Evidence and Importance to Measure and Report

Evaluation and Measure Submission Guidance

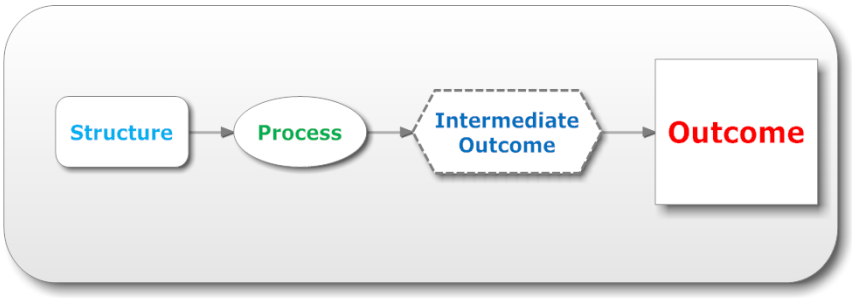
07/10/11

Measure Evaluation Guidance

- Reports on guidance for measure evaluation:
  - Evidence for the Focus of Measurement and Importance to Measure and Report
  - Measure Testing and Scientific Acceptability of Measure Properties
  - Measure Harmonization
- Updated Measure Evaluation Criteria
- Revised Measure Submission Form
  - Most changes related to guidance on evidence (1c)
  - Some changes related to taxonomy (primarily response options, e.g., setting)
  - Some clarification in wording/instructions

Evidence for Measure Focus

NQF National Quality Forum



- Hierarchical preference for
  - Outcomes linked to evidence-based processes/structures
  - Outcomes of substantial importance with plausible process/structure relationships
  - Intermediate outcomes
  - Processes/structures

Most closely linked to outcomes

www.qualityforum.org

1. Impact, Opportunity, Evidence–Importance to Measure & Report

NQF National Quality Forum

Extent to which the specific measure focus is evidence-based, important to making significant gains in health care quality and improving health outcomes for a specific high-impact aspect of healthcare where there is variation in or overall less-than-optimal performance.

a. High impact
b. Gap in performance
c. Evidence supports measure focus

4

www.qualityforum.org

Evidence Guidance: Key Points

- Explicit, transparent information on the quantity, quality, consistency of the body of evidence (not selected individual studies)
- Measure developers can/should use evidence assembled, reviewed, and graded by others
- Preferred grading systems–GRADE, USPSTF
- Rating scale for quantity, quality, consistency
- Expert opinion is not evidence
- Exception for health outcomes
- Does not replace need for expertise and judgment

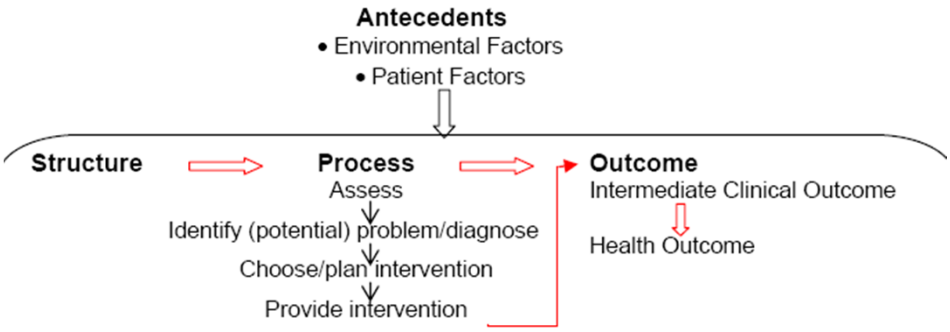
Examples – Evidence Reviews

- AHRQ Evidence-Based Practice Centers
- National Guidelines Clearinghouse
- PubMed – Clinical Queries for systematic reviews
- Cochrane Collaboration Reviews
- Others

Evidence for Measure Focus

NQF National Quality Forum

- What are you measuring?
- What is the evidence about?



7

www.qualityforum.org

Quantity, Quality, Consistency

NQF National Quality Forum

- **Quantity**: Total number of studies (not articles or papers)
- **Quality**: Certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence
- **Consistency**: Stability in both the direction and magnitude of clinically/practically meaningful benefits and harms to patients (benefit over harms) across studies in the body of evidence

8

www.qualityforum.org

Quality

NQF National Quality Forum

- Related to study factors
- Study design (e.g., RCT, non-RCT)or flaws (lack of allocation concealment or blinding; large losses to follow-up; failure to adhere to intention to treat analysis; stopping early for benefit; failure to report important outcomes)
- Directness/indirectness to the specific measure (regarding the population, intervention, comparators, outcomes)
- Imprecision (wide confidence intervals due to few patients or events)

Evidence Rating Scale

NQF National Quality Forum

- Evidence Report, See Table 4, p.20

Definition/ Rating	Quantity of Body of Evidence	Quality of Body of Evidence	Consistency of Results of Body of Evidence
Definition			
High			
Moderate			
Low			
Insufficient to Evaluate			

Evaluation of Subcriterion 1c

NQF National Quality Forum

Quantity of Body of Evidence	Quality of Body of Evidence	Consistency of Results of Body of Evidence	Pass Subcriterion 1c
Moderate-High	Moderate-High	Moderate-High	Yes
Low	Moderate-High	Moderate (if only 1 study, high consistency not possible)	Yes, but only if it is judged that additional research is unlikely to change conclusion that benefits to patients outweigh harms; otherwise, No
Moderate-High	Low	Moderate-High	Yes, but only if it is judged that potential benefits to patients clearly outweigh potential harms; otherwise, No
Low-Mod-High	Low-Mod-High	Low	No
Low	Low	Low	No

Evaluation of Subcriterion 1c

NQF National Quality Forum

Quantity of Body of Evidence	Quality of Body of Evidence	Consistency of Results of Body of Evidence	Pass Subcriterion 1c
Exception to Empirical Body of Evidence for Health Outcome For a health outcome measure: A rationale supports the relationship of the health outcome to at least one healthcare structure, process, intervention, or service			Yes, if it is judged that the rationale supports the relationship of the health outcome to at least one healthcare structure, process, intervention, or service
Potential Exception to Empirical Body of Evidence for Other Types of Measures If there is no empirical evidence, expert opinion is systematically assessed with agreement that the benefits to patients greatly outweigh potential harms.			Yes, but only if it is judged that potential benefits to patients clearly outweigh potential harms; otherwise, No

Generic Rating Scale

- **Used with 1a, 1b**

Rating	Definition
High	Based on the information submitted, there is high confidence (or certainty) that the criterion is met
Moderate	Based on the information submitted, there is moderate confidence (or certainty) that the criterion is met
Low	Based on the information submitted, there is low confidence (or certainty) that the criterion is met
Insufficient	There is insufficient information submitted to evaluate whether the criterion is met (e.g., blank, incomplete, or not relevant, responsive, or specific to the particular question)

Importance to Measure and Report

- Must pass all 3 subcriteria
  - High impact (1a)
  - Opportunity for improvement (1b) **
  - Evidence (1c)
- Insufficient evidence cannot be rated and the measure would not pass 1c or Importance to Measure and Report

** *Measures being reviewed for endorsement maintenance may qualify for reserve status if they address an important aspect of quality but fail to demonstrate a gap in performance and certain other criteria are met. Such measures should be rated on all evaluation criteria.*

NQF
National Quality Forum

Measure Submission: Section 1c, Evidence

Disclaimer

NQF
National Quality Forum

- The following are illustrations of the <u>type of information NQF is seeking</u> on the submission form
  - Not intended as an example for one measure
  - Not intended to represent the only or best approach to measure development and testing
  - Undesirable examples are indicated with an X
- The key points are
  - Provide the information requested
  - Provide substantive information and data in the measure submission form
  - Provide information that demonstrates the criteria are met

1c.1. Structure-Process-Outcome Relationship

NQF National Quality Forum

(Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-outcome; process-health outcome; intermediate clinical outcome-health outcome)

Example 1 – Measure focus is proximal to desired outcome

Measuring provision of pelvic floor muscle training will improve care (Note: Does not provide requested information)

Note: The following provides the requested information

The measure focus is the process of pelvic floor muscle training for urinary incontinence

This process leads to desired outcomes as follows:

Pelvic floor muscle training >> decreased urine leakage

17

www.qualityforum.org

1c.1. Structure-Process-Outcome Relationship

NQF National Quality Forum

(Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-outcome; process-health outcome; intermediate clinical outcome-health outcome)

Example 2 – Measure focus distal to desired outcome

Note: The following provides the requested information but the measure focus is not proximal to the desired outcome as preferred

The measure focus is the process of measuring hemoglobin every month for ESRD dialysis patients

This process leads to desired outcomes as follows:

Measure Hb >> Assess/interpret value >>

Diagnose /Identify problem >> Identify treatment options >>

Administer the appropriate treatment >> Impact on Hb >>

Impact on morbidity/mortality

18

www.qualityforum.org

1c.1. Structure-Process-Outcome Relationship

NQF National Quality Forum

Example 3 – Measure focus is health outcome

The measure focus is the outcome of hospital readmission for AMI patients. It is considered a proxy for the health outcome of deterioration in health status.

Multiple care processes can influence deterioration in health status after discharge and hospital readmission (e.g., appropriate treatment/intervention, meds, clinical stabilization, care coordination /transition).

Comprehensive care transition management/ care coordination can lead to decreased hospital readmissions as described below.

Comprehensive care transition management/ care coordination

Leads to: Early reconnection to primary care; appropriate level of follow-up care; patient understanding of self-monitoring, self-management, & follow-up care

Leads to: Continuity of treatment plan; early identification & intervention for adverse changes

Leads to: Stable/improved health status (fewer readmissions)

19

1c.2. Type of Evidence

NQF National Quality Forum

(Check all that apply)

☒ Clinical Practice Guideline

☐Selected individual studies (rather than entire body of evidence)

☒Systematic review of body of evidence (other than within guideline development)

☐Other 1c.3. Describe

Notes:

1. Evidence is not required for health outcome measures but is desirable when available for a key process or structure
2. Selected individual studies or a list of references from a literature search are not considered systematic reviews of a body of evidence.
3. Items 1c.4 – 1c.14 are required whether evidence is from guideline or some other source of systematic review

20

1c. Evidence

NQF National Quality Forum

1c.4. Directness of Evidence to the Specified Measure

(State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population.) **Example 1 – Direct evidence**

The body of evidence addresses the effect of pelvic floor muscle training in women with urinary incontinence (stress, urge, mixed) in comparison to no treatment or inactive control treatments on primary outcomes of: symptomatic cure reported by the woman, symptoms of cure or improvement reported by the woman, and symptom and condition-specific quality of life (QoL). Secondary outcomes included number of leakage episodes, number of micturitions during the day/ during the night, pad and paper towel testing, measures of pelvic floor muscle contraction

The measure focus is on pelvic muscle training in women with urinary incontinence

1c.5. Quantity of Studies in the Body of Evidence

(Total number of studies, not articles.) **14**

1c. Evidence

NQF National Quality Forum

1c.4. Directness of Evidence to the Specified Measure

(State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population.)

Example 2 – Indirect evidence (Note: Provides the requested information, but not direct evidence for the specified measure)

The body of evidence addresses the relationship between low hemoglobin values and increased mortality in patients with chronic renal disease; and Hb treatment targets and mortality

The measure focus is on the frequency of measuring the hemoglobin level in patients with ESRD on dialysis

1c.5. Quantity of Studies in the Body of Evidence

(Total number of studies, not articles.) **5**

1c.6. Quality of Body of Evidence

(Summarize the certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence resulting from study factors. Please address: a) study design/flaws; b) directness/indirectness of the evidence to this measure (e.g., interventions, comparisons, outcomes assessed, population included in the evidence); and c) imprecision/wide confidence intervals due to few patients or events)

Methodological quality was judged from the trial reports, two of which were only abstracts.

a) Only two trials adequately reported the randomization process so that it was sure there was adequate concealment. Eight of the 14 studies used blinded outcome assessments. The proportion of dropouts was 0-20%. Based on concealment and blinding, 2 trials were at low risk of bias, 6 at moderate risk, and 6 at high risk of bias.

b) the evidence is directly relevant to the focus and target population of the proposed measure – pelvic floor muscle training in women with urinary incontinence.

c) Sample sizes were small to moderate (26-170) in 12 of 14 studies and only 3 reported an a priori power calculation.

1c.7. Consistency of Results across Studies

(Summarize the consistency of the magnitude and direction of the effect across studies)

Of the 14 trials, 12 reported data suitable for analysis. Meta-analysis was not possible due to study heterogeneity.

Comparison of PFMT vs. no treatment, placebo, or control on various outcomes as follows:

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

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)

1c.7. Consistency of Results across Studies cont.

(Summarize the consistency of the magnitude and direction of the effect across studies)

QoL – 2 studies "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)

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

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

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

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)

1c.8. Net Benefit

(Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms across studies)

The estimates of benefit were reported for the various outcomes in 1c.7.

"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)

1c. Evidence

1c.9. Grading of Strength/Quality of the Body of Evidence

Has the body of evidence been graded?

⊙Yes ○No

1c.10. If the body of evidence was graded, *identify the entity that graded the evidence including balance of representation and any disclosures regarding bias*

2 review authors

Chantale Dumoulin, School of Rehabilitation, University of Montreal

Jean Hay-Smith, Dept. of Women and Children's health, Dunedin School of Medicine, Dunedin, New Zealand

Editorial Group – Cochrane Incontinence Group

1c. Evidence

1c.11. System Used for Grading the Body of Evidence

○ USPSTF ○ GRADE ⊙ Other

Note: Do not indicate a system and a grade unless those methods for systematically reviewing and grading the evidence were followed

USPSTF – US Preventive Services Task Force

GRADE – Grades of Recommendation, Assessment Development, and Evaluation Working Group

1c. Evidence

NQF National Quality Forum

1c.12. If other, identify and describe the grading scale with definitions

Cochrane Handbook for Systematic Reviews for Interventions, 5.0.2, updated September 2009 http://www.mrc-bsu.cam.ac.uk/cochrane/handbook502/

Risk of Bias

Low risk: Plausible bias unlikely to seriously alter the results. Low risk of bias for all key domains.

Unclear risk: Plausible bias that raises some doubt about the results. Unclear risk of bias for one or more key domains.

High risk: Plausible bias that seriously weakens confidence in the results. High risk of bias for one or more key domains.

1c.13. Grade Assigned to the Body of Evidence

An overall grade of methodological quality was not assigned. The risk of bias was reported for the individual studies. Based on concealment and blinding, 2 trials were at low risk of bias, 6 at moderate risk, and 6 at high risk of bias.

1c. Evidence

NQF National Quality Forum

1c.14. Summary of Controversy/Contradictory Evidence

No controversy or contradictory evidence reported.

1c.15. Citations for Evidence, other than guidelines which are addressed below

List of references from literature search

(Note: The citations should be provided for the systematic review of the evidence described in items 1c.2-1c.13)

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.

Note: If guideline is used for evidence, the reference should be provided in 1c.17 and 1c.18

1c. Evidence

1c.16. Quote verbatim, the specific guideline recommendation
(Including guideline number and/or page number)

1. Pelvic floor retraining (Kegel) exercises should be recommended for women presenting with stress incontinence.(I-A)

1c.17. Clinical Practice Guideline Citation

Society of Obstetricians and Gynaecologists of Canada, Robert M, Ross S, Farrel SA, Easton WA, Epp A, Girouard L, Gupta C, Lajoie F, Lovatsis D, MacMillan B, Schachter J, Schulz J, Wilkie DH. Conservative management of urinary incontinence. J Obstet Gynaecol Can 2006 Dec;28(12):1113-8.

1c.18. National Guideline Clearinghouse or Other URL

http://www.guideline.gov/content.aspx?id=13390&nbr=006801

1c. Evidence

1c.19. Grading of strength of guideline recommendation

Has the recommendation been graded?

⊙ Yes ○No

1c.20. If the guideline recommendation was graded, identify the entity that graded the recommendation including balance of representation and any disclosures regarding bias.

Society of Obstetricians and Gynaecologists of Canada (SOGC)

No information available on representation and disclosures

1c. Evidence

1c.21. System used for grading the strength of guideline recommendation

○ USPSTF ○ GRADE ⊙ Other

1c.22. If other, identify and describe the grading scale with definitions

Quality of Evidence Assessment

I: Evidence obtained from at least one properly designed randomized controlled trial.

II-1: Evidence obtained from well-designed controlled trials without randomization.

II-2: Evidence obtained from well-designed cohort (prospective or retrospective) or case–control studies, preferably from more than one center or research group.

II-3: Evidence obtained from comparisons between times or places with or without the intervention. Dramatic results from uncontrolled experiments (such as the results of treatment with penicillin in the 1940s) could also be included in this category.

III: Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.

Continue next slide

1c.22 continued

Classification of Recommendations

A. There is good evidence to recommend the clinical preventive action.

B. There is fair evidence to recommend the clinical preventive action.

C. The existing evidence is conflicting and does not allow to make a recommendation for or against use of the clinical preventive action; however, other factors may influence decision-making.

D. There is fair evidence to recommend against the clinical preventive action.

E. There is good evidence to recommend against the clinical preventive action.

I. There is insufficient evidence (in quantity or quality) to make a recommendation; however, other factors may influence decision-making.

1c.23. Grade assigned to the recommendation I-A

1c. Evidence

NQF National Quality Forum

1c.24. Rationale for Using This Guideline Over Others

Two other guidelines have a similar recommendation.

Hartford Institute for Geriatric Nursing (HIGN).

National Collaborating Centre for Women's and Children's Health/National Institute for Health and Clinical Excellence (NCCWCH/NICE).

1c. Evidence

NQF National Quality Forum

1c.25. Based on the NQF descriptions for rating the evidence, what was your assessment of the quantity, quality, and consistency of the body of evidence?

Quantity

⊙ High ○ Moderate ○ Low

1c.26. Quality

○ High ⊙ Moderate ○ Low

1c.27. Consistency

○ High ⊙ Moderate ○ Low