NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

WHAT GOOD LOOKS LIKE — EVIDENCE — PROCESS EXAMPLE #1

The following example is intended only for illustration of the type of information requested for the Steering Committee's evaluation of the evidence. The examples are not intended as requirements or to be replicated exactly—the key point is to provide substantive information and data in the measure submission evidence attachment so it is clear about the evidence that does or does not exist to support the measure focus.

Please contact NQF staff if you have questions, corrections, or suggestions to improve the example.

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number
Measure Title: Women with urinary incontinence who receive pelvic floor muscle training
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: 6/27/2013

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).

EXAMPLE

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

- ☐ Health outcome: Click here to name the health outcome
- ☐ Patient-reported outcome (PRO): Click here to name the PRO
  PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors
- ☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☒ Process: pelvic floor muscle training for urinary incontinence in women
☐ Structure: Click here to name the structure
☐ Other: Click here to name what is being measured

_____

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.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) and 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

Key Points

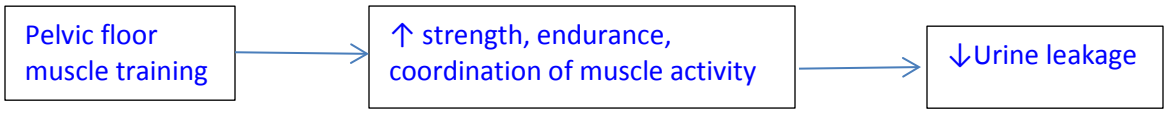
- See NQF guidance for rating quantity, quality, consistency of body of evidence and report from the evidence task force available at the Measure Evaluation webpage.
- A **systematic review** of the evidence is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include quantitative synthesis (meta-analysis), depending on available data (IOM, 2011).
- A body of evidence includes all the evidence for a topic, which is systematically identified, based on pre-established criteria for relevance and quality of evidence.
- Expert opinion is not considered empirical evidence, but evidence is not limited to randomized controlled trials
- There is variability in evidence reviews, grading systems, and presentation of the findings; however, the information should be reported as requested in this form so the Steering Committee can evaluate it according to NQF criteria and guidance.

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.

Key Points

- Indicate the causal pathway – do not just make a general statement.

- Do not discuss evidence in this item – it should be presented in the appropriate sections as indicated by the source of the evidence noted in 1a3.1.



PFMT may be prescribed to increase strength (the maximum force generated by a muscle in a single contraction); endurance (ability to contract repetitively, or sustain a single contraction over time); coordination of muscle activity, or timing to suppress urge, or a combination of these.

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*):

Incontinence in women. In: Schröder A, Abrams P, Andersson KE, Artibani W, Chapple CR, Drake MJ, Hampel C, Neisius A, Tubaro A, Thüroff JW. Guidelines on urinary incontinence. Arnhem, The Netherlands: European Association of Urology (EAU); 2009 Mar. p. 28-43.

URL: http://www.uroweb.org/gls/pdf/Urinary%20Incontinence%202010.pdf
http://www.guideline.gov/content.aspx?id=16386&search=urinary+incontinence#Section424

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

Key Points
- Do not summarize, paraphrase, or shorten the recommendation

5.2 Initial treatment of urinary incontinence (UI) in women (p.29)
For women with stress, urgency or mixed urinary incontinence, initial treatment includes appropriate lifestyle advice, physical therapy, a scheduled voiding regime, behavioural therapy and medication (Table 7, Figure 3). Some recommendations are based on good and consistent evidence of effect. However, many other recommendations are based on insufficient level 1 or 2 evidence and are essentially hypotheses requiring better evidence of their benefit.

From Table 7: Initial treatment for UI in women
PFMT should be offered as first-line conservative therapy to women with stress, urgency, or mixed UI

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

Key Points
• Should include BOTH grade and definition of the grade • Not all grades are on a letter or number scale • Grades for a recommendation are different from grades for quality of evidence (although related) – make sure it is the appropriate grade for a recommendation

A - Based on clinical studies of good quality and consistency addressing the specific recommendations and including at least one randomised trial

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.*)

B - Based on well-conducted clinical studies, but without randomised clinical trials
C - Made despite the absence of directly applicable clinical studies of good quality
Modified from Sackett et al. (2, 3).

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

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*):

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

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

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.*)

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

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*):

Cochrane Systematic Review
Dumoulin C, Hay-Smith J; Pelvic floor muscle training versus no treatment or inactive control treatments for urinary incontinence in women; Cochrane Database of Systematic Reviews 2010, Issue 1. Art. No.: CD005654, DOI: 10.1002/14651858.CD005654.pub2.

URL:
http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD005654.pub2/abstract;jsessionid=5B59498CFC062003F250C00B2B8CEFEE.d01t03

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**
The systematic review identified quality of evidence based on risk of bias. System for determining risk of bias was explained in Chapter 8 of Cochrane Handbook for Systematic Reviews for Interventions, 5.0.2, updated September 2009 http://www.mrc-bsu.cam.ac.uk/cochrane/handbook502 .

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.

Key Points

- If more than one systematic review of the evidence identified above (in 1a.4, 1a.5, and 1a.6), you may choose to summarize below 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.
- If more than one systematic review of the evidence is summarized below, provide a separate response for each review for each question and clearly identify which review is the basis of the response – do not combine systematic reviews.
- If the only systematic review of the body of evidence relevant to your measure does not make details available about the quantity, quality, and consistency of the body of evidence; respond to the following questions with what is known from the systematic review. (For example, it is not useful to report that 5,000 articles were reviewed for an entire guideline because it provides no information on the quantity of studies in the body of evidence for a particular process of care.)

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

The information in the following questions in this section is based on the Cochrane Systematic Review cited in section 1a.6.

The systematic review addressed pelvic floor muscle training for women with urinary incontinence in comparison to no treatment, placebo or sham treatments, or other inactive control treatments.

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

Key Points
- Should include BOTH grade and definition of the grade
- Not all grades are on a letter or number scale
- Grades for quality of evidence are different from grades for the recommendation (although related) – make sure it is the appropriate grade for the quality of the body of evidence

An overall grade of methodological quality was not assigned. In the systematic review, individual study quality was graded on a scale for risk of bias.

Based on the reported adequacy of allocation concealment and blinding, two trials appeared to be at low risk (Bø 1999; Castro, 2008), six at moderate risk (Bidmead 2002; Burgio 1998; Burns 1993; Kim 2007; Miller 1998; ; Yoon 2003;), and six at high or possible high risk of bias (Aksac 2003; Henalla 1989; Henalla 1990;Hofbauer 1990; Lagro-Janssen 1991;Wells 1999). Interestingly, the more recent trials tended to be of lower risk for bias based on the trial reports." (p. 20)

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

Risk of bias	Interpretation	Within a study	Across studies
Low risk of bias.	Plausible bias unlikely to seriously alter the results.	Low risk of bias for all key domains.	Most information is from studies at low risk of bias.
Unclear risk of bias.	Plausible bias that raises some doubt about the results.	Unclear risk of bias for one or more key domains.	Most information is from studies at low or unclear risk of bias.
High risk of bias.	Plausible bias that seriously weakens confidence in the results.	High risk of bias for one or more key domains.	The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results.

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: <u>1998-2008</u>

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*)

14 randomized controlled trials

1a.7.6. What is the overall quality of evidence <u>across studies</u> 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*)

Based on the reported adequacy of allocation concealment and blinding, two trials appeared to be at low risk (Bø 1999; Castro, 2008), six at moderate risk (Bidmead 2002; Burgio 1998; Burns 1993; Kim 2007; Miller 1998; ; Yoon 2003;), and six at high or possible high risk of bias (Aksac 2003; Henalla 1989; Henalla 1990;Hofbauer 1990; Lagro-Janssen 1991;Wells 1999). Interestingly, the more recent trials tended to be of lower risk for bias based on the trial reports." (p. 20)

Methodological quality was evaluated from the trial reports. Therefore, the quality of reporting might have affected the judgement of methodological quality. Two of the included studies were published only as abstracts (Bidmead 2002; Henalla 1990). Limited methodological detail was given, which made it particularly difficult to judge the quality of these trials. In addition, few data were reported.
In one way, it was disappointing that only two trials sufficiently described the randomisation process so that the review authors could be sure there was adequate concealment. On the other hand, it was encouraging, given the difficulties of blinding participants and treatment providers to PFMT, that eight of the 14 studies used blinded outcome assessors. Generally, the proportion of dropout and withdrawals was in the region of 0 to 20%. Sample sizes were small to moderate in 12 of the 14 studies, and only three trials reported an a priori power calculation. Two trials stated that intention to treat principles were used for the primary analysis, and one stated that intention to treat analysis did not change the findings of the primary analysis.

Sensitivity analysis on the basis of trial quality was not considered appropriate in view of the small number of trials contributing to each comparison. It is not known to what extent the variable quality of the trials has affected the findings of the review. It is interesting to note that of all the studies contributing data to the analysis, the largest treatment effect (for cure and improvement, and leakage episodes) was observed in a trial at the high risk of bias. This might be an example of the apparent overestimation of treatment effect (about 30%) observed in trials with inadequate or unclear concealment of random allocation (Egger 2002)." (p. 20)

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) <u>across studies</u> 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*)

Meta-analysis was not possible due to study heterogeneity.

Comparison of PFMT versus no treatment, placebo, or control was studied for a variety of outcomes as follows:

Outcome: Patient Perceived Cure – 2 studies with consistent direction in favor of PFMT but differences in magnitude of effect (risk ratio 2.34-16.80)

Outcome: Patient Perceived Cure or Improvement – 3 studies with consistent direction in favor of PFMT but differences in magnitude of effect (risk ratio 2.26-20.0). The authors concluded "Overall, the

differences in likelihood of cure or improvement after PFMT compared to control suggested by the review are sufficient to be of interest to women." (p.18)

Outcome: QoL – 2 studies
Hopkins Symptom Checklist, for psychological distress (SCL-90-R)
Global severity: 50.8 (12.8) vs. 51.4 (10.9); mean difference -0.6, 95% CI -5.3 to 4.1

Norwegian Quality of Life Scale
90.1 (9.5) vs. 85.2 (12.1); mean difference 4.9, 95%CI -1.1 to 10.9

The authors concluded "Based on evidence from single trials, there is improved condition specific QoL in women treated with PFMT compared to controls, but there might be less or no effect on generic QoL." (p.18)

Outcome: Leakage Episodes – 5 studies with consistent direction in favor of PFMT but differences in magnitude of effect. "there were statistically significantly fewer leakage episodes (-0.77 to -2.92) with PFMT" (p.18)

Outcome: Number of Voids per Day – 1 study with significantly fewer (-3.1) with PFMT

Outcome: Number of Voids per Night – 1 study with no significant difference

Outcome: Short pad Test Number Cured – 3 studies with consistent direction in favor of PFMT but differences in magnitude of effect (risk ratios 5.54-16.24)

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

Three of four studies that reported adverse events stated there were none with PFMT. The other trial recorded a few minor effects of PFMT (for example discomfort with training), and all of which were reversible with cessation of training. Although randomized trials are probably not the most appropriate way to address safety, neither these data nor the content of PFMT suggest that PFMT is likely to be unsafe. (p. 19)

The authors concluded that "PFMT is better than no treatment, placebo, drug, or inactive control for women with stress, urge, or mixed incontinence. Women treated with PFMT were more likely to report cure or improvement, report better QoL, have ewer leakage episodes per day and have less urine leakage on short pad tests than controls. (p.21)

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 <u>each</u> new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

_____

1a.8 OTHER SOURCE OF EVIDENCE

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

1a.8.1 What process was used to identify the evidence?

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

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

WHAT GOOD LOOKS LIKE — EVIDENCE — PROCESS EXAMPLE #2

The following example is intended only for illustration of the type of information requested for the Steering Committee's evaluation of the evidence. The examples are not intended as requirements or to be replicated exactly—the key point is to provide substantive information and data in the measure submission evidence attachment so it is clear about the evidence that does or does not exist to support the measure focus.

Please contact NQF staff if you have questions, corrections, or suggestions to improve the example.

Measure Number (*if previously endorsed*)**:** 38T
Measure Title: Periconception folic acid supplementation
IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: 38T

Date of Submission: 6/27/2013

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).

EXAMPLE

1a.1.This is a measure of: (*should be consistent with type entered in De.1*)
Outcome
- ☐ Health outcome: 38T
- ☐ Patient-reported outcome (PRO): 38T
  PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors
- ☐ Intermediate clinical outcome (*e.g., lab value*): 38T

☒ Process: Folic acid supplements for women who may become pregnant and in early pregnancy to prevent neural tube defects

☐ Structure: 38T

☐ Other: 38T

_____

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.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) and 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.

_____

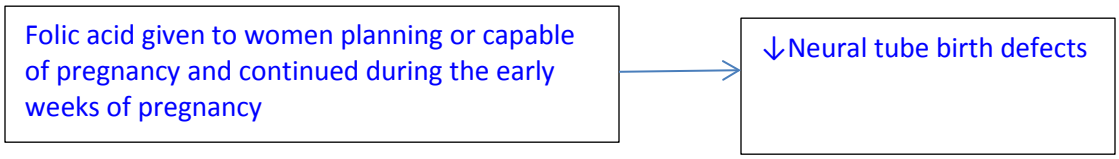
INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

Key Points
- See NQF guidance for rating quantity, quality, consistency of body of evidence and report from the evidence task force available at the Measure Evaluation webpage.
- A **systematic review** of the evidence 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 quantitative synthesis (meta-analysis), depending on available data (IOM, 2011).
- A body of evidence includes all the evidence for a topic, which is systematically identified, based on pre-established criteria for relevance and quality of evidence.
- Expert opinion is not considered empirical evidence, but evidence is not limited to randomized controlled trials
- There is variability in evidence reviews, grading systems, and presentation of the findings; however, the information should be reported as requested in this form so the Steering Committee can evaluate it according to NQF criteria and guidance.

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.

Key Points
- Indicate the causal pathway – do not just make a general statement.

- Do not discuss evidence in this item – it should be presented in the appropriate sections as indicated by the source of the evidence indicated in 1a3.1.



Folic acid supplementation in women planning or capable of pregnancy,and continued during the early weeks of pregnancy reduces the occurence of neural tube birth defects.

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*):

American College of Obstetricians and Gynecologists (ACOG). Neural tube defects. Washington (DC): American College of Obstetricians and Gynecologists (ACOG); 2003 Jul. 11 p. (ACOG practice bulletin; no. 44). [81 references]

URL: http://www.guideline.gov/content.aspx?id=3994&search=folic+acid+supplement

The American College of Obstetricians and Gynecologists (ACOG) reaffirmed the currency of the guideline in 2008.

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

Key Points
- Do not summarize, paraphrase, or shorten the recommendation

Practice bulletin no. 44
Periconceptional folic acid supplementation is recommended because it has been shown to reduce the occurrence and recurrence of neural tube defects (NTDs).

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

Key Points

- Should include BOTH grade and definition of the grade
- Not all grades are on a letter or number scale
- Grades for a recommendation are different from grades for quality of evidence (although related) – make sure it is the appropriate grade for a recommendation

Level A - Recommendation is based on good and consistent scientific evidence

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.*)

Level B - Recommendations are based on limited or inconsistent scientific evidence.
Level C - Recommendations are based primarily on consensus and expert opinion.

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

Same URL as in 1a.4.1 http://www.guideline.gov/content.aspx?id=3994&search=folic+acid+supplement
See tab for related content

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*):

U.S. Preventive Services Task Force. Folic Acid for the Prevention of Neural Tube Defects: U.S. Preventive Services Task Force Recommendation Statement. AHRQ Publication No. 09-05132-EF-2, May 2009. http://www.uspreventiveservicestaskforce.org/uspstf09/folicacid/folicacidrs.htm
Folic Acid for the Prevention of Neural Tube Defects: U.S. Preventive Services Task Force Recommendation Statement U.S. Preventive Services Task Force Ann Intern Med May 5, 2009 150:626-631.

URL: http://www.uspreventiveservicestaskforce.org/uspstf/uspsnrfol.htm

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

No numbering provided; date is May 2009
The USPSTF recommends that all women planning or capable of pregnancy take a daily supplement containing 0.4 to 0.8 mg (400 to 800 µg) of folic acid.

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

A - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. Offer or provide this service.

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.*)

Grade	Definition	Suggestions for Practice
A	The USPSTF recommends the service. There is high certainty that the net benefit is substantial.	Offer or provide this service.
B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.	Offer or provide this service.
C	*Note: The following statement is undergoing revision.* Clinicians may provide this service to selected patients depending on individual circumstances. However, for most individuals without signs or symptoms there is likely to be only a small benefit from this service.	Offer or provide this service only if other considerations support the offering or providing the service in an individual patient.
D	The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.	Discourage the use of this service.
I Statement	The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.	Read the clinical considerations section of USPSTF Recommendation Statement. If the service is offered, patients should understand the uncertainty about the balance of benefits and harms.

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

URL: http://www.uspreventiveservicestaskforce.org/uspstf/grades.htm

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*):

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

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.

Key Points

- If more than one systematic review of the evidence identified above (in 1a.4, 1a.5, and 1a.6), you may choose to summarize below 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.
- If more than one systematic review of the evidence is summarized below, provide a separate response for each review for each question and clearly identify which review is the basis of the response – do not combine systematic reviews.
- If the only systematic review of the body of evidence relevant to your measure does not make details available about the quantity, quality, and consistency of the body of evidence; respond to the following questions with what is known from the systematic review. (For example, it is not useful to report that 5,000 articles were reviewed for an entire guideline because it provides no information on the quantity of studies in the body of evidence for a particular process of care.)

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

The information in the following questions in this section is based on the USPSTF review cited in section 1a.5 unless the ACOG is specifically identified.

Folic acid supplementation in pregnant women was studied for its effect on the occurrence of neural tube defects in newborns.

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

Key Points

- Should include BOTH grade and definition of the grade
- Not all grades are on a letter or number scale
- Grades for quality of evidence are different from grades for the recommendation (although related) – make sure it is the appropriate grade for the quality of the body of evidence

High Certainty of Net Benefit - The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.

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

Levels of Certainty Regarding Net Benefit

Level of	Description

Certainty*	
High	The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.
Moderate	The available evidence is sufficient to determine the effects of the preventive service on health outcomes, but confidence in the estimate is constrained by such factors as: • The number, size, or quality of individual studies. • Inconsistency of findings across individual studies. • Limited generalizability of findings to routine primary care practice. • Lack of coherence in the chain of evidence. As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion.
Low	The available evidence is insufficient to assess effects on health outcomes. Evidence is insufficient because of: • The limited number or size of studies. • Important flaws in study design or methods. • Inconsistency of findings across individual studies. • Gaps in the chain of evidence. • Findings not generalizable to routine primary care practice. • Lack of information on important health outcomes. More information may allow estimation of effects on health outcomes.

* The USPSTF defines certainty as "likelihood that the USPSTF assessment of the net benefit of a preventive service is correct." The net benefit is defined as benefit minus harm of the preventive service as implemented in a general, primary care population. The USPSTF assigns a certainty level based on the nature of the overall evidence available to assess the net benefit of a preventive service.

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: 1992-2009

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*)

Initially 1 large randomized, controlled trial (RCT) for the 1996 review.
The recent evidence synthesis included 4 studies published since 1996: 1 cohort study, 2 case control studies, and 1 meta-anlalysis.

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*)

One cohort study rated as fair quality.
Studies will be graded "fair" if any or all of the following problems occur, without the fatal flaws noted in the "poor" category : Generally comparable groups are assembled initially but some question remains

whether some (although not major) differences occurred with follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention to treat analysis is done for RCTs.

Two case control studies – one rated fair quality and one rated good quality.
Good: Appropriate ascertainment of cases and nonbiased selection of case and control participants; exclusion criteria applied equally to cases and controls; response rate equally to or greater than 80 percent; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to confounding variables.
Fair: Recent, relevant, without major apparent selection or diagnostic work-up bias but with response rates less than 80 percent or attention to some but not all important confounding variables.

One meta-anlaysis rated fair quality.

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*)

The Czeizel cohort study reported that 1 NTD and 9 NTDs occurred in the supplemented and unsupplemented women, respectively, for an adjusted odds ratio (aOR) of 0.11 (95% CI, 0.01–0.91); the odds ratio (OR) was adjusted for birth order, chronic maternal disorders, and history of previous fetal death or congenital abnormality. The meta-analysis also found a protective effect of folic acid-containing multivitamins in NTDs with an OR of 0.67 (95% CI, 0.58–0.77) in case-control studies and an OR of 0.52 (0.39–0.69) in RCTs and cohort studies. Both the Czeizel study and the meta-analysis found a statistically significant association between folic acid supplementation and a reduction in cardiovascular congenital abnormalities. In addition, there was a significant effect of folic acid-containing multivitamin use on congenital limb defects in the meta-analysis. No consistent effect of folic acid-containing multivitamins, either on orofacial clefts or on urinary tract congenital abnormalities, was seen in the Czeizel study or the meta-analysis.

The 1995 case-control study reported an OR of 0.65 (95% CI, 0.45–0.94) for use of folic acid-containing supplements in the 3 months before conception, and an OR of 0.60 (95% CI, 0.46–0.79) for supplement use in the 3 months after conception. The 2003 study by Thompson and colleagues reported an OR of 0.55 (0.25–1.22) for regular use, and an OR of 0.92 (0.55–1.55) for some use of folic acid-containing supplements, but neither of these findings was statistically significant. Several differences in these case-control studies may explain differences in results. The 2003 Thompson study was smaller and adjusted for dietary folate intake. Additionally, the exposure timeframes were different: the Shaw study measured exposure in 2 time frames, 3 months before and 3 months after conception, while the Thompson study combined these same 6 months of periconception time into one measure of exposure.

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

USPSTF:
The recommendation statement concluded: “Adequate evidence suggests that folic acid from supplementation at usual doses is not associated with serious harms. In its current review, the USPSTF found no evidence on drug interactions, allergic reactions, or carcinogenic effects.”

The evidence synthesis found one fair quality retrospective cohort study that addressed whether folic acid supplementation in women of childbearing age increases the risk of harmful outcomes for either the woman or the infant. After adjusting for age and parity, the authors reported an OR of 1.59 (95% CI 1.41–1.78) for twin delivery after preconceptional folic acid supplementation. In a subgroup analysis of women who did not report IVF, the risk of twinning was lower and non-significant (OR 1.13, 95% CI 0.97–1.33), as expected given the increase in multiple gestation associated with IVF and other assisted reproductive technologies. The odds of having twins of unlike sex, an outcome used as a proxy for dizygotic twinning, were increased in women taking folate, (OR 1.43, 95% CI 1.12–1.83). The authors then adjusted for both a 45% underreporting of supplementation as well as an estimated 12.7% of unidentified IVF pregnancies. When the likely underreporting for folic acid use and IVF were accounted for, the OR for twin delivery after preconceptional supplementation fell to 1.02, and was no longer statistically significantly greater than the risk for women who did not take folic acid (95% CI, 0.85–1.24).

ACOG:
Risks of folic acid supplementation. The risks of higher levels of folic acid supplementation are believed to be minimal. Folic acid is considered nontoxic even at very high doses and is rapidly excreted in the urine. There have been concerns that supplemental folic acid could mask the symptoms of pernicious anemia and thus delay treatment. However, folic acid cannot mask the neuropathy typical of this diagnosis. Currently, 12% of patients with pernicious anemia present with neuropathy alone. With folic acid supplementation, this proportion may be increased, but there is no evidence that initiating treatment after the development of a neuropathy results in irreversible damage. A small number of women taking seizure medication (diphenylhydantoin, aminopterin, or carbamazepine) may have lower serum drug levels and experience an associated increase in seizure frequency while taking folic acid supplement. Monitoring drug levels and increasing the dosage as needed may help to avert this complication.

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 <u>each</u> new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

_____

1a.8 OTHER SOURCE OF EVIDENCE

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

1a.8.1 What process was used to identify the evidence?

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