	725	0.087	*	*	865	79	944	0.084	*	*
7	518	40	558	0.072	*	*	1475	133	1608	0.083	*	*	1993	173	2166	0.080	*	*
8	326	26	352	0.074	*	*	1001	95	1096	0.087	*	*	1327	121	1448	0.084	*	*
9	1539	84	1623	0.052	*	*	5767	397	6164	0.064	*	*	7306	481	7787	0.062	*	*
10	572	46	618	0.074	*	*	1524	159	1683	0.094	*	*	2096	205	2301	0.089	*	*
11	4621	277	4898	0.057	*	*	13996	1170	15166	0.077	*	*	18617	1447	20064	0.072	*	*
12	1167	72	1239	0.058	*	*	4027	263	4290	0.061	*	*	5194	335	5529	0.061	*	*
13	246	23	269	0.086	*	*	775	78	853	0.091	*	*	1021	101	1122	0.090	*	*
14	2009	140	2149	0.065	*	*	5168	428	5596	0.076	*	*	7177	568	7745	0.073	*	*
15	52	4	56	0.071	*	*	226	14	240	0.058	*	*	278	18	296	0.061	*	*
16	4860	254	5114	0.050	*	*	17460	1415	18875	0.075	*	*	22320	1669	23989	0.070	*	*
17	517	42	559	0.075	*	*	1346	187	1533	0.122	*	*	1863	229	2092	0.109	*	*
Total or Mean	**26660**	**1610**	**28270**	**0.057**	*	*	**83129**	**6910**	**90039**	**0.077**	*	*	**109789**	**8520**	**118309**	**0.072**	*	*
*	*	*	*	**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	0.01418	0.0043	**0.8414**	*	Median n	3936	0.02718	0.0082	**0.9702**	*	Median n	5167	0.02218	0.0067	**0.9721**
Minimum Patient Volume (Floor)	*	Min n	56	0.01418	0.0043	0.1944	*	Min n	240	0.02718	0.0082	0.6647	*	Min n	296	0.02218	0.0067	0.6662

*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 this access to LARC measure. 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 patient-reported outcome performance measure that evaluates the patient-centered 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 #2904 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. In 2018, NQF #2904 became publicly reported as part of CMS' Adult and Child Core Sets of Health Care Quality Measures. 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 NQF #2904, as several grantees shared the codes 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 #2904 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. 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 #2904 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 #2904, 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 #2904 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 #2904 for a quality improvement initiative, only to find that some providers did not see many clients who wish to use a LARC method. The initiative may have also been affected by concurrent statewide and provider-based initiatives to improve access to LARC as well as most and moderately effective methods, and application deadline 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 in a year prior to the measurement year, but 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) to examine access to LARC methods. 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.

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 Maternal and Infant Health Initiative https://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Quality-of-Care/Maternal-and-Infant-Health-Care-Quality.html Quality Improvement (Internal to the specific organization) 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 Washington State Health Care Authority https://www.hca.wa.gov/about-hca/reproductive-health OPA Title X Family Planning Program https://rhntc.org/resources/contraceptive-access-change-package https://opa.hhs.gov/evaluation-research/title-x-services-research/family-planning-annual-report OPA Title X Family Planning Program

*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 #2904 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) and describe one planned use for NQF #2904 in the Core Quality Measure Collaborative.

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 #2904 was added in 2018, which allows all 50 states to report the measure scores on a voluntary basis. While CMCS has collected NQF #2904 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 #2904 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 #2904 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 #2904 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 #2904 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 dually eligible was continuously enrolled in traditional fee-for-service Medicaid. Since 2017, LM has calculated and publicly reported NQF #2904 by health plan via its Medicaid Quality Dashboard [1]. In 2019, about 279,100 eligible women ages 15-44 were included in the NQF #2904 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 #2904 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 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 over 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 #2904 at the health plan level for the past six years. Approximately 196,568 eligible women ages 15-44 comprise the NQF #2904 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 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 #2904 results are calculated at the level of facility for internal quality improvement, and about 31,084 women ages 15-44 comprise the NQF #2904 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 #2904 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 this application.

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 #2904 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 #2904 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 #2903 (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].

Core Quality Measure Collaborative (CQMC) – Planned Use

The CQMC is a diverse coalition of health care leaders representing over 75 consumer groups, medical associations, health insurance providers, purchasers, and other quality stakeholders, all working together to develop and recommend core sets of measures by clinical area to assess and improve the quality of health care nationwide. Convened in 2015 by America's Health Insurance Providers (AHIP) and the CMS, CQMC is housed at NQF. In the second half of 2020, CQMC released updated core measure sets for specific clinical areas after a careful consensus-based review and deliberation among the collaborative's member organizations against CQMC's rigorous inclusion criteria. CQMC intends for its core sets of measures to be used in value-based payment programs, reported at the clinician level in outpatient settings, and could support multiple care delivery models. However, some measures selected for CMQC core sets focus on the inpatient setting and are endorsed by NQF at the levels of facility and health plan. Along with NQF #2902, the current CQMC Obstetrics and Gynecology core measure set added NQF #2904 for its members to use for quality assessment [4].

References

[1] Louisiana Department of Health. (n.d.). Medicaid Managed Care Quality Dashboard. Louisiana Department of Health. 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. (n.d.). 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

[4] National Quality Forum. (n.d.). Core Quality Measures Collaborative. National Quality Forum. Retrieved December 22, 2020 from http://www.qualityforum.org/cqmc

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 #2904, 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 #2904 alongside its two complementary measures (NQF #2902 and #2903) and emphasize 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 LARC 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 #2904, 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 #2904 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/long-acting-reversible). The latest specification available is for measurement year 2019. OPA updates its measure pages after annually updating the measure specification, code sets, and syntax.

Users can submit questions to OPA about NQF #2904 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 #2904 to OPA to respond directly to users needing help with measure calculation and interpretation. Most requests came from state Medicaid programs reporting measure scores for CMS Adult and Child Core Sets of Health Care Quality Measures.

Starting in 2016, OPA has provided technical assistance to state Medicaid programs calculating NQF #2904. 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 #2904 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 #2904 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 #2904 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 #2904's calculation, importance, 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 #2904 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 #2904 and OPA's contraceptive care measures (https://www.nationalfamilyplanning.org/file/Onepager_Contraceptive-Measures_-Messages-for-Health-Care-Settings.pdf).

PPFA´s 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 #2904 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 #2904 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 #2904 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 contraceptive provision measures and answer any questions states have about calculating and reporting on the measures.

CMS' Center for Medicaid and 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 #2904 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, including LARC methods, 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). Some sites achieved

this improvement through implementation of strategies to improve LARC access. The four best practices identified in the Contraceptive Access Change Package were:

1) Stock a broad range of contraceptive methods;
2) Discuss pregnancy intention and support patients through evidence-informed, patient-centered counseling;
3) Develop systems for same-visit provision of all contraceptive methods, at all visit types;
4) 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 #2903 (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 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 #2904 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 #2904 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 #2904 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 #2904. 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 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 intrauterine devices and contraceptive implants, which may simplify the measure code sets and numerator calculation.
- Consider state-specific policies for coding administrative claims for prescription contraceptive medications for 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 #2904 numerator).
- As described in 3c.1, multiple states stated that the calculation of NQF #2904 was complex and time-consuming, even with OPA's published SAS programming code. While the syntax has been simplified since NQF #2904'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 #2904 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 #2904 scores publicly for analysis and discussion.
- OPA continues to receive feedback on appropriate implementation and 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. As we specifically state on our website, NQF #2904 "should be used as an access measure; very low rates (less than 1-2%) may signal barriers to LARC provision that should be addressed through training … [and] quality improvement processes" (https://opa.hhs.gov/evaluation-research/title-x-services-research/contraceptive-care-measures/long-acting-reversible). OPA also notes that the measure "should NOT be used to encourage high rates of use as this may lead to coercive practices. This is especially important given the historical context of coercive practices related to contraception" (https://opa.hhs.gov/evaluation-research/title-x-services-research/contraceptive-care-measures/long-acting-reversible). OPA encourages states to use NQF #2904 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 #2904 with the new PCCC measure.

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

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 obtained a LARC method prior to the measurement year. These users also asked if it makes sense to only count clients receiving a LARC method during the year.

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.

The Generic Product Identifier (GPI) code system requires a license fee to utilize, which may not be possible for all states calculating NQF #2904 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 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 responded to users with questions about a lookback period by explaining that measure does not count LARC methods that are "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 reporting units to identify very low rates, and it enables year over year comparisons. Thus, for the purposes of identifying units with LARC provision less than 2% that could use a quality improvement intervention, the current specification is appropriate.

For this application, OPA calculated NQF #2904 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 #2904 scores by client characteristics to monitor quality improvement initiatives and identify reporting units with very low rates of LARC provision to women who wish to use these highly effective methods. To address the concerns around appropriate measure implementation and interpretation, OPA will continue to promote use of NQF #2904 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 #2904 with the new PCCC measure.

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.

As the steward for NQF #2904, we at HHS Office of Population Affairs (OPA) have noted that the measure "should be used as an access measure; very low rates (less than 1-2%) may signal barriers to LARC provision that should be addressed through training …[and] quality improvement processes" [1].

In the United States, policy changes have alleviated some barriers to LARC provision. One study conducted using data from community health centers (e.g. federally qualified health centers, rural health centers, county health departments) reports that an improvement in LARC use to more than 2% occurred after implementation of the U.S. Affordable Care Act (ACA) in states without Medicaid expansion (2013: 1.8%; 2014: 2.2%, 2016: 2.4%) that were included in the analysis. Medicaid expansion states in the study population also experienced an increase in #2904 scores during this time, but those percentages were already greater than 2%. Adolescents receiving services in Medicaid expansion states experienced a larger increase in LARC use after ACA implementation than their counterparts in non-expansion states. Very low to low LARC provision rates persisted in adolescents in this study who obtained care in non-expansion states (2013: 1.80%; 2014: 1.87%; 2016: 2.10%). Overall, this analysis indicates that an association exists between Medicaid expansion and improvements in access to and use of LARC [6]. However, the data presented in 1b illustrates that several programs still have LARC provision rates that are less than 2%, which signals the need for continued efforts to expand access to LARC.

The following programs have LARC provision rates that are less than 2%:

- The Iowa Medicaid Enterprise (IME) population. The LARC rate was almost twice as high in the state-funded family planning program compared to the general Medicaid program, which suggests that it may be worth investigating potential barriers to LARC provision in the general Medicaid program.
- The New-York Presbyterian (NYP) Hospital/Columbia University Irving Medical Center Ambulatory Care Network (ACN) population. Although this network does not specialize in family planning services, about 58% of its primary care facilities provided LARC to less than 2% of its female patients of reproductive age. While some access to LARC exists within NYP ACN, there might be possible barriers to LARC provision to explore in the facilities with very low rates.

- Data from the Planned Parenthood Federation of America (PPFA) indicate that a few health centers had LARC rates below 2%, which suggests that there may be some locations within these two affiliates in which clients may not have adequate access to LARC methods. This may be in part due to the Final Rule, which led to a substantial number of PPFA facilities losing access to Title X funds for services related to contraception [4, 5]. Due to the availability of the LARC access measure, PPFA can identify and follow up with these affiliates and facilities to assess what barriers may exist and determine how to overcome them so that clients are given the opportunity to obtain LARC if they choose to do so.
- The number of Title X grantees with LARC rates below 2% increase in 2019, possibly due to the Title X regulations that promoted single method providers, as long as the Title X project as a whole provided the full range of contraceptive methods [4, 5]. This may mean low-income women receiving care at a federally-funded Title X clinic do not have access to the full range of contraceptive methods, including LARC methods.

Although IME had a very low NQF #2904 measure score, some clinician group/practices had 100% LARC provision. 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 LARC 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.

It is important to note that the developers of the measure contend that NQF #2904 should be used only to monitor access to LARC; and that it could be harmful to set a high benchmark for this measure, because doing so may incentivize coercive practices [2, 3].

References

[1] Office of Population Affairs. (n.d.). Long-Acting Reversible Contraceptive Methods. 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/contraceptive-care-measures/long-acting-reversible.

[2] 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

[3] Gold, R.B. (2014). Guarding Against Coercion While Ensuring Access: A Delicate Balance. Guttmacher Policy Review, 17(3), 8-14.

[4] Gold, R.B., Cross, L. (2019, August 29). 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

[5] Frederiksen B, Salganicoff A, Ramaswamy A, Gomez I. (2020, December 21). 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/

[6] 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 a LARC method [1-3]. OPA reaffirms our commitment to client-centered care through the following actions taken during development and testing of NQF #2904.

Although existing research [4,5] show a high percentage of women will choose LARC when given the opportunity, the focus of this measure is on ensuring access to LARC by monitoring very low rates of use (e.g., below 2%). Further, we explicitly state on the measure website that this measure should not have a benchmark encouraging high rates of use, and that it would be an inappropriate measure to use in pay-for-performance or similar programs. If the measure is used as intended (i.e., to assess lack of access), this should remove pressure on providers to inappropriately "promote" LARC methods.

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 funding 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 (PCCC) measure (NQF #3543), 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 inferior 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 #2904 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 from 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

2903 : Contraceptive Care – Most & Moderately Effective Methods

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 #2903 and focuses on use of most (sterilization, IUD, implant) and moderately (injectable, pill, patch, ring) effective methods of contraception, of which LARC methods are a subset.

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 underscore that the primary intent of the LARC measure is to identify populations in which LARC use is noticeably low to determine if access is limited. It could be harmful to set a high benchmark for this measure, because doing so may incentivize coercive practices related to contraception [1-3]. 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 [4]. 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 #2904 in tandem use.

We share UCSF's hypothesis that the PCCC will serve as a balancing measure for the contraceptive 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 LARC 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.

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] University of California San Francisco. The Person-Centered Contraceptive Counseling Measure. https://pcccmeasure.ucsf.edu/. Accessed 22 Dec 2020.

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_2904_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 Division of Reproductive Health: Jiajia Chen PhD, Shanna Cox MSPH, Gladys Martinez PhD, Kimberly Daniels PhD, Ekwutosi Okoroh, MD

MPH, Antoinette Nguyen MD MPH FACOG, Lisa Romero PhD

HHS Health Resources and Services Administration: Rui Li PhD

Iowa Department of Public Health and Iowa Medicaid Enterprise: Debra Kane PhD RN, 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, 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 F. 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 PhD MPH, Emily J. Decker MPH, Lorretta E. Gavin PhD MPH.

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.